RAF / 87 / P03

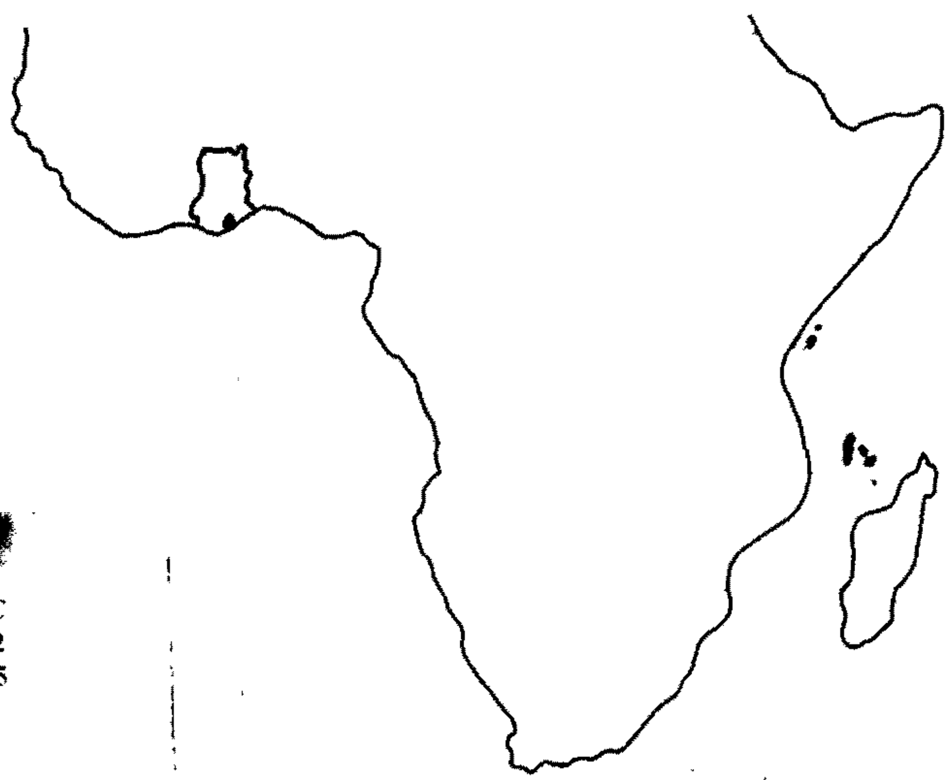
WORKBOOK

On

Demographic Data Evaluation and Analysis

Based on

ECA Sub Regional Training Workshop for Anglophone Countries held at RIPS, Accra, 1 – 19 August 1988



ECONOMIC COMMISSION FOR AFRICA / REGIONAL INSTITUTE FOR POPULATION STUDIES

RAF / 87 / P03

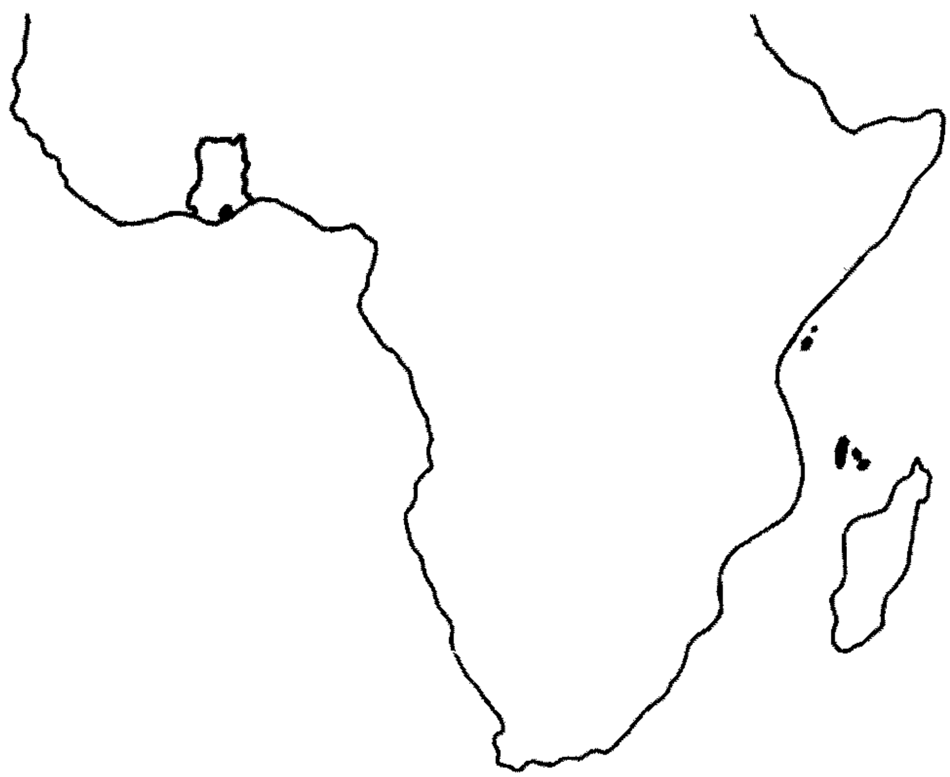
WORKBOOK

On

Demographic Data Evaluation and Analysis

Based on

ECA Sub Regional Training Workshop for Anglophone Countries held at RIPS, Accra, 1 – 19 August 1988



ECONOMIC COMMISSION FOR AFRICA / REGIONAL INSTITUTE FOR POPULATION STUDIES

C O N T E N T S

FOREWORD

Population is at the centre of all planning exercises. It is inconceivable to visualise that in a modern society meaningful developmental activities can be carried out without first considering the population - its size, distribution over space, growth and change over time, in addition to socio-economic characteristics. Thus, demographic data collection, evaluation and analysis have assumed great importance in the day-to-day administration and functioning of societies.

When a large number of countries in Africa attained political independence, the leaders were faced with the problem of having to plan socio-economic development with very little, if at all, of information on their populations. In quite a large number of countries, even the total number of inhabitants was more of a guess than anything based on hard facts. As regards other characteristics like spatial distribution, components of population change and socio-economic attributes, even wild guesses were not possible. Thus, a major priority of African countries was to take stock of their human resources. ECA, with the assistance of the United Nations Population Fund (UNFPA), has been assisting many countries undertake scientific population censuses. Currently, every English speaking African country has undertaken atleast one census, and a large number have two or more post independence censuses.

As planning needs became more and more specific and demanding, it has been felt that mere knowledge of the size of a census-based population is inadequate. Unfortunately, since other systems of data collection like civil registration and vital statistics are incomplete and deficient in most countries, much reliance is placed on censuses and sample surveys. For efficient planning, it is necessary to have information at very detailed geographic areal levels which can only be obtained by censuses. A census however, is a major operation involving the entire population of a country in one way or the other and involves several steps and stages in its execution. With very little experience in census taking, with populations unexposed to statistical enquiries , with a large segment of the population illiterate and non numerate, it is not surprising that in many countries the quality of the information collected leaves much to be desired. Hence, there is always a need to evaluate the coverage, completeness, and general reliability of data from these sources.

To study population dynamics, the census has become the major, if not the only source of information; thus several types of data must necessarily be collected to understand the phenomenon of population growth and change. Indirect methods of inferring fertility, mortality and migration are resorted to in view of the failure of direct questions to elicit accurate information. As such, analytical techniques have become the most efficient and common tool used by statisticians and demographers for deriving reliable information from the often defective and incomplete data collected.

The Kilimanjaro Programme of Action on Population (KPA) adopted by African countries in 1984 stressed that population should be a key factor in the formulation of development strategies and plans and that population and development are interrelated. Accordingly, it recommends that national efforts to create greater awareness of the interrelationship between population and development should include the provision of more information on the dynamics of population change and the impact of such change on current and future development.

Many African countries have embarked on structural adjustment and economic recovery programmes. With the best of intentions, all these exercises would become futile if policy makers do not have the means of utilising available data to demonstrate the impact of rapid population growth on development possibilities.

Lack of trained manpower is also one of the serious handicaps faced by African countries; it has hampered the analysis and utilisation of demographic and socio-economic data. This, to a certain extent, has been ameliorated by the establishment of regional demographic training institutes and through inservice or on-the-job training courses, middle level training programmes and other related activities. In spite of these, it is quite common for much of the data so assiduously collected not to be promptly analysed, evaluated, interpreted and used.

Training courses often cater for only a small number of persons. Moreover, even those who have participated in short term courses need refresher courses or access to training materials and manuals in order to apply the knowledge acquired effectively. Even those who are not fortunate to have had such training can benifit if suitable manuals or work books are made available. It is in the light of this that it was felt that the training materials used at the recently concluded Sub regional Training Workshop on Demographic Data Evaluation and Analysis at the Regional Institute for Population Studies (RIPS) Accra, Ghana would serve such purposes.

This workshop was one of the training exercises envisaged under the Canadian-funded UNFPA Project for Sub Saharan Africa for1988-92. A similar workshop is to be organized for French-speaking African countries in 1989. Although the RIPS workshop dealt only with English-speaking countries, the experiences gained will also be useful to French-speaking and other countries in the region. At the same time, the deliberations of the planned workshop for French-speaking countries would also be beneficial to English-speaking countries. This workbook may then be suitably adapted to include experiences of all countries in the continent.

Lecturers from ECA, RIPS and the UN Population Programmes and Projects Branch (PPPB) of the Department of Technical Co-operation for Development (DTCD) participated in the training workshop and prepared training materials. A committee of representatives from ECA and RIPS was responsible for editing and adapting the materials for publication as a workbook. It is my hope that the workshop will assist Member States in exploiting and utilising their demographic and related data more fully.

Adebayo Adedeji,

United Nations Under-Secretary General
and
Executive Secretary
Economic Commission for Africa.

CHAPTER 1

USES OF DEMOGRAPHIC DATA IN PLANNING

1.1 Introduction

1. Development planning is the complex process through which governments lay down procedures for promoting the economic objective of growth in the broad macro-economic sense of augmenting national production capacity on the one hand, and the social objective of ensuring equitable distribution of income among the population and also facilitating the development of their potentials on the other. The process entails the essential stages of formulation, implementation, monitoring, evaluation and modification of policies, strategies, programmes and projects for local, national, sub-regional and regional development. Accordingly, development planning has a human, a temporal as well as a spatial dimension.

2. Each stage and dimension of the planning process calls for the use of demographic data. Thus initially, population data are useful for making feasibility or preliminary assessment of the situation at any level (local, national, rural, urban etc.). They are useful for implementing development goals in the sense of supporting the translation of strategies into specific policies, programmes and projects. In planning, demographic data can be used explicity (directly) or implicitly (indirectly) for the achievement of short-, medium- or long-term objectives. Whatever the time frame, it is now becoming increasingly the practice to create an institutional framework for integrating demographic data in planning.

1.2 Need for Demographic Data in Social and Economic Planning

3. A salient feature of the current literature on development planning in Africa is the growing recognition of the cardinal role of population as a factor in determining the level, pattern and rate of development. In this respect, what is most significant is the recognition accorded to the interrelationships between population variables and other economic, social and cultural factors. This growing recognition is, after all, germane, if the avowed objectives of a man-centred development planning are to be attained. A worthy and desirable goal of all economic and social development is the improvement of the standard of living of the population. It is, therefore, crucial that development planning should take into account the nature, evolution and characteristics of the population to be catered for in the plan period. At any level of socio-economic planning, there is what can be legitimately regarded as the target population, the size and composition of which varies according to the focus of the plan and the nature of the services to be provided.

4. Furthermore, economic-demographic consideration should of necessity transcend mere examination of current conditions and also deal with future trends and development in accordance with national goals, policies and objectives. In this regard, the integration of the various sectors of the national economy is very important. Recognizing the significance of integration and harmonization of all development policies, Africa's strategy for development in the 1970s pertinently states that, "A strategy for the development and utilization of human resources must dovetail with other policies, plans and strategies for overall economic and social development. It must, in particular, be closely related to national policies on population employment and income distribution." (ECA, 1973) The more recent strategy enunciated in the Lagos Plan of Action emphasizes the need for better collaboration between demographers, planners and statisticians in their work at the national level.(OAU, 1981). The importance of population variables in development planning has been further articulated in the Kilimanjaro Plan of Action for African Population and Self-Reliant Development (KPA) as well as the United Nations Programme of Action for Economic Recovery and Development (UNPAERD).

1.3 Nature, Quality and Gaps in Data Supply

5. Population variables are both determinants and consequences of the development process. On the one hand, when viewed as consumers and producers, in interaction with other resources, population can be a means of achieving planned development. On the other hand, population changes occur partly as a result of economic and social development. Thus, population serves the dual function of being both desirable ends and means for further development.

6. Decisions as to the type of demographic data needed for development planning should be based on the relevance of data to the aims and objectives to be attained. Selected data should be geared to the articulation and achievement of the aims, objectives and targets of the plan in both the short- and long-term perspectives; they should be such as could support and promote the realization of the economic, social and political goals set in conformity with the national aspirations for the improvement of the quality of life and living conditions and for the promotion of balanced development, use and distribution of resources. In this connection, importance of the link between demographic factors and production and utilization particularly agricultural output and the general supply of services should be emphasized.

7. Demographic variables, like many social phenomena, are not easy to measure and control. This is mainly because of their continuous change and statistical fluctuations arising from the unpredictability of human behaviour and the varied extraneous factors (physical, economic, social and cultural)

which influence and often determine the levels, patterns and forms of demographic development and processes. Furthermore, ease of measurement is hampered by the complex interrelationships among demographic variables which are also intricately related to other non-demographic factors. This situation normally complicates the difficulty with which planners and demographers are able to convince others of the importance of programmes aimed at achieving certain demographic targets.

8. Apart from censuses, other important source of demographic and socio-economic data for planning includes vital registration systems, (births, deaths, and marriages) which in Africa are incomplete and biased. Sample surveys (demographic, social and economic), labour statistics, international migration records as well as data on school enrolment, are yet other sources.

9. But although much remains to be done, there is evidence that African countries have made some progress with regard to data collection. An ECA review shows clearly that since the end of World War II, more and more countries have conducted censuses.(ECA, 1982). But in spite of these growing use of censuses, its frequency has been very irregular in the majority of countries when compared to the norm of holding decennial censuses, for example.

10. Closely related to the irregularity of data supply (census, surveys, vital registration etc.), is the general problem of the scarcity and poor quality of demographic data. Even with the already mentioned progress in establishing and promoting census and survey programmes, notably under the African census and survey programmes of the United Nations Economic Commission for Africa, the gap, quantitatively and qualitatively, remains very wide. The coverage of demographic variables during enumerations has been inadequate and, more often than not, collected data in the region have content as well as enumeration errors which mar their quality and reliability for planning purposes.

11. The inadequacies of the data call for skill and expertise in making sure that, as much as possible, the inaccuracies are reduced to a minimum without necessarily distorting the basic picture. Already, there are a number of tools that could be employed in dealing with incomplete and inadequate data.(Brass et al, 1968, Brass, 1975, Som, 1971, UN, 1967 and UN, 1983) Some of these have been widely used by planners and researchers with some effect and success. However, the use to which the techniques can be put in order to improve the data base should be made with a great deal of circumspection. Good and balanced judgement should be exercised to ensure that the use of corrective techniques are not abused and that the adjustments, contrary to prevailing cultural, historical, social and economic conditions, do not just mechanically smoothen away

real and relevant irregularities that may be more revealing of other significant demographic behaviour patterns. For instance, where migrations have occurred on a large scale and for some time in the past, the application of stable population models to observed age data can lead to some misleading inferences and estimates. Therefore, bad use of the techniques can impose a greater blight than the originally observed errors.

1.4 Complementarities in the supply and use of Data

12. In providing demographic data for planning, it is vital to bear in mind the complementarities in the process. As well as the complementarity between planned projects and needed resources in terms of capital, manpower, raw materials, technology and many others, there is that between a comprehensive, integrated planning for economic and social development at the national and subnational levels and that between the macro and sectoral levels of planning. A consideration of these provides a frame of analysis which moves from the general to the particular and which promotes comprehensive and detailed planning that caters for the interest of all sections of the community, especially the under-privileged.

13. Subnational and sectoral planning should be undertaken on the premise that generally, subnational units and sectors of the economy cannot be planned independently. For example, planning for rural development cannot be done apart from that for the urban area. Also educational objectives cannot be divorced from the manpower targets of a country just as health planning cannot overlook agricultural planning for more food and better nutrition. Because of the specific nature and focus of sectoral planning, it requires increasingly more refined population data and other statistics.

14. Sectoral planning should generally cover all areas of social policies and programmes as well as human resources utilization. Ideally therefore, subjects such as education, health, food supply and nutrition, housing and household, manpower and employment should be planned individually or in related combination. In particular, sectoral services and needs of the economy have to be assessed for population at risk. Long-term planning calls for the vigilance of planners and demographers in making periodic reviews and adjustments of the relationship between demographic and other variables. This exercise should form an integral part of the planning programme. It should be done in order to take account of periodic fluctuations in data and to be able to incorporate new and better data as they become available.

1.5 A Minimum Data Supply Base and Use

15. Within flexible limits, it is desirable to maintain a minimum data supply for social and economic planning. As can be seen from Table 1.1 data should be provided for planning the

Table 1.1:

Source and Use of Data for Planning Some Basic Needs

Basic Need/ Service	Demographic and related Data Required	Use of Data	Source of Data
1. Food and Agriculture	Size and growth of total population by age and sex; agricultural population by age and sex; economically active population.	Determination of food demand and supply; setting food targets.	Census and surveys; agricultural census
	Vital rates (births, deaths, migration); rate of growth of population.	Projection of food demand and supply; setting of food targets.	Vital Registration; census;
	Agricultural labour force; employment in agriculture	Supply of manpower; determination of output.	Agricultural census, manpower surveys
	Population settlement and pattern of distribution of land and availability.	Determination of optimal land use pattern; agricultural labour migration	Census and surveys; special land use survey
	Migration volume and direction	Labour supply, demand and movement	Census and surveys; agricultural Census; Demographic Survey, etc.
	Total and per caput food intake	Production targets	Census and surveys; agricultural survey; nutrition survey
	Employment pattern (seasonality, part-time, under employment) etc.	Labour supply, use and productivity	Agricultural and employment survey
	Quantity of food produced	Monitoring and evaluation of production targets.	Agricultural census and survey
	Infrastructure: roads, markets, storage, etc.	Adequacy of support services	Agricultural census and survey.

Table 1.1 (contd.)

Basic Need/ Service	Demographic and related Data Required	Use of Data	Source of Data
2. Education	Size and growth of total population by age and sex by locality and rural-urban residence	Determination of target population; level and pattern of demand for education by cycle	Census, education survey
	Geographic/administrative distribution of population by age and sex	Optimal allocation of education facilities i.e. in areas of greatest need.	Census, education survey
	Vital rates (births, deaths, migration)	Ditto	Census, vital registration, Demographic survey
	Size and growth of school age population and school enrolment by age and sex for first, second and third cycle educational by locality levels.	Ditto	Ditto
	Level of literacy by age and sex and by rural-urban locality	Determination of met and unmet literacy needs	Census, demographic Survey
	Availability of schools and places (primary, secondary, university, technical, vocational) by locality by rural-urban residence, and administrative divisions	Determination of met and unmet needs for educational places and facilities by locality and administrative regions	Educational service records and statistics
	Manpower	Training and budget allocation	Census and educational surveys, rosters of teachers, administrative records, labour force and manpower surveys, plans and policies.

Table 1.1 (contd.)

Basic Need/ Service	Demographic and related Data Required	Use of Data	Source of Data
3. Manpower/ Employment	Size and growth of total population by age and sex, locality and rural-urban residence	Determination of size, growth and distribution of labour force	Census, manpower survey
	Vital Rates (births, deaths, and migration)	Determination of size and projection of labour force	Census, vital registration, sample surveys
	Number of jobs created or to be created by occupational categories, residence, socio-economic groups, enterprises, formal and informal sectors by locality	Potential employment opportunities	Manpower surveys
	Total number of students enrolled in formal education by locality, areas of specialization, etc. and output from educational institutions (higher, technical, vocational).	Determination of enrolment rates and entrants into the labour market	Education survey, census and administrative records
	Level and rate of employment by locality, age and sex, education	Appraisal of the status quo.	Manpower surveys
	Average household income by age and sex, locality, and type of employment.	Study of prevailing pattern of household income distribution	Social and economic surveys to measure levels of living
	Number of planned capital accummulation projects by locality, and rural-urban residence	Potential employment opportunities.	Development plans; Administrative reports

Table 1.1 (cont.)

Basic Need/ Service	Demographic and related Data Required	Use of Data	Source of Data
4. Health	Size and growth of total population by age and sex, locality, rural-urban residence, etc.	Planning of supply targets (hospitals, clinics, doctors, beds, etc.) by locality	Census
	Geographical/administrative distribution of population by age and sex;	Balanced distribution of services	Census
	Vital rates (births, deaths, still births, maternal mortality; infant and childhood mortality etc;)	Projections of future needs; age and sex specific analysis of mortality	Vital registration, hospital records, demographic surveys, and census
	Age-sex specific morbidity data (maternal, infant, childhood, adult, and old age morbidity)	Disease pattern and incidence	Hospital records and statistics, special surveys
	Inventory of health problems, personnel, services/infrastructure/resources/facilities by locality, administrative divisions and accessibility	Planning present and future supply and location of services	Administrative health status report; special surveys on health facilities and personnel
	Inventory of disadvantaged groups (mothers, nomads, rural and urban poor, children, disabled, etc.)	Problem-targeted health and welfare planning	Socio-economic and health surveys
	Inventory of health, nutrition and environmental conditions by locality, administrative divisions etc.	Planning for social and community health and welfare	Administrative Health reports and records

Table 1.1 (cont.)

Basic Need/ Service	Demographic and related Data Required	Use of Data	Source of Data
5. Housing and Shelter	Size and growth of total population by locality, rural/urban residence, age and sex	House construction, demand and target setting by locality	Census
	Size and composition of households by age and sex	Design and planning of houses	Census, demographic surveys
	Household headship rates by age and sex	Demand and supply of new houses; Household income and expenditure	Housing surveys, demographic surveys
	Population/households with services, facilities	Provision of amenities	Census, housing surveys and house listing operations
	Housing stock and future requirements i.e. types and conditions of available housing units	Replacement and renewal of old houses and construction of new ones	Administrative records; registered annual number of building plans etc. Census, house listing operations
	Annual number of houses constructed with basic facilities (water, toilet, baths, etc.)	Ditto	Ditto

Table 1.1 (cont.)

Basic Need/ Service	Demographic and related Data Required	Use of Data	Source of Data
6. Transport and Communication	Size and growth of total population by age and sex, locality, rural/urban residence administrative units, etc.	Determination of volume of services required by locality, etc.	Census
	Population that use transport and communications services;	Determination of volume and supply of service	Administrative records on transport, communications
	Population that does not have access to transport and communication services	Planning location of new services	Special transport and communication surveys
	Households that have radios, T.V.s, telephones, etc.	Determination of demand and supply of services	Special surveys, census
	Localities served by post offices, postal agencies, etc., distance to the nearest service	Ditto	Administrative records special surveys; census

essential basic needs of the population. In the first place, each country should be able to obtain, by complete enumeration at a census, data on the total population by localities especially by urban and rural residence. If conducted on a regular basis, this will facilitate the calculation of the rate of growth of the national as well as rural-urban populations and planning for their development needs. The collection of age and sex data at this stage is highly recommended in the interest of enriching the data base for socio-economic analysis and planning and for projections.

16. To cater for qualitative changes and the dynamics and components of the reported population growth, a minimum attempt, given the scarcity of resources, should be made to collect data on the characteristics of the population on a sample basis. Such characteristics include sex and age, marital status, literacy and educational attainment, economic activity, fertility, and mortality. In this regard, it is significant that the Strategy for the African Region in the International Development Strategy for the United Nations Third Development Decade called for greater account to be taken of the high rate of population growth, uncontrolled urbanization, high levels of mortality and fertility, the gap in meeting the needs of large groups for health, education, shelter, food, employment, etc., and greater opportunities for the advancement of women and the young.(ECA. 1979)

17. Now, each of the suggested characteristics can be put specifically at addressing the above issues raised in the Strategy and indeed to several other uses. The sex-age composition which is the product of past fertility, mortality and migration trends influences very much the trend of the increase of the population. It also influences the size of the school age population, the labour force and female population in the reproductive ages. Also, the sex-age structure is vital for making component and sectoral projections. For many other variables, age and sex composition provides a good basis for in-depth analysis through cross-tabulation. For example, tabulation of marital status by age and sex permits an objective study of marriage in relation to family formation and establishment of new households. Generally, studies of trends in marital status are important for understanding fertility trends and variations.

18. In order to remedy effectively the high illiteracy situation in many African countries and in order to assess the size of available manpower, data on literacy and educational attainment are necessary. In the long-run, such data are important for assessing and evaluating the success and effectiveness of school systems and programmes within regions and localities.

19. At all stages of the educational pyramid, planning which should integrate successive stages, should be made. This involves projections of attendances at all educational levels and estimating the needs of the system in terms of the supply of teachers, equipment and the construction of new schools, colleges and universities. The projections and estimates require to be fed with data on enrolment at successive stages taking account of drop-out and wastage rates. Because educational problems differ between sub-regions of the countries and in particular between rural and urban areas, projections of school age population and attendance should be prepared separately for these areas. For these sub-regional and local planning, population data are needed for each administrative district or local area, depending on the case in question.

20. Collection and analysis of data on economic activity have, in the past, been difficult as shown by attempts to apply international classifications of the active population and of occupations. Nevertheless, it seems that with adequate care, the determination of the size of the economically active population is valuable for measuring participation rates by age, the size and composition of available manpower and for projecting it into the future. Furthermore, data on employment status, occupation and industry are valuable assets for assessing the utilization of manpower resources and for formulating future policies and programmes for their development.

21. Population projections are considered one of the most valuable instruments which demography contributes to social and economic planning in terms of programming for the future. In addition to the future figures for the total population, it is necessary also to consider figures for the country's administrative divisions, localities of various sizes, and urban and rural sectors. In the African experience, decisions regarding the projection period and the relative accuracy required are heavily influenced by the inadequacy of demographic data and knowledge basic for the projections. There is no doubt that the enlightened guesses about the future courses of fertility, mortality and migration requires solid data bases. In this regard, it should be stated that vital rates estimated from limited data can only be used as stop-gaps.

22. Given the gross inadequacy of facilities, health planning in various countries of the region should commence with an inventory of resources. Demographic data needed for the planning of all types of health services so that populations at risk can be estimated both for present and future needs include total population numbers, age-sex structures and geographic distribution of population. These are also needed for the

planning of general medical services, and targets for the number of doctors, health personnel, hospital beds and other facilities usually expressed per 1,000 population. The planning of maternal and child health care services need accurate data on fertility, still-births, maternal mortality, infant mortality, and mortality rates for children up to the ages included in the service, and also data on morbidity of mothers and children, as much as is available. The age and sex structures of populations are also important because a number of diseases are to some extent age-group or sex-group specific.

23. Because reasonable access to basic health services is essential, all relevant demographic data should be available and projections made for administrative units. Urban-rural classifications are also important, mainly on account of very great differences in population size and density, calling for special types of rural health services. Adequate provision of water supply and sewerage to meet demand also depend on population densities.

24. It should also be noted that demographic data in the health sector are required not only for planning but also for evaluation of the progress made in implementation since some demographic characteristics serve as major indicators of health levels.

25. There are clear indications that future demand for food in Africa will be determined more by population than income growth. The type of data collected and used in analysing the results of agricultural census include variables such as size and composition of urban and rural populations, agricultural and farm populations, the economically active population, employment in agriculture, and related problems of seasonality of activity, part-time employment and underemployment. These and some others not listed provide some indications of the interest of agricultural planners in demo-graphic measures.

26. A useful first-step would be to obtain measures of the total size of the population and its rate of growth. On a macro level, the rate of increase in the supply of food should, resources permitting, be high enough to sustain the rate of population increase. Estimates of <u>per caput</u> food intake should use the proper total population base projected objectively into the future. Both base and projected estimates of requirements should also take account of the age and sex structure. Apparently, the amount and type of food consumed vary with age. Thus, it is obvious that the needs of children are not generally the same as those for working age and aged persons. In terms of food production, children and very old persons contribute little or nothing and yet their consumption needs have to be met.

27. Still on food production, data on the size and growth of the agricultural population in relation to land use and arable population densities should be made available. An important demographic factor which affects the size of the agricultural labour force is migration to non-agricultural job centres in towns, plantations and mines. It would, therefore, be necessary to have data on the volume and direction of migration flows and assess their impact on the agricultural labour force. Because of the close relationship between the size of the agricultural labour force and arable population densities, data on subregional and local population characteristics and distribution should be handy. The tendency is for people on the land to move from rural areas of high arable densities to those with low densities. Data on the pattern of population distribution will therefore be useful in planning future expansion of land use and development of resettlement schemes for co-operative farming.

28. Actually, the relevance of rural-urban population distribution and characteristics cannot be under-estimated. Food production, after all, is carried out in the rural areas where the agricultural labour force dwells. The towns produce very little or no food for its inhabitants who generally depend on the supply from the rural areas. Knowledge of the relative size, growth and characteristics of the population in the rural area would therefore, be essential in forecasting the level of food demand and supply.

29. The reasons for the importance of long-term planning in the education and health sector apply to food with even greater force. If food production is to increase faster than the growth of population so that increasing future demands would be met, then long-term estimates based on reliable projections would be required. The future needs, given other constraints, should involve making projections of the total population and especially the population dependent on agriculture and the agricultural labour force, all of which affect the projected food demands and supplies. The projections, no doubt, would also use age and sex data and also consider regional as well as rural-urban differences in population structure and distribution.

REFERENCES

Brass W. et al,
The demography of tropical Africa, Princeton University Press, 1968.

Brass W.
Methods of estimating fertility and mortality from limited and defective data, Laboratories for population statistics, An occasional publication, University of North Carolina, October, 1975.

Organisation of African Unity
Lagos Plan of Action for the economic development of Africa, 1980-2000, International Labour Institute, Geneva 1981.

Som R.K.,
Recall lapse in demographic enquiries, Asia Publishing house, Bombay, 1971.

United Nations (ECA)
Africas strategy for development in the 1970's, Addis Ababa, 1973.

United Nations (ECA)
Strategy for the African region in the international development strategy for the United Nations Third Development Decade, E/CN.14/INF/107/Ref 1, Addis Ababa, 1979.

United Nations (ECA)
Conference of planners, statisticians and demographers, ST/ECA/PSD.2/28, Addis Ababa 1982.

United Nations
Methods of estimating basic demographic measures from incomplete data, Population studies 42, New York, 1967.

United Nations
Indirect techniques for demographic estimation, Population studies 81, New York, 1983.

CHAPTER 2

DATA EVALUATION AND ADJUSTMENT

2.1 Data Collection and Collation

1. Demographic data and estimates, together with related socio-economic statistics, are the essential pre requisites for effective and objective monitoring and evaluation of population trends and development at various spatial levels. Thus such data and their analysis have been recognised as essential in the day to day administration and functioning of modern societies.

2. In the 1960's, the Economic and Social Council of the UN requested the Secretary General to develop proposals for the intensification of action in the field of economic and social development with particular reference to the need to review facilities for the collection, collation, analysis and dissemination of statistical and other information required for charting economic and social development and for providing a constant measurement of progress towards the objectives of the Development Decade.

3. Resolution 1054 B (XXXIX) of 16 july 1965 recommended that the member States of UN undertake to carry out population and housing censuses during the period 1965-1974 and that they take into account the international recommendations in order that the censuses may meet national requirements and facilitate the study of population and housing problems on world-wide basis.

4. For the 1980 round, the Statistical Commission accepted the proposal that the Regional Commission should take the lead in developing recommendations particularly with respect to the topics to be covered and the tabulations to be prepared, in order that regional needs could be met more effectively.

5. Considering the fact that most of the countries in Africa emerged from colonial rule to national independence only in the 1960's and the hard reality that at independence they were faced with the necessity for planning for economic and social development of their population with not much information on even the total size of the population let alone information on spatial distribution, dynamics and characteristics, it is but natural that the first priority of the nascent nations was the collection of as much data as possible on the population and as expeditiously as feasible.

6. With statistical institutions not developed and the population not much exposed to statistical enquiries, it was an arduous task to conduct data collecton exercise. The census became the only source for data in the absence of other systems of data collection. This necessitated the loading of censuses with all types of items including stock and flow data. With no adequate trained manpower, with a population not exposed to reporting on various demographic and socio-economic aspects, it was but to be expected that the data collected will not be

without errors, biases and other deficiencies. Hence it was imperative that the data collected should be evaluated and, if necessary, adjusted.

7. The collection of data being mainly focussed on its eventual use in planning and policy formulation, required to be analysed, interpreted and disseminated. Thus the importance of fuller utilisation of information collected was kept in focus even though the experience has not been commensurate with the requirements.

2.2 Types of data collected and collated

8. The UN and ECA recommendations on the topics to be covered and the list of tabulations are given below both for 1970 and 1980 rounds. Also given are illustrative lists of topics covered by Anglophone countries in the region in their 1970 and 1980 rounds and tabulations prepared in a few selected countries to highlight the situations in the region as to demographic data collection and collation.

List of topics recommended by UN

1970 round

I. Geographic and migration characteristics

(a) Place where found at time of census, (b) Place of usual residence, (c) Place of birth, (d) Duration of residence, (e) Place of previous residence, (f) Place of work.

II. Demographic and social characteristics

(g) Sex, (h) Age, (i) Marital status (j) Citizenship, (k) Nationality/ethnic/tribe.

III. Fertility and mortality

(l) Children born alive, (m) Children living, (n) Age at marriage, (o) Duration of marriage, (p) Marriage order.

IV. Household characteristics

(q) Relationship.

9. Among these 17 topics, the items d,e,f,j,k,n,o and p were only suggested as useful topics but others were recommended as priority items. African countries adopted a very similar set of topics but excluded items e,g,j and p. However they added five additional items as follows:
(i) Births during last 12 months, (ii) Deaths during last 12 months, (iii) Survival of parents, (iv) Form of marriage, and (v) Number of spouses.

10. The African recommendations included items a to c, f,h,i, k to m and q among priority list and items n and o as useful categories for collection. Among the additional items, numbers i to iii were priority and iv and v as useful items.

1980 round

11. The UN recommendations were given in two lists -A and B - depicting respectively priority and useful topics. Compared with 1970, two topics viz., f and p were deleted but: place of residence at a specified date in past; Live births within the 12 months preceding the census; and Deaths of infants within the 12 months preceding the census and maternal orphanhood, were added.

12. The African recommendations did not include topics e,j and p. However, several other topics given below were included: (i) Survival of parents, (ii) Form of marriage, (iii) Number of spouses, (iv) Deaths during last twelve months, (v) Births during last 12 months, and (vi) Deaths of infants born during the last 12 months.

13. In 1980 round, the UN recommendations included items a to e, g to j, l, m and q as also place of residence at a specified time in past among priority listing. All other items were considered as useful.

14. The African recommedations included items a to d, f to i, k, m and q as also topics (i), (iv), (v) and (vi) among priority listing and others as useful.

15. Thus we note that the African countries in their regional recommendations reflected the dearth of data in the continent and the quest for filling up gaps in information. As evidence of their genuine desire to collect a wide variety of data on demographic, social and geographic topics, Table 2.1 gives the types of data collected during the 1970 and 1980 rounds on these aspects in the 18 Anglophone countries. Not only has there been an increase in the number of topics included in the censuses, it has also been noted that the depth and breadth of the scope of the topics were also widened from 1970 experiences and the ever increasing needs for data for planning and policy making. It is also noted from Table 2.2 that at least the quality of age-sex reporting as measured by digit preference indices and joint scores have certainly improved.

Table 2.1 TOPICS CANVASSED DURING 1970 AND 1980 ROUND CENSUSES IN ANGLOPHONE COUNTRIES

COUNTRY	1970/1980 round census years	Relationship	Marital Status	Age at marriage /duration of marriage	No of wives	Children ever born	Children living/ dead	Most recent birth/ birth last year	Survival of recent/ last year birth
Botswana	1971	X	X			X	X	X	X
	1981	X	X			X	X	X	
Ethiopia	-								
	1984	X	X			X	X	X	X
Gambia	1973	X				X	X	X	X
	1983	X	X			X	X	X	X
Ghana	1970	X							
	1984	X							
Kenya	1969	X	X			X	X	X	
	1979	X	X			X	X	X	
Lesotho	1966	X	X						
	1976	X	X			X	X	X	X
Liberia	1974	X	X			X	X	X	
	1984	X	X			X		X	
Libya	1973	X	X			X	X	X	
	1984	X	X			X	X		
Malawi	1966	X							
	1977	X	X		X	X	X	X	

Table 2.1 (CONT.)

TOPICS CANVASSED DURING 1970 AND 1980 ROUND CENSUSES IN ANGLOPHONE COUNTRIES

COUNTRY	1970/1980 round census years	Survival of parent	Survival of spouse	Deaths in household	Disability/ physical status	Resident status	Place of birth	Ethnic/ tribe/ nationality	Citizenship
Botswana	1971	X						X	X
	1981			X					X
Ethiopia	-								
	1984			X	X	X	X	X	
Gambia	1973	X					X	X	
	1983	X				X	X	X	
Ghana	1970						X	X	
	1984					X	X		X
Kenya	1969	X					X	X	
	1979	X					X	X	
Lesotho	1966							X	
	1976			X		X	X	X	X
Liberia	1974			X			X	X	X
	1984	X		X			X	X	X
Libya	1973			X	X	X	X	X	X
	1984			X	X		X	X	
Malawi	1966						X	X	
	1977	X		X		X	X	X	

Table 2.1 (CONT.)

TOPICS CANVASSED DURING 1970 AND 1980 ROUND CENSUSES
IN ANGLOPHONE COUNTRIES

COUNTRY	1970/1980 round census years	Place of residence	Usual residence	Previous residence	Duration of residence	Residence at a specfied time in past	Place of work/ study	Members outside country
Botswana	1971							X
	1981		X		X	X		X
Ethiopia	-							
	1984				X			
Gambia	1973							
	1983							
Ghana	1970							
	1984	X	X					
Kenya	1969							
	1979			X				
Lesotho	1966							
	1976	X		X	X			
Liberia	1974				X			
	1984				X			
Libya	1973			X	X			
	1984			X	X	X		X
Malawi	1966							X
	1977							

Table 2.1

TOPICS CANVASSED DURING 1970 AND 1980 ROUND CENSUSES IN ANGLOPHONE COUNTRIES

COUNTRY	1970/1980 round census years	Relationship	Marital Status	Age at marriage /duration of marriage	No of wives	Children ever born	Children living/ dead	Most recent birth/ birth last year	Survival of recent/ last year birth
Mauritius	1972	X	X	X		X	X	X	X
	1983	X	X	X		X	X	X	X
Nigeria	1973		X						
	-								
S/Leone	1974	X				X	X	X	X
	-								
Sudan	1973	X	X			X	X	X	X
	1983	X	X	X		X	X	X	X
Swaziland	1966					X	X	X	
	1976	X				X	X	X	X
Tanzania	1967					X	X	X	
	1978	X	X			X	X	X	
Uganda	1969					X	X	X	
	1980	X	X						
Zambia	1969	X	X			X	X	X	
	1980	X	X	X		X	X	X	X
Zimbabwe	1969	X	X						
	1982	X	X			X	X	X	X

Table 2.1 (CONT.)

TOPICS CANVASSED DURING 1970 AND 1980 ROUND CENSUSES IN ANGLOPHONE COUNTRIES

COUNTRY	1970/1980 round census years	Survival of parent	Survival of spouse	Deaths in household	Disability/ physical status	Resident status	Place of birth	Ethnic/ tribe/ nationality	Citizenship
Mauritius	1972						X	X	
	1983							X	
Nigeria	1973						X	X	
	-								
S/Leone	1974	X					X	X	
	-								
Sudan	1973	X					X	X	
	1983	X					X	X	
Swaziland	1966					X	X	X	
	1976	X					X	X	
Tanzania	1967			X			X	X	X
	1978	X	X				X		
Uganda	1969	X					X		
	1980					X	X	X	
Zambia	1969			X	X		X	X	X
	1980				X		X	X	X
Zimbabwe	1969						X	X	X
	1982	X					X	X	

Table 2.1 (CONT.)

TOPICS CANVASSED DURING 1970 AND 1980 ROUND CENSUSES
IN ANGLOPHONE COUNTRIES

COUNTRY	1970/1980 round census years	Place of residence	Usual residence	Previous residence	Duration of residence	Residence at a specfied time in past	Place of work/ study	Members outside country
Mauritius	1972				X			
	1983		X			X	X	
Nigeria	1973							
	-							
S/Leone	1974				X			
	-							
Sudan	1973							
	1983		X		X			
Swaziland	1966							
	1976							X
Tanzania	1967							X
	1978					X		
Uganda	1969							
	1980							
Zambia	1969					X		
	1980					X		
Zimbabwe	1969		X		X		X	X
	1982			X				

Table 2.2
COMPARATIVE EVALUATION OF QUALITY OF AGE-SEX REPORTING
ANGLOPHONE COUNTRIES IN 1970 AND 1980 ROUND CENSUSES

Country	Year	Age ratio		Sex	Joint	Myers Index	
		Male	Female	Ratio	Score	Male	Female
Botswana	1971	5.7	5.6	7.1	32.6	8.5	9.6
	1981	3.4	5.2	3.6	19.4	7.1	6.8
Gambia	1973	15.1	25.3	14.7	84.5	44.7	51.0
	1983	15.1	24.9	14.2	82.6	42.9	50.7
Ghana	1970	8.3	11.6	7.6	42.7	26.8	31.2
	1984	9.0	10.3	5.7	36.4	27.4	33.4
Kenya	1969	5.1	7.4	6.7	32.6	21.2	22.2
	1979	3.6	4.0	4.2	20.2	12.4	15.4
Lesotho	1966	10.0	9.7	6.5	39.2	11.8	12.8
	1976	9.2	8.7	5.3	33.8	10.4	9.6
Liberia	1974	11.2	14.6	11.8	61.2	27.5	29.4
	1984	7.5	11.1	7.3	40.5	20.0	26.8
Libya	1973	5.6	11.0	6.1	34.9	9.8	25.8
	1984						
Malawi	1966	15.1	16.1	5.3	47.1	-	-
	1977	14.4	9.2	8.3	48.5	12.9	16.6
Mauritius	1972	5.5	5.2	3.0	19.8	2.0	2.7
	1983	5.2	4.9	3.2	19.6	2.1	2.3
Sierra Leone	1974	13.0	16.1	8.5	54.6	44.4	47.2
	1985						
Sudan	1973	17.7	20.7	9.6	67.2	48.2	56.0
	1983						
Swaziland	1966	17.6	11.4	9.4	57.2	15.4	16.2
	1976	8.9	6.2	5.6	31.9	14.1	14.7
Tanzania	1967	16.6	15.1	11.8	67.1	31.4	39.0
	1978	10.0	9.0	8.8	45.4	25.2	32.1
Uganda	1969	8.3	12.3	8.9	47.3	24.6	33.3
	1980						
Zambia	1969	14.4	8.0	14.6	66.2	14.1	15.3
	1980	5.5	9.4	7.7	38.0		
Zimbabwe	1969	10.0	9.9	8.8	46.3		
	1982	13.7	8.6	8.7	48.4		

16. Again, the tabulations prepared by countries on the basis of the data collected, have been more in line with analytical needs unlike in the past when analysis was an after thought.

Tabulations recommended by UN (Basic demographic characteristics)

1970 round

A. Geographical and migration characteristics

(a) Total population and population of major and minor civil divisions, by sex.

(b) Population in localities by size class of locality and sex.

(c) Population of principal localities and of their urban agglomerations by sex.

(d) Native and foreign born population by age and sex.

(e) Foreign born population by country of birth, age and sex.

(f) Native population by major civil division of birth, age and sex.

(g) Population by duration of residence in locality and major civil division, age and sex.

(h) Population by place of usual residence, place of previous residence and sex.

B. Demographic and social characteristics

(i) Population by single years of age and sex.

(j) Population by five year age groups and sex.

(k) Population by marital status, age and sex.

(l) Population by country of citizenship, age and sex.

(m) Population by nationality and/or ethnic/tribal group, age and sex.

C. Fertility and mortality characteristics

(n) Female population 15 years of age and over by age and number of children born alive.

(o) Female population 15 years of age and over by age and number of children living.

(p) Female population 15 years of age and over, in first marriage, by age at marriage, duration of marriage and number of children born alive.

D. Household characteristics

(q) Population in households by relationship to head of household, marital status and sex.

17. Among these tables f and q were recommended but were given lesser priority and tables g,h,l,m, and p were suggested as 'other useful tabulations.' Thus only a basic list of ten tables were recommended as first priority tables to be produced on demographic variables from the 1970 censuses.

18. For the African region, the list included only tables a to e, i,j, n and o. These were in line with the major objectives of the census of that period to, 'determine the number and present distribution of the population' and to a certain extent to 'determine population structure based on selected socio-economic characteristics.'

1980 round

19. The tables suggested for 1980 included all those recommended for 1970 in addition to the following:

(i) Population ..... years and over by place of usual residence, place of residence at a specified date in the past, age and sex.

(ii) Female population .... to 49 years of age, number of live births within the 12 months preceding the census, and deaths among these live births.

(iii) Population by maternal orphanhood and age.

20. There were also more details included in the tables in respect of geographic specifications and age groupings. Also duration of residence was tagged on to table h of the 1970 recommendations.

21. The African minimum recommendations were very similar to those for 1970 but added the tables k and m from the 1970 recommendations and thus excluded only table f,g,h,l, p and q. However, quite a large number of additional tables were suggested which were not included in the 1970 series. These are:

i). Population by sex, age and relationship to head of household.

ii) Population by sex, age, region of birth and region of enumeration.

iii) Population by sex, age, place of usual residence and place of enumeration.

iv) Population by sex, age, place of birth (in relation to place of enumeration) and duration of present residence.

v) Population.... Years of age and over by sex, age, place of work and place of usual residence.

22. One can easily note that the African recommendations included a wide variety of questions on migration and the tabulation plan reflect the concern of the region in regard to mobility and spatial distribution of the population.

23. We shall now take a few case studies from 1970 and 1980 round censuses and see what kind of tabulations were made and how changes have occured. As we have already noted that between 1970 and 1980 there was also a general improvement in the scope of a census in terms of items being canvassed, it is but natural to expect that the types and varieties of tables produced also should reflect this explosion; especially since the census is the only source of national data and most countries rely on the census to collect all types of information.

A. <u>Liberia</u>

<u>1974 census</u>

1. Summary population totals at national and county/territory level.

2. Broad population characteristics by county/territory and clan.

3. Population by sex by county/territory by district and clan.

4. Population by sex: citizens and aliens born in Liberia; citizens of Liberia; citizens born abroad; tribal affiliation.

5. Population by sex, by citizenship, by place of birth (in Liberia and outside) by tribal affiliation.

6. Population by county/territory, district and sex.

7. Number of localities by size, population and by sex and county/territory.

8. Population by single years of age and sex, by county/territory and district.

9. Persons born in Liberia and elsewhere by age and sex.

10. Place of birth of persons born in Liberia by place of residence, and sex.

11. Foreign born population by citizenship, place of birth and sex, county/territory.

12. Citizenship of the Population by age and sex by county/territory.

13. Length of residence in county/territory of enumeration by age and sex, county/territory.

14. Marital status of Population 10 years plus by age, sex and county/territory.

15. Number of households and population by size of household, county/territory.

16. Relationship to head of household by age and sex.

17. Number of children ever born for women 15 plus by current age, county/territory.

18. Number of children surviving by women aged 15 plus by current age, county/territory.

19. Number of births between February 1, 1973 and January 31, 1974 by age of women, county/territory.

<u>1984 census</u>

1. Number of localities, by size, population and sex.

2. Population by single years of age and sex.

3. Persons born in Liberia and born elsewhere by age and sex.

4. Place of birth of persons born in Liberia by age and sex.

5. Foreign born population by country of birth, citizenship and sex.

6. Citizenship of population by age and sex.

7. Length of residence in county/territory of enumeration by age and sex.

8. Population by length of residence in county/territory of enumeration, place of birth and sex.

9. Marital status of the population 10 plus by age and sex.

10. Number of households and population by size of household.

11. Heads of households by sex, age and activity status.

12. Relationship to the head of household by age and sex.

13. Ethnic affiliation of the population by age and sex.

14. Religious affiliation of the population by age and sex.

15. Religious affiliation of the population 10 plus by marital status and sex.

16. Religious affiliation of the population by ethnic group and sex.

17. Religious affiliation of women 10 plus by number of children ever born, and current age.

18. Number of children who died for women 10 plus by current age of women.

19. Number of children surviving for mothers 10 plus by current age of mother.

20. Number of births between February 1, 1983 and January 31, 1984 by age of women.

Sudan
1973 census

Provincial level (10 provinces)

1. Population by place of residence, mode of living and sex.

2. Population of towns by sex.

3. Population and households by size of houshold and institutions.

4. Population by age and sex.

5. Population by marital status, age and sex.

6. Population born in Sudan and outside by age and sex.

7. Population by place of birth, age and sex.

8. Population by nationality, age and sex.

National

1. Population by relationship to head, marital status, sex and province.

2. Population by place of birth and province of enumeration.

3. Socio-economic characteristics for 35 selected towns (urban).

4. Urban population by sex and size class to towns.

1983 census
Provincial level (18 provinces)

1. Population distribution by mode of living, area and peoples councils.

2. Population by single and five year age groups, sex urban/rural ahd nomadic.

3. Households by size and population, urban/rural, nomadic.

4. Households by size, population and number of family nucleus.

5. Population 12 plus in private households by sex, relationship and marital status.

6. Heads and other members of households by age, relationship, urban/rural and nomadic.

7. Population by sex, age, marital status, urban, rural.

8. Ever married women 12 plus by age, number of children ever born, urban/rural.

9. Ever married women 12 plus by age, number of children surviving, urban/rural.

10. Mothers 12 plus by sex and survival of last birth during 3 years preceding the census, urban/rural.

11. Population by age, sex and survival status of mother, urban/rural.

12. Usual residents by province of birth (urban/rural), sex and broad age groups (urban/rural).

13. Non Sudanese by continental origin by sex, broad age groups and place of birth.

14. Usual residents of large towns (50000 plus) by sex, broad age groups and place of birth (urban/rural).

<u>National</u> (proposed)

In addition to the 18 provincial table volumes, it is proposed to aggregate the provincial tables to produce regional and national tables. Also additional tables at national level are envisaged. Tentative list is given below:

1. Urban population by province of birth and province of usual residence (urban/rural) and sex.

2. Population (urban/rural) by duration of residence in province, age and sex (urban/rural).

3. Ever married females by age, children ever born by sex of children (urban/rural).

4. Ever married females by age, number of children surviving by sex of children, (urban/rural).

5. Mothers 12 plus with at least one child under 15 years of age living in the same household, by age of mother and by age of children (urban/rural).

24. A look at the types of analyses carried out during the 1970 and 1980 rounds also will give us an idea as to the widening scope of demographic and related data analysis and the importance countries have attached to such detailed studies.

A. <u>Liberia</u>
<u>1974 census</u>

1. <u>Land and Life</u>

Historical background; geographic features; people and social life; education, health and religion; the government, economy.

2. <u>Quality of census data</u>

Introduction, coverage error; content error; collection, processing; and presentation error; comparability of data of censuses of 1962 and 1974.

3. <u>Population distribution, urbanization and migration</u>

Introduction;population distribution and density by counties; population growth by counties; urbanization; urban population growth; concentration of urban population growth; international migration; internal migration.

4. <u>Sex and age composition</u>
General; overall sex ratio; spatial distribution of observed overall sex ratio; observed age distribution; reported and adjusted distribution of population by age and sex, 1974.

5. <u>Fertility</u>
Introduction; level and pattern of fertility.

6. <u>Mortality</u>
Introduction; level and pattern of mortality; rural/urban differentials.

7. <u>Ethnic affiliation and citizenship</u>
Ethnic groups-components and growth; spatial concentration of tribes; urbanization; age; sex situaions; citizens and aliens-size, growth, spatial distribution; citizens and aliens by sex and age.

8. <u>Population projection</u>
Introduction; total populations; sex and age structure, rural/urban differences; components of growth of projected population.

1984 census (proposed outlines)

1. Evaluation of Data

Evaluation and appraisal of data collection system, instruments and procedures. Pilot testing, training, quality controls, field checks etc. Data processing and tabulation. Coverage evaluation. The PES and findings and results. Content evaluation. Age and Sex. Single year and five year group data. Digit preference, age preference, omission, duplication, shifting. Indices for digit preference. Age and sex ratio and UN Joint scores. Comparison with 1974. Cohort comparison-survival ratios (cohort and overall). Age-sex distribution-percentage distribution, cumulated percentage distribution. Derived ratios like child women, dependency ratios. Mean and median ages by sex. Growth rates by age group and sex. Population pyramids and other pictorial representation of age sex data. Manipulation of reported (adjusted) age sex data. Reverse survival, forward survival methods. Comparison with models. Derivation of fertility and mortality parameters from age-sex data. One census and two census methods. Comparability of parameters with estimates from past and current data and information available from other sources.

2. Population size, growth, and age-sex structure

Population size as from 1984 census. PES results and adjustments. Comparison of 1984 population with projections and estimates. National projections, other projections from UN, ECA, other international agencies. Growth rate of population. Growth of age-sex segments. Implications on dependency, youthfulness of population, old age groups and working ages. Sex composition and changes over time and space. Implications of age-sex distribution for population growth, fertility, mortality and migration. Effects of age-sex composition on needs for education, health, food, housing, employment etc. Comparison of growth rate with past and with other African and developing countries. Implications of growth rate and age sex distribution of future population growth size, structure etc.

3. Ethnic and religious composition

Socio-cultural patterns of the various ethnic and religious groups. Spatial distribution and changes in size, age-sex structure. Attitudes and beliefs impinging on education, economic activities and way of life.
Kinship groups, marriage, polygamy, divorce. Attitudes and practices relating to child bearing and child rearing. Health and living conditions. Migration and movements. Internal and international migrations. Housing patterns and amenities. Estimates of fertility, mortality and population change.

4. <u>Population distribution</u>

Population distribution over space. Changes over time. Growth and declines of localities. Causes and consequences. Density and land utilization. Spread of population and impact on socio-economic development planning. Size class of localities. Lorenz Curve and Gini Concentration Ratios. Rural development programme. Availability of education, employment opportunities, housing amenities, health facilities etc., by rural residence. Changes over time. Density charts and dot diagrams.

5. <u>Nuptiality and fertility</u>

Socio-cultural factors in nuptiality, fertility and family formation. Value of children and other societal norms affecting level of fertility. Proportions married, mean and median age at marriage. Polygyny, divorce, widowhood and separation. Childlessness. Perception of high infant and child mortality. Role of children in economic pursuits. Influence of education. Changing pattern and age at marriage.

Level and pattern of fertility. Use of current (past one year births) and retrospective (children ever born) data to estimate level of fertility. Comparison with previous estimates.

Differentials by socio-economic and geographic characteristics. Education, religion, ethnic group, economic activity, rural-urban etc.

Family planning and contraceptive practices-traditional, other. Perception of large family size among population. Population policy indications and governmental perceptions and pronouncements. Fertility prospects.

Estimates of fertility from other sources-age-sex distribution. One census and two census methods. Coale Demeny, Reverse Survival Methods.

6. <u>Health, morbidity and mortality</u>

Hospital and health facilities. Food and nutrition. Economic and social conditions affecting health, morbidity and mortality. Maternal and child health programmes. Immunizations. Provision of safe water supply, improvements of living conditions - toilets and other environmental improvements. Morbidity and causes of death. Vital and health statistics.

Level of infant, child and adult mortality from child survival and mother survival data. Comparison with 1974 data on age sex composition, child survival. Deaths in household by sex.

Life tables by sex. Comparison with 1974 and previous information. Trend of mortality.

Mortality differentials - socio-economic and geographic implications for future course of mortality.

7. Migration and urbanization

Estimates of internal migration by growth rate method. Place of birth data and duration of residence information. Life time and intercensal migration using two census age, sex and place of birth data.

Sources and destinations of internal migration, Causes and consequences. Depletion of population in rural and remote areas and fast growth of urban localities, towns and especially Monrovia. Selectivity of migration by age and sex. Trend of migration over time.

Urbanization and growth of towns. Trend of urbanisation, characteristics of urban and rural populations. Changes over time.

Impact of rapid urbanisation on urban areas - housing, provision of education, health facilities, employment, food etc.

Future prospects of migration and urbanization in Liberia. Policies and programmes to modulate migration and urban growth.

International migration. Role of international migration in manpower needs.

8. Population projections

Evaluation of basic demographic parameters-age, sex, fertility, mortality and migration. Adjustment of base age sex data and parameters. Assumptions of scenarios on fertility, mortality, migration. Projections of national population by age and sex. Urban-rural projections and other sub-national projections, e.g. for counties, districts and other lower level administrative ecological regions as needed. Sectoral projections like school, economically active, households etc. Estimation of food, health and housing requirements. Implications of alternative and speculative projections. Policy implications and choices. Comparison of projections with other international projections.

B. Sudan
1973 census

1. Land and people

Introduction; the land; administrative aspects; transportation, communications; the people-beliefs, tradition, culture, population size, fertility, mortality and migration; population growth and prospects.

2. Population count, age sex and other characteristics - an evaluation

Introduction; the 1955/56 count; other data; the 1973 census; census questionnaire and enumeration; population count; population distribution; the nomads; the cotton pickers; sex ratio; survival ratios; age distribution; fertility, mortality and migration; population size; conclusions and recommendations.

3. <u>Mortality analysis</u>
Introduction; past findings and other data on mortality; data from 1973 census; mortality level; mortality differentials; summary, conclusions and recommendations.

4. <u>Fertility analysis</u>
Introduction; past data and findings; data from 1973 census; fertility level; fertility differentials; age pattern of fertility; fertility trends and prospects; summary, recommendations and conclusions.

5. <u>Population distribution, migration and urbanization</u>
Introduction; past data and findings; data from 1973 census; estimates of migration by age, sex and province; urban growth; urbanization by province and growth of towns; causes and determinants of migration; consequences and some implications of migration; external migration; future prospects of migration, urbanization and settlement; policy recommendations; conclusions.

6. <u>Population projections</u>
Introduction; brief review of previous population projections; projections based on 1973 census; estimates of 1975 base data; assumptions; total population by single calendar years; population by age, sex and by single calendar years, 1975-85; population by provinces; city populations; summary and findings.

<u>1983 Census (Proposed outlines)</u>
<u>Brief Outlines of Analytical Reports</u>

1. <u>Evaluation and Adjustment of Basic Demographic Data</u>
Introduction; Importance and need for data evaluation; quality checks and controls in the various stages of data collection, data preparations and tabulation of data; types and varieties of data collected. Quality of age-sex data collected; Coverage of population and of various age-sex groups. Sampling in census. Comparison with 1973 census, surveys etc. (SFS, data from education ministry etc.); adjustment of data and lessons learnt from the 1983 census.

2. <u>Fertility Analysis</u>
Introduction; Data situation quality; Socio-cultural factors - marriage, nuptiality and value of children; level of fertility. Childlessness, sub fertility; patterns of fertility - age, marital status; differentials; socio-economic, regional (rural/urban, provincial); fertility and family structure; fertility trends. Cohort fertility; fertility prospects. Population policies and family planning implications of fertility. Summary and conclusions.

3. Mortality Analysis

Introduction; Health situation, hospitals and housing conditions. Environment. Nutrition and food consumption. Attitude and practices. Causes of death, morbidity. Sources and quality of data; level of mortality. Infant, child and general mortality; mortality patterns. Age-sex patterns; mortality differentials-sex, age, socio-economic, regional (rural/urban, provincial); mortality trends. Changing pattern of diseases and causes of death; prospects of mortality decline and changes; health policies and planning for development. Health for all by year 2000. Population policies and their impact on health, nutrition and mortality. Implications of mortality on population. Summary and conclusions.

4. Migration Analysis

Introduction; Spatial distribution of population in Sudan. Environmental, socio-economic situation and political factors of movements; data availability and quality; population distribution, dispersion and density. Causes and implications; estimation of migration - Direction, volume and characteristics; causes and consequences of migration. Urbanization and growth of towns. Employment, housing and environment. Health and diseases, drought; trends in migration: age-sex, direction and volumes; prospects of population movements. Implications on agriculture, land utilization, employment; implications of migration on population policies. Summary and conclusions.

5. Household and Housing

Introduction; Problem of housing-big cities and migration; housing quality in rural and urban areas; sources and quality of data. Household and housing data from censuses, surveys and administrative operations. Master plans for cities; socio-economic determinants of housing. Households and families. Migration and its impact on housing. Family size. Housing amenities and facilities in towns; projection of households and housing. Estimation of housing need.

6. Population Projections

Introduction; Importance and need for population projections. Types of Projections. Period and frequency of projections. Available population projections and their evaluation in respect of results of 1983 census. Lessons learnt from the comparison. Sources and quality of data. Components of population change. Assumptions for base data and future periods; projection methodology. Mathematical and other methods. Component method. Projection of sub-national, sectoral populations. Variants of fertility. Hypothetical projections. Impact of dynamics of population change and their implications; implications of population projections for social, economic, regional planning and policy-making. Summary and conclusions.

2.3 Sources of Errors and Biases

25. Evaluation or appraisal means the measurement of achievement against goals. Evaluative tools are often necessary to gauge success or failure of efforts made in regard to massive and multipurpose data collection. Involving as they do several enumerators, varied types of respondents, variety of topics, difficult logistical problems and a wide range of procedures and processes, they provide ideal ground for evaluative studies because at every step and stage of data collection, compilation and presentation there are possibilities to introduce errors, biases, deficiencies etc.

26. It is not possible to avoid completely every type of error etc, but what one attempts to achieve is to minimise these deficiencies and, in fact, what we usually need is an indication of the range of variation (perhaps not too wide) in the data. Thus an evaluation and appraisal of the data is a pre-requisite before they are interpreted, utilised etc.

27. Evaluation has three main objectives:- (i) to identify the types and sources of error or biases in order to know which groups, items or methodology produced these deficiencies so that appropriate preventive measures can be taken; (ii) to measure accuracy, i.e. provide the range of variation of the data for the users in the appropriate analysis and applications of the statistics and (iii) to adjust the data taking account of the varieties and amounts of errors in the data.

Types and sources of errors and biases

28. Evaluative studies probe into the quantitative and qualitative aspects of the data. Hence two types of errors being investigated by such studies fall under the broad headings:- coverage and content errors. Coverage deals with the completeness and quantitative aspect of the enumeration whereas content pertains to qualitative characteristics like age, sex, marital status etc. There could be overlap of content and coverage as when one wishes to evaluate sub populations of given age-sex groups where not only size but characteristics play a part. It is noted that coverage may be affected by or affects content characteristics. The usual omission of young children from enumeration is a case in point. Conversely, content might affect coverage. For instance, if a question on military service is asked or any other question which might have connotation like conscription, taxation etc., then it can result in omission of the affected group. Omission or wrong reporting also might be caused by the questionnaire designs as when some detailed questions are only for certain segments of the population. For example, questions on labour force, marital status, fertility and so on which are for specific population groups, might induce the enumerators either to omit some of the persons in order to avoid additional work or shift them to groups which do not require such extra

labour. Exaggeration, on the other hand, can arise when some benefits are involved as in rationing, old age benefits, medical aid, voting rights etc. Enumerators might also be tempted to bloat up the groups if payments are according to numbers canvassed.

29. One common feature of most census questionnaires has been that they have ten lines for a possible ten-member household. When the number exceeds ten in any household, the instruction is to note down in the schedule that it is being continued in the next schedule, write identification particulars in the next schedule and number the members as 11, 12 .... Instances have been noted of enumerators or respondents not proceeding as intended because of fatigue or inertia. This results in either complete omission of some members of large households or in putting the additional members as if they constitute a separate household. These would result in coverage and/or content errors in data as observed in the 1973 census of Sudan and in the pilot census of Kenya, 1979. Inspite of the fact that such an error was noticed in the 1970 pilot census of Botswana and accordingly the 1971 questionnaire had 12 lines, in 1981 again the questionnaire was prepared with 8 lines and a similar error has been noted in the data.

30. Another common observation is the one on questionnaire design whereby some items are relegated to obscure corners of the schedule. For instance, current fertility and mortality from households are usually put in a corner or at the bottom of the questionnaire. Enumerators either ignore these as in the 1973 census of Libya or do not care to match them with other information in the schedule like infants in the household, survival status of last years birth etc. In some cases usual members of households might have given birth during the year, but at enumeration time might have left the household either due to migration or death. If information is collected on usual members present, then information on these births would be lost, especially if the mother is dead. If the infant also died, then that event also would not be reported.

31. Pre-coding of a questionnaire has a lot of advantages but some problems may arise, if care had not been taken at the preparation of the questionnaire. As an example, in the 1978 census of Tanzania children ever born was pre-coded and the code for 'not applicable' and '8 or more children' was the same. This resulted in difficulties in analysis and interpretation. Also pre-coding, unless carefully planned, may result in much loss of information and definitely will restrict analysis.

32. A salient point to keep in mind while preparing survey schedules is that it should be as clear as possible, should contain only what is essential for the purpose in hand, should not be loaded with questions which may arouse fear, suspicion, anger or hatred and should be easy to administer. Proper

education of both respondents and enumerators will go a long way in ensuring good quality data.

33. Geographic preparation is an important element in data collection. If enumeration areas are not demarcated and identified, it is possible that enumerators may either omit or duplicate some areas, resulting in coverage and content errors. In this connection it is necessary to mention that listing of structures, dwelling units etc. would go a long way in ensuring accuracy of data.

34. A common observation in surveys has been the omission of one person households. This would vitiate the data both quantitatively and qualitatively. Enumerators should call back these households where nobody is found at time of visit and in case all efforts fail, then they must try to obtain as much information as possible from neighbours. The problem of remote area dwelllers, nomads, mobile groups, sparsely populated hamlets, difficult terrain and environment and so on make enumeration indeed very difficult and must be planned well in advance. Ignorance, superstition, fear, hostility, anger and a general feeling of insecurity all could mar an enumeration and every effort must be made to ensure coverage and quality of information. Education, publicity and creation of rapport between the data collection agency and the respondents are channels for assuring cooperation and improving quality of data. In this vein, avoidance of questions which may arouse suspicions, fear etc. should be kept in mind. It is better to plan ahead and avoid bottlenecks than to adjust the data after collection. No amount of doctoring will cure the data of the errors, biases etc.

35. The method of administering the questions, interpreting and recording them are important elements in ensuring data quality. Since the population may not understand the language or subject matter of enquiry, it is essential that training manuals and translations be used in local language to assist understanding by respondents and interpretation by enumerators. Care should however be taken that the questions are not mutilated in translation and information biased. In the Fiji fertility survey it was noticed that since the word in local language for 'former' and 'first' being similar (pahale) there was interchange in translation resulting in problems.

36. Since there are usually many languages spoken in most Anglophone African countries, the census questionnaire is usually in English or at the most in one of the major language necessitating translation. But translations are always liable to be misinterpreted; this risk is all the greater when the translation is destined to be addressed to respondents with varying levels of competence in the language concerned. There are also instances of misunderstanding of questions. It looked from a perusal of census schedules of the 1975 census of Somalia that since the question on fertility was addressed to

only females aged 14 years and above, some enumerators interpreted that only female children should be recorded and actually in some areas only female children were noted down. Another instance of misinterpretation because of difficulty in translation seems to have occured in the 1973 census of Sudan where a child was literally taken as to be naturally very young. This resulted in mothers being reported as childless even in cases where from the schedule it was obvious that she was living with a grown up son or daughter. Here another misinterpretation was due to the question on 'children living with you here.' A mother staying with her son or daughter was staying with them and not they with her. It also looked that some enumerators interpreted that the question on children ever born etc. was only for the wife of head and not for other members of household, especially if they are old, widowed, separated etc. These resulted in large proportions of females being not reported regarding the number of children ever born to them and they were all categorised as 'not stated.' Even genuinely childless women got mixed up with the 'not stated' and the problem of 'zero error' as noted by El Badry occured but with a difference and that is, there were only few cases reported as zero parity and there was no variation over age, making the problem of estimation of the 'true childless woman' extremely difficult. In this connection the questionnaire design of the 1980 census of Zambia looks superior, as it requires all eligible women to be relisted and their fertility history recorded.

37. The method of data collection plays an important role in data quality. For instance whereas it may be easier to carry out a de facto count, in some cases it may be advisable to utilize the de jure concept in enumeration. Many times compromises are to be made. Again, in certain situations a reference date is adhered to whereas in others a floating reference date is used. Criteria for inclusion/exclusion of members of households also should be clearly spelt out so as not to vitiate results. In the 1955/56 sample census of Sudan, the de jure concept was utilized and the enumeration lasted for several months. Obviously no specific reference date could be fixed. It is suspected that these contributed to an inflation of the population. In the 1973 and 1983 censuses therefore the de facto concept with fixed reference date was adhered to. If a census is taken on the de facto concept and then a PES is conducted, even only after a very short interval, using de jure concept as in Swaziland 1986 census, the problems of matching and coverage evaluation becomes difficult. In the 1980 census of Zambia, in addition to de facto concept, there was a problem of no reference date in census. An instance of effect of looseness in the use of inclusion/exclusion criteria is the 1978/79 labour force and migration survey of Lesotho. Apparently, absentees were excluded if they were away for a month and included if they were to return soon (soon was instructed as one week). The result was an omission of school children and those temporarily away (possibly looking for work) and this vitiated the main focus of the survey.

38. An important stage of data preparation is coding. Great care is needed at this stage, as otherwise, the quality and usefulness of the data collected are reduced. The question of special coders for specific items versus one coder for all items, the method of payment - either per items coded or number of hours worked - all can affect data quality and should be evaluated. Another problem is that once an item is coded and transferred to the tape, it will not be possible to draw any additional information except by referring to the original schedules themselves. An illustration is the coding of place of birth in the 1973 census of Sudan which resulted in loss of valuable information. This could have been avoided by using a three instead of the two digit code used in order to distinguish between intra district urban-urban and rural-rural migrants. Another instance is the omission of relationship column while coding data from the 1976 census of Lesotho depriving analysts of a valuable piece of information.

39. An interesting example of coding errors vitiating data is that from the 1976 census of Lesotho. Examination indicated that most of those coded as 'not stated' were in fact 'zero parity, women reported with 88 children was due to wrong coding etc. At the punching stage further errors are introduced, which if not checked and rectified could foul up the data. In the 1976 census of Lesotho and in the 1975 census of Somalia, it was observed that in many cases zero parity which should have been '00' got punched as '10' inflating number of children. Shifting of columns while punching, skipping lines etc. also resulted in queer results. Verification should be thorough in order to escape from such situations. At least there should be no further deterioration of data quality once it has been brought to the office. An interesting problem noticed in data processing especially where the data processing expert does not know statistics or demography and the demographer/analyst did not instruct the data processing expert fully about the various terms and coding of data is that noticed in Sierra Leone at the 1985 census. Here the number of children ever born and those who died were coded numerically in two digits such as 00, 01, 02, ... The not stated was given the code 99. In tabulating the data, the code 99 was considered as 99 children ever born or those who died vitiating both the parity and proportion of surviving children. An observation in this connection is in order. It was noticed in several cases that record keeping, especially of in and out flow of census or survey schedules was not satisfactory and this can result in loss of information. Another observation pertains to identification numbering of individuals while transferring information to cards, tapes etc. There should be uniqueness so that identification becomes easy and accurate. This also will ensure that there is no duplication or omission of records while keying in data and it also assists in matching exercise.

40. A most important but difficult logistical problem is to ensure adequate stock of questionnaires in each of the enumeration areas, especially when terrain is difficult, transportation inadequate and communication takes a long time. Both too much and too little numbers of schedules would be costly. This is valid for all other aspects of data collection in the field as well. A problem noted in the 1976 census of Lesotho was that there was shortage of questionnaires in some areas inspite of total number printed being adequate. This happened obviously because larger numbers were dispatched to outlying areas and hence less were left behind in Maseru. At the last minute while the enumeration was in progress, reprinting was resorted to but a mistake seem to have been made in respect of forgetting that the census booklet should have every tenth page as a long questionnaire. Only the short questionnaire was apparently printed and used in some parts of Maseru culminating in reduced sampling fraction, non representative nature of sample, etc. In the 1985 census of Sierra Leone, instead of the revised questionnaire, only the pilot questionnaire was printed and used at the main census.

41. Before we conclude this section it is in order if we talk a little bit about the crucial role of the enumerator and supervisor. In addition to training and providing manuals etc., the enumerator should be impressed about the necessity for good quality work in the short time usually available for data collection. Close and continuous supervision is essential in order to avoid poor quality of information. Especially in the earlier stages of data collection, there is need for checking all the completed schedules and errors, mistakes etc. pointed out and investigated. If several enumerators make similar mistakes, then it may be advisable to have group meeting for discussions and clarification of doubts.

42. Since it is difficult, if not impossible, to assign characteristics in the office after the data have been brought from the field and imputation should be avoided as far as feasible, it is necessary to impress on enumerators and supervisors that all of the information should be filled in the schedule in the field itself and not left blank, before they are turned in to the office. It may be easier and more reliable to assign the particulars while still in the household than after leaving the locality because the information could be corraborated with other members of the household, neighbours, physical or physiological factors, social status and a host of other variables related to the information one is looking for. In the 1974 sample census of Zambia age was not recorded for about a fifth of the population making the data of very little value. Another type of omission is brought out by enumerators who leave the sex column blank because to them the deciphering is easy from the name. This results in errors at the later stages of data processing when the sex has to be assigned. This is due to the fact that not only are names same in some cases, many times they are similar and the difference may be in the last vowel as in Spanish names like Francisco (male) Francisca (female) and in Arabic names the female name

is same as the male one but with a vowel 'a' at the end as in Zuhair (male) Zuhaira (female). When handwriting is illegible or the writing is affected by atmospheric factors smudging the writing, then assignment of sex may result in errors.

43. In any enumeration there will be cases with missing entries for characteristics - usually called 'not reported, 'unknown' etc. There are several methods of imputing values for such characteristics which are consistent with other characteristics. Hot and cold deck methods are well known. However, serious reservations have been raised about imputations etc. and hence every effort must be made to collect all information in the field itself (Banister, 1980).

44. Finally a word or two about the kingpin of any survey - the respondent - is essential in the understanding of data quality and quantity. It is the respondent who supplies the information and he plays a vital role in the success or failure of a survey. His cooperation is essential to achieve the aims of an investigation. However, due to ignorance, superstition, fear or other feelings or again sometimes deliberately the respondent may supply incorrect or biased information. Instances of respondent resistance in supplying information to enumerators or wilfully falsifying the data are noted when the data collection is associated with governmental activity like taxation, conscription or allocation of benefits as was reported during the 1964 census of Libya and same is suspected at the 1962 census of Nigeria. But generally the experience in developing countries is that very rarely does the respondent deliberately try to give false information or refuse to provide information. Mostly it is ignorance or superstition which affect data quality. Publicity, education campaign and other avenues have been applied to improve the reliability of data collected. In this connection the role of young persons, especially school or college children has been noted to be worth trying.

45. Even though it may be difficult to see all the reported members in a household either because they are physically absent from the household or are present but cannot be seen because of ill health, socio-cultural reasons (purdah), age (infancy) etc., it may be a good idea to see as many as feasible and get information from individuals particularly on personal characteristics.

46. Several methods of preventing errors and biases at the planning and data collection stages have been indicated in the previous discussions. There is need also to control the quality of the data after it is collected especially at the subsequent stages like storage, editing, coding, keying, processing, tabulation, publication etc. 'Quality' refers to fitness for use, or the degree to which a specific product satisfies the wants of the users and 'control' means the planned cycle of activities by means of which intended objectives and quality standards are achieved. Quality

control, thus encompasses the entire collection of activities by which fitness for use is achieved (POPSTAN, 1979).

47. A quality control system encompasses the following; (a) establishing quality standards, (b) determining verification techniques, (c) measuring quality through record keeping, and (d) taking timely and appropriate action to maintain the desired quality. Since a census involves a complex arrangement of people and machines operating on material and information, there is enough scope to introduce quality control at various stages of data production and especially at the editing, coding, keying, processing, tabulation and publication stages. It is recommended that countries plan to integrate quality checks and controls at various levels from the early stages of preparing for a census so that a good quality product can be ensured. Prevention is better than cure.

48. We have listed the various channels and avenues through which errors and biases in the data could enter, what types of deficiencies result therefrom and what preventive and remedial actions need be taken. These are only illustrative and do not claim to cover the entire gamut of types and sources of defects in the information collected. What is really aimed at is to point out that any slackness or negligence in any one of the several aspects of data production can lead to deterioration in data quality and at the same time stressing the importance of quality assurance.

49. Since the cure should be specific to a disease, so is the case with data adjustment or correction. Once the source and types are identified and the error and its magnitude detected, appropriate remedial measures could be instituted to make the imformation usable. This also enables one to plan future data collection in a better way. The more one knows about the problem, the easier it becomes to take care of it.

50. In the next section we present methods, techniques and tools for the detection and evaluation of possible errors, biases and deficiencies in a given data with particular focus being put on basic demographic data, i.e., population size, its structure (age and sex) and dynamics (fertility and mortality).

2.4 Methods of data evaluation and error detection

51. How do we evaluate demographic data? There are two broad types of techniques of data evaluation and error detection - the direct and the indirect methods. Among the direct method are the reinterview surveys through post-enumeration checks. Internal and external consistency checks, analytical tools for derivation of parameters leading to consistency-convergency criteria (i.e., the derived parameters should be consistent with each other and with the other types of data and information available and they should form a convergent set and not differ too widely among themselves) constitute the indirect methods.

52. Even though it is difficult to utilize indirect methods efficiently to estimate coverage error, there are several instances where such coverage errors were in fact detected by these methods.

Direct Method - Reinterview Surveys

53. Surveys conducted shortly after the field operations are completed are generally referred to as post-enumeration surveys. They are carried out on representative sample areas taking account of the heterogeneity in the data and keeping in view that one should be able to generalise and get information on important special segments of the populaion. Thus population clusters, EA (enumeration areas) or other identifiable geographical units are utilised. These surveys are carried out for coverage and content error checks and for collecting additional information. For evaluating content, a sample of households or persons enumerated in the enquiry can be reinterviewed. For evaluating coverage, however, an area sample is required. Two basic designs have been used - (a) one which stresses attaining high accuracy by using better qualified enumerators, choosing the most knowledgeable respondents to provide information, better interview methods etc., and (b) one which strives for an independent repetition of the first enquiry under essentially the same conditions. Matching and reconciliation is then carried out. Evaluation of coverage is thus more difficult with a de facto census and it is more so when there is no reference date.

54. The timing of the PES should be such as not to be too close to the enquiry or too far from it. For independence of the operations, it is essential that all field data are received back in the office. PES has been used as a single system to replace the figures obtained in the first enquiry or as dual system for correction and adjustment.

55. Very few African countries have utilised the PES as an instrument for checking the quality of enumeration in terms of coverage and content. According to ECA, of the Anglophone countries that completed the 1970 round of censuses in Africa, no ad hoc PES was undertaken in: Botswana (1971), Gambia (1973), Lesotho (1976), Libya (1973), Mauritius (1972), Sierra Leone (1974), Somalia (1975), Sudan (1973), Swaziland (1976), Tanzania (1967- only a regional study was made), Uganda (1969) and Zambia (1969). In the 1960 census of Ghana, the content and coverage evaluation survey was combined with a supplementary enquiry to collect much needed additional information and on the basis of experience, the 1970 supplementary enquiry put more stress on collecting additional information.

56. Some of the countries, however, succeeded in getting conclusive results from PES. Liberia (1974) and Cameroun (1976) are important among them who could assess the coverage of population by sex and broad age segments. In the case of

Malawi (1966) and Kenya (1969), the PES clearly indicated underenumeration in the census and these formed the basis for correcting census totals.

57. In Libya (1973) and Mauritius (1972) there were no PES but data quality and coverage were evaluated by analytical tools. In Sudan (1973), on the other hand, in addition to analytical tools, there were resurveys and check surveys carried out in specific areas and among groups where doubts were raised on enumeration quality. The resurvey in Nigeria in 1963 consequent on the census of 1962 is well known for the controversy it generated. Such surveys might have serious implications. Any group not satisfied with an enumerated population figure might demand a recount and stop only when its expectation is fulfilled-however unscientific the expectation be. It is thus essential that no scope be given for such a contingency.

58. The general recommendation by ECA during the 1970 round of population censuses in African countries was that ad hoc post enumeration surveys should be done for evaluating the census data. It was stressed during the second meeting of the African Census Programme country experts held in Addis Ababa in 1974 that the census evaluation should consist of coverage rather than content checking when a PES is undertaken to evaluate data. For errors of content, the data should be evaluated by using methods of demographic analyses.

59. For the 1980 round of censuses, the working group on recommendations for the 1980 population and housing censuses in Africa (July 1978) urged that, "a post enumeration survey for coverage evaluation should be planned, organised and executed as an integral part of a population census." (ECA, 1986)

60. Despite this recommendation, only 15 out of 47 countries which carried out censuses between 1975 and 1984 carried out coverage evaluation surveys as part of their census programmes. These are: Algeria (1977), Botswana (1981), Burkina Faso (1975), Burundi (1979), Cameroun (1976), Cote d'Ivoire (1975), Equatorial Guinea (1983), Ghana (1984), Kenya (1979), Liberia (1984), Madagascar (1975), Mozambique (1980), Senegal (1976), Zaire (1984) and Zambia (1980).

61. From the evidence available so far on the post enumeration surveys carried out in the 1980's, it appears that the main objective of coverage evaluation may not be attainable. In the case of Zambia (1980) the sample design of the post enumeration survey 'was found to be defective and the results cannot therefore be used to detect coverage and/or content errors.' As such the PES data was "used for training in coding and other aspects of data processing" (Zambia, 1987). In the case of Botswana (1981) although the PES results have not been published, problems faced by the PES included "fatigue on the part of the field staff," "refusals by respondents" and

on the whole "the quality of enumerators and supervising staff was poorer than those used for the census enumeration." (Botswana, 1983)

62. The results of the Ghana (1984) PES have not yet been published. According to the preliminary report of the census, it is stated that although "the enumeration was somewhat prolonged" (having lasted from 11-30 June 1984 compared to the census which lasted from 11-22 March 1984) and there were some "difficulties in accessibility" due to rains, field enumeration presented few problems. (Ghana, 1984)

63. Concerning the use of other methods of evaluating census coverage, "the current situation in the region regarding household sample surveys and administrative records seem to indicate that the PES would be, for the present, the only feasible source of data for census evaluation studies." (ECA, 1987) It appears that these same problems faced by the 1980 Round PES were similar to those faced by the 1970 Round PES also. There was general lack of enthusiasm in many African countries during the 1980's for the use of a PES to evaluate census coverage. This situation can be improved by eliminating the problems facing PES in Africa. Improvements in census cartographic preparations, better delimitation of enumeration and sample area boundaries and training of enumeration staff in map reading can minimize problems associated with identifying the areas to be covered by field staff. The problem of matching households and persons names can also be helped by more thorough identification specifications. There is also need for adequate budgetary provision and remuneration of staff, and thorough training of evaluation technical staff to ensure high quality execution of the PES after the exacting demands and resultant fatigue from the census. Among the problems associated with PES discussed by the working group on recommendations for the 1990 Round of Population and Housing Censuses in Africa, were the following: (a) Manual matching resulting in considerable delays; (b) inadequate provision on the questionnaires for identifying households and housing units; (c) recording on names of respondents, (d) inability of enumerators to read maps thus failing to cover selected sample areas accurately, (e) failure to achieve independence between the PES and census, and (f) importance of the timing of the PES.

64. As regards use of data from PES/Census comparison, the Working Group felt that estimates of coverage error should not be used to correct the total count obtained from census but could be used in adjusting the age and sex distribution of the population for purposes of population projections. However, the Working Group noted that use of the PES for content error evaluation was gaining acceptance. As the accuracy of estimates derived through evaluation using analytical techniques can be adversely affected by errors in the census data it was suggested that African countries might consider

using the PES not only for coverage error evaluation but also for evaluation of content error. For the future however, it appears that African censuses will rely more heavily on the PES as the best way of evaluating coverage errors. For content errors use of demographic analysis techniques can be supplemented by the PES data.

65. Even though the experience of African countries in carrying out PES is still inadequate, the common feature of those few PES carried out is that they utilised the available area frames and maps and did not attempt to evaluate the adequacy of the cartographic and other areal preparations of the census. Thus, it seems that what was done was an appraisal of population count within the delimited areas and there was no attempt to check completeness of areal coverage.

66. Some of the major factors contributing to the failure of PES were: (i) Population mobility; (ii) unrecognisable boundaries of sample areas or inadequacy of geographic preparation; (iii) varying names of persons and also many common names, different ways of spelling names etc., and; (iv) general fatigue of enumerators and perhaps of the repondents. An example of the last factor is shown by the PES conducted after the 1981 census of Botswana.

67. With the experiences from the 1970 and 1980 rounds of African censuses, the 1990 round census recommendations on coverage evaluation included post enumeration surveys as a desirable component in the census evaluation programme but keeping in view the experiences of the past, it was stressed that unless tests show that the technical problems involved in matching and other important aspects of a PES could be controlled to ensure satisfactory results and adequate budgetary provision and the necessary technical expertise would be available, the exercise should not be attempted. It further re-emphasized that in view of the relatively limited experience of African countries with the PES and the known problems in carrying it out, countries which opt to include it in their census evaluation programme may use it to evaluate only the census coverage. Content error should be evaluated by techniques of demographic analysis and the PES should positively avoid being a vehicle for collecting supplementary information. (ECA, 1987)

Indirect Methods - Analytical Techniques

68. Among the indirect methods we have two broad categories - the external and the internal consistency checks. In the former, we compare the enquiry result with other external evidences to arrive at the relative acceptability of the former. In the internal consistency checks, we analyse the data and check whether a broad picture of the reliability of the various information collected emerges. The reliability or otherwise is measured by the criteria of consistency and/or

covergency of the parameters derived from the data and other tools of analyses of data like the age ratio, sex ratio and joint score methods of age-sex data evaluation, consideration of growth rates, survival ratio, child women ratio, dependency ratio, proportions under various ages by sex, mean age of population (by sex), mean age of fertility, mean age of mothers at birth of children and also progression of age data over the age range. Vital parameters derived from the data by the application of diverse analytical tools, especially developed for the analyses and interpretation of defective and incomplete data are compared for consistency (i.e, not conflicting with each other and with other evidences) and convergency (i.e., lying within a reasonably narrow margin of variation). Stable and quasi stable models, analogy etc and similar techniques also are powerful tools for data evaluation and error detection.

69. Even though, as indicated earlier, indirect methods are not very powerful in the evaluation of coverage, still under certain conditions these tools could be used with care to arrive at possibilities of errors in coverage in addition to content error evaluation.

External Consistency Checks

70. One of the most important external consistency check is through the use of what is called the 'balancing equation.' Since a population changes by births, deaths and migration, this method utilises any information on these to appraise population figures. For example, if there are two enumerations, one could utilise data from vital statistics and migration figures to obtain a balancing equation:

$$P_1 = P_0 + B - D + I - E, \text{ where}$$

P_1 and P_0 are respectively populations at the second and first enumerations. B, D, I and E are respectively the births, deaths, immigration and emigration figures during the interval between enumerations. This equation can be applied for age-sex groups provided such detailed data are available.

71. The table below compares, for each sex, the 1983 enumerated population with the expected population based on the 1972 census enumeration, registered births and deaths, and total international arrivals and departures for Mauritius.

Table 2.3 Intercensal population change, 1972-83

	Male	Female	Total
Enumerated population 30.6.72	413,580	412,619	826,199
Births July 1972 - June 1983	126,186	121,944	248,130
Deaths July 1972 - June 1983	41,433	31,001	72,434
Arrivals July 1972 - June 1983	908,066	674,062	1,582,128
Departures July 1972- June 1983	935,937	688,624	1,624,561
Expected population June 1983	470,462	489,000	959,462
Enumerated population	481,368	485,495	966,863
Excess of enumerated over expected	+10,906	-3,505	+7,401
Excess as % enumerated 1983 population	+2.3	-0.7	+0.8

72. It is observed that the total enumerated population is about 7,400 more than expected. The evaluation of the 1972 census had shown a deficit of about 10,000. Their survivors in 1983 would be about 8,500. It is therefore reasonable to assume that, on the whole, the 1983 enumeration has been better than that of 1972. However, whilst the deficit in 1972 was observed in both sexes, the gain of 7,400 in 1983 is made up of a gain of 10,900 males and a loss of 3,500 females. This finding appears not only unacceptable on its own, but also inconsistent with the estimated deficit of 5,900 males and 4,100 females in 1972. There seems to be a problem with the sex distribution of one or more of the components entering the balancing equation.

73. A systematic sample of 5,000 census questionnaires, selected from 55 enumeration areas chosen with probability proportional to size, was checked for errors in coding of sex. The sample, representing about 2.5% of all questionnaires, indicated that sex had been properly coded at the census. Coding of sex on live birth and death cards was also checked on a sample basis by matching the name of the person with the sex code on computer print-outs of births and deaths records. Again there was no indication of erroneous coding or punching. It was finally found that the problem was coming from the international migration data. Sex is not asked explicitly on the embarkation-disembarkation card and has to be deduced from the name of the person if the title (Mr., Mrs., Miss) is not clearly indicated.

74. More reliable results are perhaps obtainable if the analysis is carried out for Mauritians only, representing 99.5% of the enumerated population. The advantage of considering Mauritians only is two-fold: firstly, the international passenger traffic data are more reliable, because sex coding is better for Mauritians; and secondly, the flow data can be made to refer only to the Population that really matters. All births and deaths relate to the Mauritian Population, and it is reasonable to consider only mauritian arrivals and departures

in the balancing equation, especially since total arrivals and departures are disproportionately large when compared to the Population. For about a million Population, with 250,000 births and 72,000 deaths for the intercensal period, there were more than one and a half million total arrivals and about the same number of departures. The number of arrivals of Mauritians only was 327,000, that is about 21% of all arrivals, and the number of departures was 372,000, constitituting about 23% of all departures.

75. The balancing equation method applied to the Mauritian population only shows (Table 2.4) that the enumerated Mauritian population is about 10,400 in excess of the expected Population, the gain being 5,300 for males and 5,000 for females. This sex breakdown looks more acceptable than the sex breakdown obtained above with the total population. Although these figures for Mauritius only are not strictly comparable with the findings of the 1972 evaluation on the total population, it is interesting to note the close agreement between the 1983 gain of 5,300 and the 1972 deficit of 5,900 for males, and the 1983 gain of 5,000 and the 1972 deficit of 4,100 for females. Generally speaking, the balancing equation method shows that the 1983 enumerated population is consistent with the 1972 census data, and with vital registration data, and migration statistics, if one allows for the observed deficit of 1972 and also that the coverage has significantly improved.

Table 2.4 Intercensal change in Mauritian Population, 1972-83

	Male	Female	Total
Enumerated Mauritian Population 30.6.72	410,696	409,979	820,675
Births July 1972 - June 1983	126,186	121,944	248,130
Deaths July 1972 - June 1983	41,433	31,001	72,434
Arrivals of Mauritians July 1972- June 1983	207,859	119,517	327,376
Departures of Mauritians July 1972- June 1983	229,816	142,468	372,284
Expected Mauritian Population June 1983	473,492	477,971	951,463
Enumerated Mauritian Population	478,814	483,010	961,824
Excess of enumerated over expected	+ 5,322	+ 5,039	+ 10,361
Excess' as % of enumerated	+ 1.1	+ 1.0	+ 1.1

76. Vital registration data on births by sex and deaths by age and sex are available for the inter-censal period, as are also international arrivals and departures by age and sex. It has therefore been possible to survive the 1972 census population to obtain the expected 1983 population by sex and single year of age. The exercise has been done for both the whole population and for Mauritians only. Comparison of the enumerated with the expected population is shown only for the Mauritian population, because of the problems in migration data for non-Mauritians.

77. It is seen that the census population is larger than the expected survivors of the 1972 population except for ages 0-4, 5-9, 45-49 and 55-59 for males, and ages 0-9, 50-54 and 65-69 for females. Underenumeration is probably the explanation for the young ages, but otherwise this observation confirms finding of evaluation of 1972 census that the deficit was not restricted to specific age-groups. The deficit of males at ages 45-49 and of females at ages 65-69 may be due to slight overenumeration in the corresponding ages noted in 1972. The same explanation does not seem to hold for the deficit of females aged 50-54, this is probably due to some over statement of age especially since the deficit in ages just below 55 is compensated for by gains in ages 55 and just above. Hence, if allowance is made for the deficit noted in 1972, and for some age errors which must be present, then the enumerated and expected population show a striking consistency with each other (Table 2.5). For more details refer to analytial report Vol 1(Maritians 1985).

Table 2.5 Comparison of 1983 enumerated Population with expected Population based on 1972 census (five Year of ages and by sex) - Mauritian population

Age Group	Male			Female		
	Enume-rated	Expected	Diffe-rence	Enume-rated	Expected	Diffe-rence
All ages	478814	473492	5322	483010	477971	5039
0 - 4	55981	56941	-960	54911	55569	-578
5 - 9	52261	52291	-30	51424	51877	-453
10 -14	47953	47782	171	46489	46206	283
15 -19	57303	56209	1094	56193	55774	419
20 -24	52972	52601	671	52069	51441	628
25 -29	44533	43005	1528	44450	44448	2
30 -34	39021	37665	1356	38456	37573	883
35 -39	26779	25694	1085	27624	26104	1520
40 -44	19786	19329	457	20352	19846	506
45 -49	19161	19411	-250	19439	19225	214
50 -54	16053	15846	207	15961	16574	-613
55 -59	17209	17670	-461	17449	16631	818
60 -64	11793	11637	156	12634	11831	803
65 -69	8593	8588	5	9947	10048	-101
70+	9416	9123	293	15532	14824	708

78. Most of the countries in the region however, lack such detailed data but approximations could be applied as in the following modification of the equation

$$P_1 = P_0 + NG + MG, \quad \text{Where}$$

NG and MG denote respectively natural and migration growths, implying that instead of the four components we need only the net effect of natural and migration growth. Yet another modification is when migration is negligible giving the equation:

$$P_1 = P_0 + NG, \text{ ie } NG = P_1 - P_0$$

$$\text{Hence } \frac{NG}{nP_0} = \frac{P_1 - P_0}{nP_0} = r_A$$

(arithmetic growth rate per year, if n is the inter censal interval)

$$\text{or } \frac{2NG}{n(P_1 + P_0)} = \frac{2(P_1 - P_0)}{n(P_1 + P_0)} = r_G$$

(approximate geometric growth rate per year, if n is the inter censal interval)

In all cases the equation can be considered for age-sex groups, provided such detailed data are available. Unfortunately such information are rarely available and one can apply the method only for the total population,perhaps by sex.

79. For instance, the non Libyan population at the censuses of 1964 and 1973 were 49000 and 210000 and immigrants and emigrants during the period respectively 2367000 and 2223000. Vital statistics are not available for the entire period, but for the period 1972-73 the natural growth rate is known from registration figures and is 1.4%. since the age-sex balance was better in the period previous to 1973, it is possible that the growth rate was higher. The balancing equation implies a growth of 1.9% which looks reasonable and hence the figures look consistent.

80. For Ghana between 1948-70 and for Libya between 1954-73, three census figures are available and the following table (Table 2.6) shows the growth rates. Migration could have affected the values but only in the case of libya, do we have figures but not by sex.

Table 2.6 Intercensal Growth Rates (Annual Geometre: Ghana and Libya (Libyans)

	Ghana		Libya		Libya (adjusted for migration)	
Female	1948-60	1960-70	1954-64	1964-73	1954-64	1964-73
Male	.0417	.0225	.0385	.0332	.0365	.0333
Female	.0418	.0263	.0378	.0354		

81. Obviously, the growth rates in the earlier period are higher than in the latter period, contrary to expectations unless out migration was heavier in the latter than the former period or immigration heavier during the first than the second period. In Ghana it is known that migration was heavier between 1948/1960 and in Libya we have adjusted for migration. Thus it indicates that the 1948 census of Ghana was an underenumeration as compared with census of 1960 and this is admitted. In Ghana between 1960 and 1970 there was an exodus of Non Ghanaians and this is responsible for the observed lower growth rate. Actually the growth rate of the Ghananian Population during 1948-70 was very consistent and thus indicates that the enumeration were resonably good. In Libya, the 1954 census was the first attempt in a vast and climatically difficult country and a huge underenumeration was known to have occured. In 1964, on the other hand, due to an impending election, there is reported to be an overenumeration. The growth rates depict the picture very well.

82. A related method using survival ratios of groups (either in specific ages or in groups of specific ages) is yet another tool, but will be considered later.

83. Other types of external consistency checks are those based on comparisons of specific age-sex groups in one enumeration with data from other independent sources. For example, children of school age, schooling population, persons of voting age, working age etc. could be compared with data from school statistics, electoral rolls, labour bureau, rationing authorities, registration of vital events, other civil registers etc. However, one has to be cautions in interpreting the findings, as sometimes, it is known that the independent external sources used might be subjected to similar types of errors as the enumeration or may have different types and magnitudes of erros and biases and hence may not be comparable and will lead to wrong conclusions.

84. As an illustration, the non Libya workers from 1973 census of 117524 is compared with statistics from manpower division which was 117344 on 20 June and 128000 on 20 September 1973. Census was on 1 August and the figures agree very closely. Another data is primary school children. Ministry of education gave 259729 boys and 196131 girls in schools in 1973 as against census values of 250556 boys and 184329 girls. Timing difference explains most of the discrepancy, which is anyway not large.

85. Comparing the student population by age and also by grade both from census and school statistics for Mauritius given in the following table, we can notice that the differences are marginal. However, a salient point to note is that the census figures are consistently higher than the school statistics, when considering age but is the reverse while considering grade data. The discrepancy is explained by the fact that school

statistics includes children under 5 and over 20 between grades P1 to S6. Also the timing of the two systems may not be strictly comparable. But the differences are quite small.

Table 2.7. Comparison of school and census statistics, 1983

Age group	Male		Female	
	School Stat.	Census	School Stat.	Census
05-09	50496	51438	49686	50632
10-14	37769	38082	34226	34992
15-19	18840	19456	16160	16758
05-19	107105	108976	100072	102382
05-12	75045	76392	72534	74130
13-19	32060	32584	27538	28252
Grade				
Primary (I-VI)	67509	67156	65746	65155
Secondary (I-VI)	40057	39916	35906	35685

86. It is important to point out that all the above methods are not conclusive and even if perfect agreement is shown between the two sets of data, it cannot be concluded that they are acceptable. This is because of compensatory errors. These are necessary conditions for acceptability of data but are not sufficient. One has to carry out other exhaustive tests before final conclusion can be drawn. Some of these will be described below, but they too are not entirely free from similar defects.

Internal Consistency Checks

87. Among internal consistency checks are those based on patterns in age,sex and other data,reasonableness or acceptability of some of the observations and derived parameters over time and space etc.

Age (Vertical Consistency Checks)

88. One of the most important items collected in all demographic enquiries is that on age. Not only are demographic phenomenon like fertility, mortality and migration closely related with age, even socio-economic and other characteristics are highly correlated with age. Also,age can be used as a tool for estimating some of the demographic parameters and in developing societies without direct data on vital parameters, it is data on age which are generally utilised to arrive at such parameters. Age is also important in using indirect information on fertility and mortality like children ever born, children surviving, survival of parents, spouses, siblings etc.

89. Thus, not only is it true that age is one of the few information collected in all demographic enquiries, it is also one of the more thoroughly examined, analysed and adjusted characteristics. This perhaps is due to the apparent easiness of measuring techniques and the practical needs and uses of age data.

90. Though age is an easy concept to understand, when it comes to measurement, there are several problems. First of all, age could be in completed years, age nearest or next birth day etc. what is a year? To a Western man or to one familiar with that system it connotes a solar year. But to a Muslim it means a lunar year and to a Chinese it may be according to their traditional calendar. To yet others, it may be quite different from all these. Again, what is a birth day? Very few people in developing societies celebrate birth days and little importance is put for the event. Age reckoning, if it exists, is based on other criteria like seniority, number of harvests since birth, seasons, floods of rivers, position of planets and so on.

91. Even though there are tremendous variations in age reckoning among societies, it is possible to convert (approximately) age in one system to that in another. Thus the real problem in age estimation is not the different systems adopted, it is ignorance or indifference as to when a person was born. Deliberate misstatements of age due to one reason or the other, is yet another dimension, and this could create problems.

92. When ages are not known to respondents, they are estimated by them or by others (including the enumerator). The first approximation is to put persons into one of the appropriate decennial group-thus a large number are reported as aged 10, 20, 30 years. When efforts are made to estimate age more closely, then mid point of the range is used culminating in persons being reported as 15, 25, ... years. This results in digit 5 becoming the second preferred digit. Still further approximation within Five year ranges brings in reported ages ending in digits 2 and 8 with minor peaks at ages 4 and 6. This is also brought in by avoidance of odd digits like 1,3,7 and 9. This kind of digit preference or age heaping has been noted in most countries in the region. However, the tempo varies over the age range-becoming more and more as age advances, so much that very few are reported at ages other than 60, 70, 80 etc. In Ghana, for example, in 1984 around 30% of persons are reported with ages ending in digits 0 and 5 but among these aged 60-69 it was near 60%.

93. Many methods have been tried to arrive at better quality of age data. One method is to ask for age next birthday or age nearest birthday. Experience with such dodges has been to shift preference to other digits. Year of birth brings in preferences in digits ending in 0 and 5 in year reported. Another problem in asking for year of birth is the possibility of several 'not stated' persons, as observed in 1974 sample census of Zambia. At the same time in 1969 census of Zambia, year of birth brought in digit preference for 9 and 4 in ages as year of birth was rounded. However, if enumerators are cautioned about digit preference, then it may lead to complete omission of ages ending in 0 and 5, as noted in a demographic

survey of Guinea. Use of historical calenders, documents etc. can bring in yet another type of digit preference or age preference or age preference pattern. As an instance, in Libya during the 1973 census, ages were estimated either from a document issued in 1964 or were estimated otherwise. Anlysis showed that whereas the former produced digit preference for 9 and 4 the latter had the usual preference for 0 and 5. Even though there was error in the data so obtained, one important finding was that for young children this method of using document might not be too much in error and hence in the future this method might improve data quality. Also there will be more consistency on age data over time even for the others whose ages in the document might be in error.

94. In addition to digit preference and age preference resulting from use of documents, calender of events etc., there are certain cases of aversion or preference for certain ages. For example, usually age 13 is avoided and instances of preference for age 30 are not rare. In the Chinese culture the word for digit 4 (cha) also means 'death' and hence ages ending in digit 4 are avoided. Such aversion has been noted in Korea, Japan and among the Chinese population of Mauritius in the past. The use of historical calenders also may result in large number of persons being reported as born during specific important events. Certain legal rights and privileges bring in reporting ages in these specific ages and sometimes a cut off point either in data collection or in certain categories of social groups might result in ages being shifted to that age from neighbouring ages.

95. For instance, in many countries the age at entry into school is 6 years and a marked tendency for a large number of children to be reported as 6 years is noted. Again, the cut off point for fertility enquiry being 49 years, there is a marked tendency for shifting large number of women to the next age i.e. 50 years to avoid the extra labour. Such shifting of persons either below or above certain specific ages has been reported especially in regard to detailed queries on labour force, marriage and fertility. Thus single year of age data are affected by several types of errors and biases, but all the same it is highly recommended that every effort be made to collect such information.

96. To detect errors in single year of age data, we can use graphical method or algebraic methods like Whipples, Bachis, Myers, Carrier, Ramachandran or similar indices. Since, age estimation error comes in at older ages and selective under reporting vitiates data at young ages, it is recommended to deal only with the age range 10-69.

97. Whipples index is obtained by summing the age returns between ages 23 and 62 inclusive and finding what percentage is borne by the sum of the returns of years ending with 0 and 5 to one fifth of the total sum. The result would lie between 100

and 500. The method can be modified to seperately study the incidence of preference for digit 0 and 5 by including in the numerator only the sum of ages ending in digits 0 or 5. These sums are then divided by one tenth of sum of ages 23 to 62. The procedure is illustrated using data from 1984 census of Ghana. The digit preference index (DPI) introduced by the US Bureau of the census, is a simple method of studying the incidence of preference or otherwise of each of the digits 0,1....9. It consists of aggregating populations with end digits 0,1,....9 seperately and comparing it with the sum of population in the total are range, usually taken as between 10 and 69. Thus the DPI is a better method than Whipples, in that one is able to discern the preference/rejection of each of the digits 0,1,...9.

98. However, the DPI has the defect that since the age range considered is 10-69 and this range has the digit 0 as the starting digit, populations ending in age 0 will automatically be larger than those in end digits 1,2,...9. Myers introduced his index (MI), to take care of this lacuna in the DPI. The method consists of calculating the digit preferences within 10 age groupings having starting digits 0,1,...9 respectively. Thus the procedure to calculate MI is the same as that for DPI, but involves calculation of 10 digit preference indices within age ranges 10-69 (as for DPI), 11-70, 12-71, ... 19-78 to assure all positions to all the digits 0,1,..2. If one wished to restrict oneself to age range 10-69, then one should start with 10-59, and proceed to 19-68. The procedure can be presented in a simple fashion as follows:

Let A_0, A_1, ... A_9 be the population with end digits 0, 1, ...9 in the age range 10-59 and B_0, B_1, ...,B_9 be the corresponding populations in the age range 20-69, then the blended sum at digit i (i=0,1, ...9) is obtained by the formula:

$$C_i = (i+1)\, A_i + (10-i-1)B_i,\ i = 0,1, ...9$$

The blended percentage at digit i is then,

$$100\, C_i/(C_0 + C_1 + \ldots C_9)$$

The Myers index is defined as the sum of the absolute deviation of the blended percentages from the expected 10%.

99. Sometimes it is also defined as the absolute value of the sum of positive (negative) deviations only. In such case the value will be half of that obtained by summing all the absolute deviations. In the former, the index will lie between 0 and 180, whereas in the latter case it will be between 0 and 90.

100. Even though the Myers method removes the bias introduced because of the starting digit in the DPI, still one can easily note that the method suffers from the fact that it is insensitive to the structure of the population and hence cannot be used for comparative purposes between populations. A method which standardises for age structure is that proposed by

Ramachandran. In this method, instead of absolute numbers A_i and B_i used by Myers method, corresponding proportions within respective age ranges are utilised to obtain the blended values.

101. Illustrations of the calculation of digit preference and Myers indices are given in Table 2.8 utilising the single year of age sex data from the 1984 census of Ghana.

102. From the single year age data from table 2.8 we can obtain the population aged 23-62 as:

71481 + 92147 + ...+ 63302 + 8333 + 15594 = 2042077

Population aged 30,40,50 and 60 = 136241 + ... + 63302 = 375439

Population aged 25,35,45 and 55 = 118513 + ... 36053 = 346020

Whipples index (digits 0 and 5) $= \dfrac{375439 + 346020}{2042077} \times 500 = 176.7$

Whipples index (digits 0 only) $= \dfrac{375439}{2042079} \times 1000 = 183.9$

Whipples index (digits 5 only) $= \dfrac{346020}{2042077} \times 1000 = 169.5$

103. Even when data are presented in five year or other age groupings, certain undulations are noticed for several reasons. Sometimes there are omissions of persons of certian age groups and other times, this may be due to pushing persons across critical age boundaries for one reason or the other.

104. One of the commonest observations in developing countries is the relative shortage of infants and children in enumeration. Deliberate omission, misunderstanding of the scope of enumeration, age estimation error etc could be the reason. Some duplication could result from de jure counting in extended family systems. Direct evidence of this phenomenon is few and difficult. In the 1974 census of Liberia, the PES indicated that the 0-4 age group was underenumerated to the extent of 5.4% for female and 1.3% for males. The indirect method is to use reverse survival technique to estimate births in the past year or past five years and compare with estimated birth rates. Calculation of sex adjusted birth rate may be better in view of effect of age distribution. If mortality estimate is not known, an appropriate model life table with reasonable life expectation may suffice for practical purposes. For one year projection backward, perhaps rate of growth might work. The estimates are not very sensitive to mortality level, if the entire population is carried backward. Yet another tool is that based on stable models or analogy. For this purpose one needs estimates of fertility and mortality or growth rate and one of fertility/mortality parameters.

Table 2.8

Digit Preference (DPI) and Myers Indices (MI) - Ghana 1984, Males

Digit	10-19	20-29	30-39	40-49	50-59	60-69	10-59	20-69	10-69	%	Dev.
0	188801	142023	136241	96500	79396	63302	642961	517462	706263	18.3	8.3
1	115176	81261	40385	23133	17196	8333	277151	170308	285484	7.4	-2.6
2	190396	97078	84163	52998	33161	15594	457796	282994	473390	12.2	2.2
3	135373	71481	38831	27442	16927	8907	290054	163588	298961	7.7	-2.3
4	145076	92147	52062	25969	26575	11604	341829	208357	353433	9.1	-0.9
5	157714	118513	96661	94793	36053	36522	503734	382542	540256	14.0	4.0
6	137661	82397	57983	30616	26807	7490	335464	205293	342954	8.9	-1.1
7	102437	81995	39741	23999	13647	7988	261819	167370	269807	7.0	-3.0
8	146944	94780	54478	43084	19766	12457	359052	224565	371509	9.6	-0.4
9	91843	55900	33490	24727	10811	6049	216771	130977	222820	5.8	-4.2
								Sum	3864877		29.0=DPI

Digit	Age Group Sum:10-59	Coeff	Product	Age Group Sum:20-69	Coeff	Product	Blended Popn.	Percent %	Dev. from 10%	Abs. Dev. from 10%
0	642961	1	642961	517462	9	4657158	5300119	17.1	7.1	7.1
1	277151	2	554302	170308	8	1362464	1916766	6.2	-3.8	3.8
2	457796	3	1373388	282994	7	1980958	3354346	10.8	0.8	0.8
3	290054	4	1160216	163588	6	981528	2141744	6.9	-3.1	3.1
4	341829	5	1709145	208357	5	1041785	2750930	8.9	-1.1	1.1
5	503734	6	3022404	382542	4	1530168	4552572	14.7	4.7	4.7
6	335464	7	2348248	205293	3	615879	2964127	9.6	-0.4	0.4
7	261819	8	2094552	167370	2	334740	2429292	7.8	-2.2	2.2
8	359052	9	3231468	224565	1	224565	3456033	11.1	1.1	1.1
9	216771	10	2167710	130977	0	0	2167710	7.0	-3.0	3
								Sum:	31033639	27.3=MI

105. The reporting of age of males and females differs when they reach puberty and post puberty ages. Young unmarried girls usually are reported as younger and those married and or have children are reported older, but with a noted tendency to be within the reproductive age. For males, the tendency is to exaggerate age of adults. Deficits in reproductive age for males also might result from their mobility, migration, work status etc. Omission of young unmarried girls is noted in some societies in Northern Africa. As marital status and fertility status are blamed for affecting age data, they conversely are affected by age misreporting as can be noted from the following table.

106. Relatively high parities at ages 20 and 25 and lower value at 40 is suspected to arise from relating parity with age. Single year tabulations of proportion ever-married, percentage of those entering first marriage and other similar tabulations of characteristics by single years of age would show up the effect of these factors on age reporting and vice-versa.

107. An idea of age exaggeration within the reproductive age span by females can be gauged from the comparatively very high values of the mean age of fertility schedule and of mother. In most of the countries in the region, they are near 29 or 30 which look too high in view of early and universal marriage, early child-bearing and little child spacing. Tabulation of proportion married, birth rate, parity, widow (ers) all could point to age reporting errors. We have already noted the situation in respect of parity. The understatement and exaggeration of ages according to marital status and parity is given in the following table based on data from Nigeria.

Table 2.9 Reported Average Parity (P), Ghana 1971 (Ashanti Region)

Age	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
Parity	0.02	0.07	0.22	0.45	0.65	1.11	1.45	1.72	2.06	2.41	2.84	3.07	3.46	3.91	4.08
Age	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
Parity	4.66	4.81	5.33	5.45	5.84	6.00	6.16	6.46	6.71	6.94	6.71	6.91	7.13	7.65	7.55

Source: Table 5 from D.C Ewbank, Age misreporting and age selective under enumeration Sources, patterns and consequences for demographic analysis, Report 4, National Academy Press, Washington, D.C., 1981.

Table 2.10 Percentage of age misreporting by age, marital status and Females parity-aged 15-24 years, Nigeria, 1969

True age	Reported age*	Marital status and parity		
		Single parity=0	Married, parity=0 & Single, parity=1	Married, parity=2+
15-19	Exaggerated	7	16	27
	Same	83	79	70
	Understated	10	5	3
20-24	Exaggerated	5	11	14
	Same	81	81	77
	Understated	15	8	9

* Ages exaggerated or understated differed from the true age by at least one five year age group.

Source: Caldwell, J.C., and Igun. A.A., An experiment with census type age enumeration in Nigeria: Population Studies Vol. 25(2), 1971.

108. It is clear that single women aged 20-24 tend to be reported in a younger age group and women aged 15-19 and to a lesser extent those aged 20-24 tend to be reported at older ages if they are married or have at least one child.

109. One of the usual phenomenon in age reporting is the exaggeration of age by persons especially at older ages, because of prestige, physical appearance, fertility status etc. However, for females, the tendency is to report a younger age so as to be still included in the reproductive age group. The high proportion of old persons found in developing societies inspite of the rather high levels of mortality is indicative of the exaggeration of ages. Another idea of age exaggeration can be obtained by calculating the mean age of the population. In several instances, it has been noted to be one or two years above the expected.

110. Yet another test for age exaggeration is by use of forward and reverse projections. In the forward projection, the old age population shrinks whereas in the reverse projection it becomes so large as to overwhelm even younger age groups.

111. It is reported that in the 1974 Liberian census a very large proprotion of the ages of individuals had to be estimated by the enumerators. Inspite of all efforts to estimate ages using historical or other calender of events, physical

appearance, physiological changes, social customs like age grade, circumcision and other grading documents and so on, still it is possible that a large number of cases remain where age estimation is impossible. Age not stated or unknown category will then be the result. So long as this group is not large, there is not much of a problem. In most cases, prorata allocation may suffice. If other informations are available to allocate these into specific age-sex groups, then this could be done. Cold and hot deck methods of imputation are well known. But it should be remembered that imputation may create problems in data analysis, interpretation and utilistation and hence should be avoided, as much as possible. As mentioned earlier, it is easier and perhaps more accurate, if information are collected at the household itself however rough they may be. Any strand of evidence may be helpful in estimating age and other characteristics in the field rather than in the office. As mentioned earlier, the more than 20% of not stated cases in Zambia 1974 made the data virtually useless.

112. When age data show undulations, they could be genuine reflections of past behaviour of fertility, mortality and migration. In most cases, if these factors could be assumed not to have much effect on specific age groups, then the progression of population from one age group to the next should be more or less smooth. An age pyramid distinguishing males and females would be a good graphical representation and would compare not only progression, but also sexwise differentials.

113. For quantification, the calculation of age ratios would be useful. Here the Zelnik or Ramachandran modifications could be utilized. Whereas the UN defined age ratio as the ratio of population in an age group to the average of population in the two adjacent age groups, the modification of Zelnik and Ramachandran used the population of the group being considered also in denominator. In cases where digit preference is predominant, the Ramachandran modification would work better and remove some of the biases in the other two methods, because while the Zelnik method gave equal right to the 3 age groups, the Ramachandran method assumed double weight for the mid-age group and hence equalised weights for all digits, thus removing digit preference biases. Also this index approaches 100 when the data is good, as it should, see table 2.11.

114. These ratios are calculated seperately by sex and of necessity the groups should all have the same number of individual ages in them. These ratios generally should not fluctuate too far from 100. Thus, an index of vertical consistency can be taken as the arithmetic mean of absolute deviations of age ratios from 100 and is called the 'Age Ratio Score.' Age ratios can be calculated by single year of ages also but these ratios will show more the digit preference than age shifting errors.

115. Sex is yet another most important demographic characteristics. Many demographic, socio-economic attributes are sex specific. Data classified by sex not only have analytical importance, many times they could be used as evaluative tools and a classification is essential in certain situations because of biological differences.

116. Even though the definition and classification of sex are easy and sex is generally ascertainable, still field data have their due share of problems arising from classification of data by sex. As mentioned earlier, there are several reasons why one sex or the other is either misreported, wrongly reported or not fully reported or even over reported.

117. The sex ratio of a population should not fluctuate from one age to another unless migration or other factors are responsible. Thus the calculation of sex ratios by age will give clues to defects in sex data by age. Sex ratio can be defined either as males per 100 females or as females per 100 males. Here we use the former definition. Some authors use the latter definition and one has to keep this in mind while interpreting data. The sex ratio at birth is one of the few stable parameters of a population and lies in a narrow range of between 100 and 108. It is stated that the sex ratio at birth in African countries lies in the range of 101-104. If only fertility and mortality have effect on a population, then the sex ratio of the population will not deviate much from some stable value. In most cases, male mortality is higher than female mortality and hence the sex ratio of the population will

Table 2.11 Calculation of Age, ratio, sex ratio and UN joint scores: Ghana 1984

			Age Ratios							Deviations (absolute values)		
Age group	Male	Female	U.N. Method Male	U.N. Method Female	Zelnik Method Male	Zelnik Method Female	Ramachandran Method Male	Ramachandran Method Female	Sex ratio	Age ratio (UN) M	Age ratio (UN) F	sex ratio
00-04	1015167	1014915							100.0			
05-09	1012787	989038	113.2	113.5	108.4	108.6	106.2	106.3	102.4	13.2	13.5	2.4
10-14	774822	728387	94.0	91.1	95.9	93.9	96.9	95.4	106.4	6.0	8.9	4.0
15-19	636599	609791	101.1	93.8	100.8	95.8	100.6	96.8	104.4	1.1	6.2	2.0
20-24	483990	572011	90.4	102.0	93.4	101.3	95.0	101.0	84.6	9.6	2.0	19.8
25-29	433585	511526	103.8	106.2	102.5	104.1	101.8	103.0	84.8	3.8	6.2	0.2
30-34	351682	391121	98.2	76.2	98.8	97.4	99.1	98.0	89.9	1.8	3.8	5.1
35-39	282353	301946	97.7	94.6	98.5	96.3	98.9	97.2	93.5	2.3	5.4	3.6
40-44	226042	247212	90.5	96.4	93.5	97.6	95.0	98.2	91.4	9.5	3.6	2.1
45-49	217219	210988	108.8	98.9	105.7	99.3	104.2	99.5	103.0	8.8	1.1	11.6
50-54	173255	179429	106.8	113.2	104.5	108.4	103.3	106.2	96.6	6.8	13.2	6.4
55-59	107084	105997	76.2	71.3	82.8	78.8	86.5	83.2	101.0	23.8	28.7	4.4
60-64	107740	118036	121.3	130.6	113.3	118.5	109.6	113.3	91.3	21.3	30.6	9.7
65-69	70506	74803							94.3			3.0
SCORE										9.0	10.3	5.7

Note: The age ratios are defined as follows:

UN method: $$\frac{200P_{x,\ x+4}}{(P_{x-5,\ x-1} + P_{x+5,\ x+9})};$$

Zelnik method: $$\frac{300\ P_{x,x+4}}{(P_{x-5,x-1} + P_{x,x+4} + P_{x+5,x+9})}$$

Ramachandran method: $$\frac{400\ P_{x,\ x+4}}{(P_{x-5,x-1} + 2P_{x,x+4} + P_{x+5,x+9})}$$

be somewhere between 94 and 98. If male mortality is abnormally high due to occupational patterns, then the sex ratio could be, say 93. On the other hand, when males have an edge over females in mortality through socio-cultural reasons, then the sex ratio of the population could be more than 98 and can go up to 101. Again, if mortality improvement occurs normally, then it is noted that the sex ratio would fall because females have an advantage over males in most cases. However, the decrease in sex ratio could only be slight.

118. Table 2.11 gives the sex ratio of the population of Ghana as obtained from the 1984 census. As expected, the sex ratio at age 0-4 is slightly above 100 and increased to 106.4 by age 10-14 and then declined. Male mortality is generally higher than female at almost all ages excepting perhaps the reproductive ages. At very young ages there could be excess mortality for either sex depending on socio-economic and cultural factors like child care etc. The low sex ratio at very young ages could also be a reflection of relative coverage or reporting of children. Fear of evil eye, preference for children of specific sex and other behavioural factors could also affect relative coverage of children. An important observation in African data is the relative shortage of males at reporductive ages. This could be due to the twin factors of pushing girls across age 10-14 to ages 15+ and women past reproductive ages to ages below 50 or age exaggetation by males, omission of males because of their mobility, work status or migration or deliberate under reporting of adult males because of conscription, poll tax, head tax etc.

119. Sex ratios by single years also can be calculated to highlight the relative preference or avoidance of specific digits by either sex.

120. To quantify the error, an index, called the 'Sex Ratio Score' is calculated by summing the absolute deviations of successive sex ratios and dividing it by the number of such deviations. Since sex ratio may not change linearly, it seems better to obtain an average of the absolute second difference appropriately weighted. A large score indicates that there is some error or bias in the data. From the data presented in table 2.11 the sex ratio score is 5.7.

121. Normally sex ratios should be calculated for only the age groups up to 60 or 69, as otherwise fluctuations would be tremendous. Also it is advisable that the age groups for both sexes should not only be similar, it is preferable to have the same five year or ten year age groups for both sexes for the entire age range.

122. The sex ratio score is an index of horizontal consistency of data and could be used to appraise data quality. Combined with the age ratios, it can be converted to a single index of data quality by appropriate weighting. The UN noted that the variability of sex ratios is much less than that for age ratios and hence apportioned a greater weight to this score than to age ratio scores. Obviously the age ratio scores were given equal weight. Based on studies on age - sex data in the 50's, the UN defined a Joint Score as:

$$\text{Joint Score} = A.R.M.S + A.R.F.S + 3\ S.R.S.,$$

where ARMS and ARFS are respectively the male and female age ratio scores and SRS is the sex ratio score.

123. Experience indicates that if the joint score is less than 20, then the data would be considered reliable. Between 20 and 40 the data may be usuable with adjustments, between 40 and 60 it is considered deficient and care and caution should be exercised in the use and interpretation of the data and massive adjustments might be necessary. Beyond 60, the data is considered grossly erroneous and it may be risky to utilize such data for any inference.

124. The joint score for Ghana 1984 from table 2.11 is

$$= 9.0 + 10.3 + 3\ (5.7) = 36.4$$

Since the score is between 20 and 40 we can conclude that the data is usable with appropriate adjustments.

Diagonal Consistency Checks

125. The age and sex ratios give the two dimensions-vertical and horizontal - in the evaluation of the consistency of data. When more than one enumeration is available there is yet a third dimension which can be considered, ie., diagonal consistency - based on survival ratios of cohorts. These ratios could be for specific age groups or for open ended ages. The latter type of ratios are not much affected by age reporting, but this becomes their weakness as a tool for evaluation. Survival ratios for single years of age could rather show the digit preference errors and only to a lesser extent the other age reporting errors.

126. The intercensal (10 year) survival ratios are given in Table 2.12 for various five year age cohorts from the 1969-1979 censuses of Kenya. The first point to note is the tremendous fluctuation of these ratios from one age to another for both sexes and especially for females, where we even observe an impossible ratio higher than unity. The ratios for the age group 0-4 for both sexes are too high and this is more for male than female contrary to experiences in other parts of the world. Another observation is the rather high values at older ages and comparatively low ones at middle ages.

Table 2.12 Ten year (intercensal) survival ratios, Kenya, 1969-1979

Age group in 1969	0-4	5-9	10-14	15-19	20-24	25-29	30-34
Male survival ratio	.9932	.9318	.8974	.9451	.8302	.8302	.9307
Mortality level	23.5	10.3	4.9	10.0	16.0	3.3	16.0
Female survival ratio	.9785	.9937	1.0334	.9934	.9169	.7912	.9146
Mortality level	20.2	22.0	24+	23.6	9.5	*	12.0

Age group in 1969	35-39	40-44	45-49	50-54	55-59	60-64	65+
Male survival ratio	.8682	.9431	.8161	.8131	.8713	.6477	.4259
Mortality level	10.1	21.1	11.0	14.9	23.6	15.2	23.1
Female survival ratio	.8352	.9459	.8211	.7875	.8140	.6617	.4683
Mortality level	5.6	19.5	8.2	9.6	17.2	13.7	24.0

* Less than the lowest available level, 1.

Table 2.13 Overall (Open ended) survival ratios, Kenya 1969-79

Age in 1969	0+	5+	10+	15+	20+	25+	30+
Male ratio	.8977	.8749	.8600	.8504	.8333	.8068	.8011
Level*	58.3	42.9	42.9	48.9	50.0	47.1	61.5
Female ratio	.9230	.9099	.8886	.8550	.8223	.7995	.8018
Level *	77.8	71.6	65.0	52.1	42.3	42.2	61.9

Age in 1969	35+	40+	45+	50+	55+	60+	65+
Male ratio	.7701	.7433	.6901	.6510	.6003	.4995	.4259
Level*	63.4	71.4	70.8	80.1	88.3	87.2	100.0
Female ratio	.7725	.7529	.6959	.6565	.6088	.5337	.4683
Level *	64.8	75.8	73.2	81.8	90.8	96.7	115+

* The levels are taken from table A.11 of Carrier N.H, and Hobcraft, J. Demographic estimation for developing societies, London, 1971.

127. There are two ways to analyse these ratios. The first one is to calculate the ratio of these values to corresponding values from a reasonable life table. Here for example, we could compare the survival ratios with corresponding values from North model level 13 or 14. For ease of comparison a ratio of these values would do well because, if the observed values are acceptable, then the calculated ratios should not deviate much from unity. Yet another method illustrated in the table is to write down the levels implied by these ratios from North model life tables. It can be noted that the ratios fluctuate, they are high for the very young and old ages and low for middle ages. For females, one value is even below the lowest available level and one above the highest and generally the female levels are lower than the male levels. To remove some of the age shifting errors we have presented also in Table 2.13 the values of overall (open ended) survival ratios for the same data.

128. It is evident that the wide fluctuation depicted by the cohort survival ratios have subsided implying that grouping makes the data much more smooth and acceptable. Here also the level at younger and older ages are much high and at middle ages lower. One value for female, ie., at age 65+ exceeds the highest level in the table.

129. In table 2.14 we present a picture similar to what is shown in table 2.12 by calculating the ratios of census to life table survival ratios (using an appropriate model life table and level of mortlity). We can note that the ratios deviate from the expected value of one, fluctuate from one age to another, are high at young and old ages, are more erratic for females and are on the whole not smooth. Any other level (higher or lower) would indicate very similar patterns. So is the case with use of another model life table (North versus say West).

130. There are several other ways of testing the acceptability of the survival ratios like those using logit transformations, calculating life expectation or life table values and so on. Some of these will be considered in the later section dealing with techniques of analytical manipulation of data.

Table 2.14 <u>Ratio of intercensal (Kenya 1969-79) to life table (North level 14) Survival ratios</u>

Age		0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44
Ratio	M	1.091	.977	.942	.976	1.013	.894	1.013	.960	1.069
	F	1.068	1.039	1.077	1.042	.970	.844	.985	.912	1.043

* Coale A. and P. Demeny, Regional model life table and stable populations, Princeton University Press, 1966.

Other Internal Consistency Checks -Structure of Population, Comparison with Models, Calculation of Ratios, Differences and so on

131. One attempt worth trying is to match observed age-sex data with models or analogies. Either the ratios of percentages at various ages or differences of cumulated percentages could be considered. In the first case, the sum of absolute deviations from unity could be the criterion, whereas for the latter, it could be the sum of absolute differences themselves. It may also be worthwhile to try if an improvement would result, if we consider the standardised absolute deviations using the expected percentages in the denominator. Also sum of squares of deviations of observed and expected percentages has been suggested.

132. Yet another approach would be to calculate ratios of population in adjacent quinary ages and study the progression. Undue fluctuations, deviation from unity, very low values and so on could be interpreted as evidence of errors in data. By a suitable modification of a method proposed by Stolnitz, these ratios could be converted to life table values and the overall reliability of the data could be assessed, based on the observed expectation of life.

133. Still another method is based on several types of ratios obtained from the reported data. Since a ratio is less susceptible to certain types of errors, this may not be an ideal procedure to follow in many cases, but one should manipulate the data in several ways before one can come out with some ideas about its quality. For example, child woman ratio, child adult ratio, dependency ratio and so on could be computed and compared with model values. Sometimes these values can be calculated for the two sexes seperately and compared for differential errors. Some results from the application of the method to the 1973 Sudanese census data are presented in table 2.15. The closeness of male values to expected than female values in most cases indicates the poorer quality of female age reporting, inspite of the fact that the ratios, obliterate to a certain extent, some of the inherent errors and what is shown is a residual. Most of the ratios are higher than the model values and point out the inconsistency between the derived demographic parameters and the age-sex structure.

134. Even though these methods are labourious, it looks that the computer can be used for carrying out the matching process, once the inputs are fed into the computer. Several other ratios or proportions could be tried.

Table 2.15 Special Ratios Based on 1973 Census of Sudan and from Models*

Ratio	Sex	Sudan	Model	Ratio	Sex	Sudan	Model
P_{5-14} / P_{5+}	M	.358	.320	P_{0-4} / P_{15-44}	M	.441	.419
	F	.333	.313		F	.389	.412
P_{5-9} / 20–49	M	.512	.397	P_{0-9} / P_{15-49}	M	.815	.694
	F	.444	.389		F	.724	.681
P_{0-14} / P_{15+}	M	.808	.792	Dependency ratio $\frac{(P_{0-14}+P_{65+})}{(P_{15-64})}$	M	1.079	.944
	F	.809	.763		F	.955	.936

* The model was selected after several trial and error and is the Coale-Demeny North Model Stable Population with level 11 and growth rate of 2.5%.

135. Another similar procedure worth trying is that suggested by Brass (Brass et al, 1968). In this method the age equivalents of the deviations of the cumulated percentages of the reported population at ages upto age 5,10,15,.... are plotted and the highest and lowest deviations are utilized to select an upper and lower stable population. It seems that an average of the upper and lower levels might fit the data on hand, in many cases.

Methods Based on Analytical Manipulation of Age-Sex Data

136. We have already referred to the application of projections (reverse and forward) for the evaluation of data. There is yet another important application of forward projection technique suggested by Coale and Demeny. Various life tables are used to project cohorts such that the projected figures tally with enumerated figures at least approximately. Fluctuation in the levels from age to age and by sex would indicate the types and magnitudes of errors. This method also can be used to arrive at a level of mortality depicted by the enumerated populations and its consistency with other information would be an index of data quality. Since survival ratios are readily available for five year age groups and in five year intervals, it is better to deal with such age groups and periods which are multiples of 5. See chapter 3 for details.

137. A similar procedure proposed by Ramachandran and Nair (Ramachandran and Nair, 1970) is based on tables of overall survival ratios which could be compared with observed ratios. The additional information needed is the intercensal growth rate and also some idea of changes therein. Tables have been presented using Coale-Demeny family of life tables. Carrier and Hobcraft present such tables based on Brass model life tables. If the intercensal growth rates are known and is either constant over time or its variation over time is known, then this method seems easier than the method of Coale-Demeny requiring several trial and error projections.

138. An allied method which can be used to evaluate data is due to Hardy. In this method the difference between those aged say, 10 years and above in the second enumeration and the total population at the earlier count ten years ago is used as a measure of mortality from which the total number of deaths and consequently death rate is obtained. This is done by sex. It seems possible to extend this method and deal with all ages: those aged 5 years and above, those aged 10 years and above and so on, at least for 5 or 7 consecutive quinary ages. The acceptability of the death rate will indicate the quality of data by age and sex.

139. An important method for checking consistency of age-sex data is that based on cumulated age distribution. We have already noted its use in comparing observed with model values. Another application when approximate stability can be assumed, is that based on growth rate, gross reproduction rate, mortality level or any other fertility/mortality information. The method can be used to estimate mortality level given the values of either the growth rate, or gross reproduction rate GRR. Conversely, given a mortality level, estimates of γ or GRR can be obtained. The acceptability of the estimates and (growth ratio), the fluctuations over age and sex will all give ideas about data quality. See chapters 3 and 4 for details.

2.5 Some Methods of Data Adjustment

140. Based on the findings of the evaluation and appraisal of the various types of basic demographic data, it is obvious that before they can be utilised for projections and other analytical purposes, it is necessary to adjust and smooth the data to remove some of the inherent errors, biases, deficiencies and other defects in the data. However, the adjusted data should not only be smooth, they should also be consistent with each other and with demographic, socio-economic realities in the past. Thus consistency with the various demographic parameters should be vouchsafed and not sacrificed at the altar of the so called smoothness of data. Also, the adjustment should be minimal and be specific to take care of the types, varieties and magnitudes of the errors and deficiencies.

141. There are several methods of achieving smoothness in data. in addition to using information from extraneous sources like post enumeration surveys, mathematical methods are the easiest and need only very little information and can be carried through mechanically. But, one should be cautious in such simplistic models submerging some of the peculiarities inherent in the data. Moreover, the types and magnitudes of errors and biases may not be the same or even similar from one age to next or one section to another or even over time. Thus one mathematical formula might compromise some of the specialities and bring forward other biases into the data. Alternatively, there are several demographic techniques-borrowing from other experiences, models etc. and sometimes a combination of these-mathematical and demographic which might work well in many situations.

Adjustment of Age Sex Data

142. Since age-sex data are paramount in demographic analyses, most attention has been devoted to adjust these. Two broad types of methods will be considered-the direct and the indirect. In the direct method, additional information-more reliable than the one being adjusted-is used to adjust the former. Usually, age-sex data from post enumeration surveys being based on more carefully planned and collected information might be used for adjusting the observed information.

143. Table 2.16 presents the effect of adjusting the male age data from the 1974 census of Liberia. It can be noted that not only did the age distribution become nearer to stable model, the estimated birth rates also were close. The sum of squares of deviations from 1 of the ratios of stable to observed was reduced. Thus if PES or other extraneous information of a better quality are available, it seems useful to use these to adjust observed age-sex data. Adjustments can be for under/over enumeration in count, age misreporting, sex wise distribution of population and any other types of deficiencies in data. Hence detailed information on marital status, migration status, relationship and other relevent factors affecting data quality should be collected both in the enumeration and the PES or other sources.

144. Ewbanks succintly puts the problem of adjusting census data using PES information as follows: PES studies usually provide only a good measure of the variance of the content error. This suggests that frequently the results of a PES will not be very useful for adjusting the data prior to analysis, since such an adjustment requires estimates of the biases in reported ages. For example, adjusting age data before applying stable population analysis requires an estimate of the net errors. For some types of analysis, even more information is needed; for example, adjusting individual observations for use in analysing differences among individuals requires information about gross errors and adjusting the data prior to applying the

Brass techniques for estimating fertility and mortality requires information about the characteristics of those who report age incorrectly. Unless the results of reinterviews can be regarded as correct, they will prove to be of little value in adjusting the data prior to the application of such analytical techniques.

Algebraic (Mathematical) Methods

145. Graphical methods or free hand drawing and the use of mechanical aids such as a French curve or a spline and weights might be tried, but the subjective nature of these should be borne in mind while interpreting the results. The oblique axis method of Carrier seems useful in certain cases. In this method the first step is to find the groupings most nearly correct (for this single year values are needed). For instance, when digits 0 and 5 are highly preferred, it has been noted that either the 2-6 or 3-7 groupings would give lower age ratios and sometimes also sex ratios. As a general principle, the groups should be chosen to minimise crossing of boundaries. This will frequently imply splitting a single year of age group, which may be achieved with sufficient accuray by dividing the number in the single year age into equal halves keeping in mind that the patterns of preference are likely to be different at different parts of the age scale.

Table 2.16
Evaluation of ajusted data by using PES information-Liberia, 1974 (Males)

Age	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44
C(a)	15.18	30.28	41.81	51.77	58.94	66.02	72.22	78.46	83.22
C(A)*	15.39	30.55	42.14	52.75	60.13	67.24	73.45	79.65	84.24
CBR	39.77	43.93	43.30	42.71	40.30	39.35	38.45	38.81	38.14
CBR*	40.39	44.43	43.80	44.07	41.83	40.90	40.13	40.63	39.98
CBR Mid range					4.53			3.41*	
Sum of squared deviations					.098			.081*	

* These pertain to adjusted values

Source: Ewbank. D.C., Age misreporting and age selective underenumeration: Sources, patterns and consequences for demographic analysis, National ACademy Press, Washington. D.C., 1981.

146. Since other types of errors are found at very young and older ages, it is suggested that this smoothing technique be restricted to ages 10 to 59 or 64 or at the most to 69 years. A minimum of about 30 years resulting in sufficient number of points for carrying out this procedure should be ensured.

147. The method consists of cumulating the population from the start of the age range (here 10 as an example) and continued by the five year groupings till the last group, say 69 years. Thus we get cumulated populations from 10 onwards for ages 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, assuming that the groupings are in the conventional five year age groups, 10-14, 15-19, ..., 65-69. If the age groups are in the 2-6 pattern, then the first age will be 12 and the first group 12-16 and the cumulated values will be for ages 12-16, 12-21, 12-26, ..., 12-66 or perhaps 12-71. Suppose 69 is the last age selected and 10 the starting age, then there are 60 single year ages in this range for which the cumulated population comes as P(10-69). An average for each age is thus: P(10-69)/60. This average need be calculated only approximately, say to the nearest ten or even hundred depending on the size of the population being considered. This average is denoted as 'A'. The next step is to calculate expected populations upto the various ages, 14,19,.........,69 by the formula: $Q(x) = P(x) - (x-9)\,A$ for $x = 14,19,\ldots,69$.

148. The Q(x) values are then plotted against x and this graph is the oblique axis ogive. The Q(x) values are then smoothed by visual inspection and a curve is drawn through the smooth values and are denoted as Q'(x). From these smoothed values, we revert back to P'(x) the smoothed values of the cumulated populations at various ages by the equation: $P'(X) = Q'(X) + (X-9)\,A$. From P'(X) we can then get the populations in five year age groups by differencing.

149. Graduation formulae like those proposed by the United Nations or a three point version of it, moving averages, Carrier-Farrag ratio method, Newtons halfing formula (quadratic interpolation), osculatory interpolation formula and other curve fitting methods are some of the other mathematical tools in data adjustment.

150. The United Nations five point smoothing formula is a moving average method utilising five consecutive five year age groups with the value to be smoothed in the middle and two groups above and two below it. The weights and method are given by the formula:

$$V_{x,x+4}=(1/16)(-U_{x-10,x-6}+4U_{x-5,x-1}+10U_{x,x+4}+4U_{x+5,x+9}-U_{x+10,x+14})$$

where $U_{x,x+4}$, ... are the reported populations at ages x to x+4, ... and $V_{x,x+4}$ is the corresponding smoothed value.

151. This formula can be conveniently denoted by the notation: $V_{x,x+4} = (1/16)\ (-1,\ 4,\ 10,\ 4,\ -1)$ where the weights -1, 4, 10, 4 and -1 are to be applied respectively to the group populations starting with the group 10 years younger than the group being smoothed and continued till the group 10 years older also is included.

152. A milder version of the above formula is represented by: $V_{x,x+4} = (1/36)\ (-1,\ 4, 30, 4,\ -1)$
In case the age groups are affected by migration, it is not advisable to include a large number of age groups in the smoothing at one time because then the formula will spread the migration into adjacent groups considering the true migration as an error in the data. In such cases, a three point version:

$V_{x,x+4} = (1/4)\ (1,\ 2,\ 1)$ is used where only three five year age groups are considered at a time with the population of the age group being smoothed being flanked by the preceding and succeeding five year groups.

153. The Newton's halfing formula or quadratic interpolation method is applied to split a ten year group into two five year groups. The assumption is that by grouping a population into ten year groups some of the undulations brought in by digit prefernce and other errors would be reduced. The grouping also could be in such a way as to include ages which are supposed to be affected. For instance with zero being predominantly preferred in most cases, the ten year groupings which could be considered are 2-11 or 3-12 which then would be split into 2-6, 7-11, or 3-7 and 8-12. Then by suitable interpolation, we can obtain the conventional ages 5-9, 10-14 ....

154. The formula for getting the first half of the ten year group is:
$P_{x,x+4} = (1/16)\ (P_{x-10,x-1} + 8\ P_{x,x+9} - P_{x+10,\ x+19})$.

The second half can be obtained either from the formula:

$P_{x+5,x+9} = (1/16)\ (-P_{x-10,x-1} + 8\ P_{x,x+9} + P_{x+10,x+19})$ or

by subtracting the first five year value from the ten year group value, i.e.,

$$P_{x+5,x+9} = P_{x,x+9} - P_{x,x+4}$$

155. Karup-King give interpolation coefficients to split even these five year groups into single year values.
Sprague gives similar multipliers for splitting ten year groups into five year or single year values using five ten year groups. The weights are:
-.0117, .0859, .5000, -.0859, .0117
to obtain the first five year group and
.0117,-.0859, .5000, .0859, -.0117 for the second five year group.
The Karup-King and Sprague methods are known as osculatory interpolation formulae.

156. The Carrier Farrag ratio method is yet another procedure for splitting a ten year group into two five year groups. In this method, appropriate ten year groups are calculated and ratios of consecutive values obtained by keeping the younger age groups in the numerator. For instance the first ratio can be:

P_{0-9} / P_{10-19}. Usually since the age group 0-4 is affected by underenumeration, it is advisable to exclude it. For the consecutive ten year groups, similar ratios can be calculated.

157. To obtain the smoothed value of the five year age groups, we take the fourth root of these ratios and apply them to the ten year groups as follows:

Let $K_1 = \sqrt[4]{(P_{5-14})/(P_{15-24})}$, then $P_{15-19} = K_1 \, P_{15-24}/(1+K_1)$

and hence $P_{20-24} = P_{15-24} - P_{15-19} = P_{15-24}/ (1+K_1)$
and so on.

158. Let us illustrate the apllication of some of these smoothing techniques to the data from Ghana. For instance, if we want the smoothed value of the population aged 20-24 we have to have several five year age groups above and below that age group depending on the formula used. Since the male age distribution may be affected by migration in addition to age reporting errors, we shall use the techniques to the female population.

159. From table 2.11 we have the five year age distribution for females. From the UN five point formula we have

(1)

$$F_{20-24} = \frac{1}{16} [-F_{10-14} + 4F_{15-19} + 10F_{20-24} + 4F_{25-29} - F_{30-34}]$$

$$= \frac{1}{16} [-728387 + 4(609791) + 10(572011) + 4(511526) - 391121]$$

$$= \frac{1}{16} [-728387 + 2439164 + 5720110 + 2046104 - 391121]$$

$$= \frac{1}{16} [9085870] = 567867$$

160. Similarly the mild version of the formula will give:

$$F^{(2)}_{20-24} = \frac{1}{36} [-728387 + 4(609791) + 30(572011) + 4(511526) - 391121]$$

$$= \frac{1}{36} [-728387 + 2439164 + 17160330 + 2046104 - 391121]$$

$$= \frac{1}{36} [20526090] = 570169$$

The three point formula will give:

$$F^{(3)}_{20-24} = \frac{1}{4} [-609791 + 2(572011) + 511526]$$

$$= \frac{1}{4} [609791 + 1144022 + 511526]$$

$$= \frac{1}{4} [2265339] = 566335$$

161. For the Newtons halfing formula we have to form ten year age groups with age group 20-24 included in one of the groups. Since population aged 0-4 is usually suspect, we start only from age 5-9 or 10-14. Let us form the ten year groups starting from 5-9.

$F_{5-14} = 989038 + 728387 = 1717425$

$F_{15-24} = 609791 + 572011 = 1181802$

$F_{25-34} = 511526 + 391121 = 902647$

162. Since the group 20-24 is in the second half of the group 15-24, we use the formula:

$$F^{(4)}_{20-24} = \frac{1}{16} [-F_{5-14} + 8F_{15-24} + F_{25-34}]$$

$$= \frac{1}{16} [-1717425 + 8(1181802) + 902647]$$

$$= \frac{1}{16} [-1717425 + 9454416 + 902647]$$

$$= \frac{1}{16} (8639638) = 539977$$

163. To use the Carrier-Farrag ratio method, we calculate

$$K^4 = \frac{F_{5-14}}{F_{15-24}} = \frac{1717425}{1181802} = 1.4532$$

and K = 1.098

Then

$$F^{(5)}_{20-24} = \frac{F_{15-24}}{1+K} = \frac{1181802}{2.098} = 563299$$

164. We can apply these methods to other age groups so long as there are sufficient numbers of points to carry through the calculations. For instance, the UN five point and mild formulae cannot give smoothed values for the first two age groups as also for at least 3 of the last age groups including the open age group where as the three point formula misses only the top one and bottom two groups. The methods also can be applied only so long as all the relevant populations are in five year age groups. On the other hand, the Newton and Carrier-Farrag methods are specifically meant for splitting ten years into five year group.

165. A combination of appropriate grouping, preliminary smoothing and osculatory interpolation might work in some cases. It is advisable to carry out most of these methods before deciding on the most appropriate smoothing. On the whole these methods are risky and will have to be used with care and monitered thoroughly.

Demographic Methods

166. Under this category, there are a whole host of methods. The best known is the one based on models or analogies. In the former, the age-sex distribution of the country is assumed to fit in with an appropriate model population based on demographic parameters. In some cases, the age-sex distribution is borrowed from the experience of another country or area which is anticipated to have similar demographic evolution but has better quality data.

167. A combination of mathematical and demographic methods is sometimes preferable. For example, for the younger and older ages the adjustments may be based on models but for the other ages, moving average or other mathematical methods may be used. The Brass logit method is another technique utilising models, mathematical tools and at the same time trying to utilise available information to the maximum. The model chosen should, as far as feasible, be near the demographic situation in the area being studied.

168. Where more than one enumeration is available, the survival ratios and other details available should be used. One approach to obtain an average age distribution tried in Tanzania, 1978 is to take the averages of persons reported as x and x+11 at the 1967 and 1978 censuses respectively and attribute an age midway between the two. For example, those aged 6 in 1967 and 17 in 1978 are related through mortality and an average of the two will depict population at exact age 12 and so on. The survival ratios also can be used first by smoothing these ratios and applying them to the enumerated age distributions.

169. UN Manual $\underline{\bar{X}}$ gives yet another method based on a quadratic regression equation of transformed cumulated proportional age dustribution utilising standard population values.

170. The method consists of calculating

$$Y(x) = \ln \left[\frac{1 + C(x)}{1 - C(x)}\right], \text{ where}$$

C(x) is the cumulated proportion upto age x in the observed population.

171. The regression equation suggested is:

$$Y(x) = A\, Y_s^2 (x) + B\, Y_s (x), \text{ where}$$

$Y_s(x)$ is calculated from a chosen standard population. The standard chosen should be near to the expected situation for the observed population.

172. The parameters A and B are estimated by using the method of selected points. Usually the numbers of available age values are divided into 2 equal groups and the means of the two groups are utilised If $\bar{Y}_1$, $\bar{Y}_2$, $\bar{Y}_{S1}$, and $\bar{Y}_{S2}$ are respectively the mean of the first and second groups for the observed and standard populations, then

$$A = \left[\frac{1}{\bar{Y}_{S1} - \bar{Y}_{S2}}\right] \left[\frac{\bar{Y}_1}{\bar{Y}_{S1}} - \frac{\bar{Y}_2}{\bar{Y}_{S2}}\right]$$

and

$$B = \frac{\bar{Y}_2}{\bar{Y}_{S2}} - A\, \bar{Y}_{S2}$$

173 Using these values of A and B, the expected values of Y(x) can be obtained from the quadratic regression equation. The values of C(x) are then obtained as

$$C(x) = \frac{e^{Y(x)} - 1}{e^{Y(x)} + 1}$$

Table 2.17

Smoothing of age data by Manual X method: Ghana 1984 -female

Age	C(x)	Y(x)	$C_S(x)$	$Y_S(x)$	Y*(x)	C*(x)	c*(x)
5	.163	.329	.181	.366	.363	.179	.179
10	.322	.668	.327	.679	.669	.323	.144
15	.438	.939	.450	.969	.951	.443	.120
20	.536	1.197	.553	1.245	1.217	.543	.100
25	.628	1.476	.640	1.516	1.477	.628	.085
30	.710	1.774	.713	1.786	1.733	.700	.072
35	.773	2.055	.774	2.060	1.991	.760	.060
40	.821	2.320	.825	2.345	2.257	.811	.051
45	.861	2.594	.867	2.642	2.533	.853	.042
50	.895	2.893	.902	2.966	2.830	.889	.036
55	.924	3.231	.930	3.317	3.149	.918	.029
60	.941	3.493	.952	3.705	3.498	.941	.023
65	.960	3.892	.970	4.185	3.923	.961	.020
70	.972	4.255	.983	4.759	4.424	.976	.015

Note: The $C_S(x)$ values are from Coale-Demeny North stable populations with level 14 and growth rate 3%

C*(x) and c*(x) are the estimated cumulated and actual proportions corresponding to age x.

$$Y(x) = A\, Y_S^2(x) + B\, Y_S(x) \qquad \bar{Y}_1 = 1.205 \quad \bar{Y}_2 = 3.240$$

$$\bar{Y}_{S1} = 1.232 \quad \bar{Y}_{S2} = 3.417$$

$$A = \left[\frac{1}{[\bar{Y}_{S1} - \bar{Y}_{S2}]}\right]\left[\frac{\bar{Y}_1}{\bar{Y}_{S1}} - \frac{\bar{Y}_2}{\bar{Y}_{S2}}\right] = \left[\frac{1}{-2.185}\right][.978 - .948]$$

$$= -.0137$$

$$B = \frac{\bar{Y}_2}{\bar{Y}_{S2}} - A\,\bar{Y}_{S2} = .948 + .0137\,(3.417) = .9948$$

174. The adjusted populations at each age group can then be obtained by successive differencing of the estimated cumulated proportions. Table 2.17 illustrates the use of the method to the female age distribution of Ghana, 1984.

175. Admitting that some of these techniques are time consuming and that it is many times necessary to try all methods before accepting any particular method, it is gratifying that soft wares are now available to carry out these procedures quite expeditiously. The UN five point and Carrier Farrag, Newton quadratic difference and Sprague osculatory formulae are available in the US Census Bureau programmes UNSMH, and SMOTH programmes.

176. The US Bureau has also another method FWDRV which utilises 2 enumerations to adjust population distribution. In addition, the Demeny - Shorter method and the Preston (PRESTO) method produce adjusted age distributions when one has data from two enumerations. PRESTO is available in MORTPAK. Another programme in MORTPAK is CENCT which adjusts one census against another enumeration.

<u>Experiences and observations on basic age-sex data adjustments during 1970 and 1980 round</u>

177. Age sex data from selected African countries (Anglophone) were smoothed by the various mathematical and demographic techniques and results for some of these are presented in table 2.18. An index of dissimilarity was calculated comparing the reported against the smoothed. The largest index was taken as indicative of better smoothing.

178. Some of the methods do not deal with the young and old ages where the problem of enumeration and reporting are quite serious. It is only the osculatory interpolation formula which technically can smooth young ages and old ages other than the open ended ages.

179. A study by ECA concluded that Brass method of smoothing is recommended for African data but cautioned about its use because the intercept and slope derived on the basis of the reported and reference standard age distributions are used to adjust the age distribution. Thus like the quasi or stable population models which are used to smooth data, the Brass method also depends on the adequacy of the selected reference standard. To the extent that there is some subjectivity in this selection, it is suggested that care should be exercised in its use.

<u>Adjustment of Fertility and Mortality Data</u>

180. Since the biggest problem in fertility and mortality data is the large omission of events, any, information on coverage would be useful to adjust the data. The Brass P/F ratio method to a large extent takes care of under reporting of births during the past year as compared with retrospective

information. The Chandrasekar-Deming method of adjustment should be used only when all the conditions needed by the method are fulfilled. The method oi adjusting for recall lapse by using a memory decay curve might also produce reasonable

Table 2.18 Reported and Adjusted female age distributions for Tanzania, 1978

Age Groups	Repor-ted	UN Method	Oscula-tory Inter-polation	Brass logit difference method	Quasi Stable Model	Carrier-Farraq
0-4	18.1	18.9	19.2	18.4	18.3	18.2
5-9	15.8	15.3	14.7	14.3	14.3	15.8
10-14	11.6	11.9	11.6	11.9	11.9	11.8
15-19	9.8	9.6	9.8	10.2	10.0	9.7
20-24	8.3	8.5	8.7	8.7	8.4	8.8
25-29	7.9	7.4	7.5	7.1	7.1	7.4
30-34	5.7	5.9	5.9	6.0	6.1	5.9
35-39	5.0	4.8	4.8	5.1	5.2	4.8
40-44	3.9	4.0	4.1	4.2	4.4	4.1
45-49	3.5	3.4	3.3	3.5	3.7	3.3
50-54	2.7	2.6	2.6	3.0	3.1	2.5
55-59	2.0	2.1	2.1	2.4	2.5	2.1
60-64	2.0	1.8	1.8	1.9	2.0	1.8
65-69	1.3	1.4	1.5	1.4	1.4	1.4
70-74	1.0	1.3	1.1	1.0	0.9	1.0
75-79	0.8	0.7	0.5	0.6	0.5	0.5
80+	0.6	0.4	0.8	0.3	0.2	0.8
Index of diss-imilarity		2.2	2.3	2.9	3.1	1.4

results. A method proposed by Ajit Das Gupta for adjusting parity at older ages might work in certain situations. To adjust for the so called 'zero error' the method proposed by El Badry may be useful but one should ascertain that the conditions for using the technique are fulfilled. Where the proportion of zero parity reported is extremely small and does not vary over age, this method might not work. Fitting of curves to parity or cumulated fertility data might also be tried. the Gompertz curve or logistic curve have been noted to be appropriate. The method proposed by Zaba may also be useful in certain instances. The Coale-Demeny formula and the Brass modification based on parity at young ages to estimate TFR seem to work in certain situations but large variations may be shown by these and in such cases an average may depict the situation.

181. For adjusting mortality data, the Brass, Sullivan, Trussel and Preston-Palloni methods might indicate directions of errors and deficiencies. Adult mortality methods, in view of age misstatement, wrong reporting of status etc., might not be near the actual situation and should be used with caution. The Brass death distribution method and the Courbage method seem worth trying when data on death are deficient.

182. Mortpak has a programme called BENHR which estimates the completeness of adult death registration based on population age distributions from two censuses and registered deaths by age for the intercensal period.

2.6 Effects of errors and biases on vital parameters, Projections etc.

183. In most of the developing countries because direct information is lacking or at best incomplete and defective, indirect estimation procedures are utilised to estimate vital rates. The most important element in the estimation procedures is the age-sex distribution. hence any evaluation of the derived parameters should keep in mind the potential sources of errors coming through age misreporting, under enumeration etc.

184. If the variance of age responses is very high, the conclusions that can be drawn from the data are very restricted. Most studies of age misreporting show a rather large random component in age reporting, suggesting that caution should be exercised in analysis at the individual level.

185. Some uses of data eg stable population analysis are insensitive to gross errors in age reporting, whereas the Brass techniques for estimating fertility and mortality might be sensitive to gross errors in age reporting if they are correlated with parity or marital status and hence stable population analyses might produce biased estimates. However, finding from stable population analysis merely suggests that the effect of age misreporting on age group 15-24 is larger in magnitude and opposite in direction to the effect of under reporting. In addition, since the individuals who are most apt to be missed are the single and childless, the underenumeration in these age groups might bias estimates of fertility rates and of ages at first union and first birth and hence it is still necessary to be cautious of results based on stable population analysis. Brass P/F and child mortality method depend on experience of young women. It is very vulnerable to gross errors in age reporting among women aged 15-29. If age reporting errors are random, it is unlikely that they would cause biases in the estimates of fertility and child mortality; however, they are clearly linked to parity and union status. Parity is most closely related to age reporting and recent fertility is least affected. Whether biases are introduced by age misreporting related to child survival rates is not known.

186. For instance, if a woman's first and only child is now deceased and she is not in a union, there is a good chance that this child might not be recorded. Hence women aged 20-24 whose only child is dead, might be classified as 15-19 and hence bias upward the proportion surviving among children of women aged 20-24.

187. Again, between ages 15-30 most individuals reach puberty, form their first conjugal union and have their first children. Along with these changes, individuals often migrate or set up new households. Each of these events marks a change in a person's place in the community, each plays an important role in population dynamics and each is of primary interest to demographers. The significance of these changes is made even more crucial by the fact that some of the techniques for estimating overall fertility and child mortality rates (especially those developed by Brass) rely heavily on the reported experiences of women aged 15-29.

188. Evidence of biases due to misreporting of age on estimates of fertility is shown by data in following table. Four age groupings: 0-4, 1-5, 2-6 and 3-7 were used to arrive at estimates of fertility from current and retrospective reports and using correction factors P_2/F_2 and P_3/F_3. The estimates even though variable, are not drastically different from each other. Also the estimates based on the different correction factors are within a narrow range, it looks as if the estimates based on 0-4 and 3-7 groupings are similar. Thus it is concluded that digit preference, age shifting etc., may not effect estimates very much.

Table 2.19 Estimates of fertility from different age groupings - Ghana, 1971

Parameter	Basis of estimate	Age range			
		15-19	16-20	17-21	18-22
TFR	P_2/F_2	7.78	7.91	7.70	7.61
	P_3/F_3	7.62	8.06	7.86	7.80

Source: Ewbank D.C., Age misreporting and age selective underenumeration: Sources, patterns and consequences for demographic analysis, National Academy Press, Washington, D.C., 1981.

189. On the other hand, Van de Walle tested the effects of various simple patterns of age misreporting on the Brass method. He noted that age misreporting affected the shape of the fertility curve and that the cumulative fertilities, computed from recorded age-specific fertility rates in the presence of age misreporting were substantially different from the original figures. This was true of parity also. Under the hypothesis of higher fertility and lower parity, he noted that the P/F ratios declined from a rather high value much above 1 at age, 15-19 to values around 1 by age 30-34. Table below presents the P/F ratios by age for the data from the 1978 census of Tanzania, which shows the declining patterns of the ratio by age, perhaps pointing out to the type of error studied by Van de Walle.

Table 2.20 P/F ratios by age - Tanzania 1978

Age group	15-19	20-24	25-29	30-34	35-39	40-44	45-49
P/F ratio	1.40	1.30	1.17	1.12	1.04	0.96	0.90

190. Table 2.21 presents life expectation based on Brass estimates of q_1, q_2, q_3 and q_5, utilising four different age groupings to study the effect of digit preference. It can be noted that the differences are smaller than those among the four q_x values. All of them show high life expectation through q_1 and a declining pattern. But in many other situations, what we see is higher values of life expectation based on q_3 and q_5 perhaps due to omission of some dead children. The single year data giving proportion of dead children for Ghana (Table 2.22) indicates very little of the problem of omission of dead children. It would be interesting to have similar tabulations for other countries where the levels show a contrary pattern.

Table 2.21 Estimates of child mortality derived from different age groupings - Ghana, 1971

Parameter	Basis of estimate	15-19	16-20	17-21	18-22
Life expectation at birth	q_1	57.10	55.86	55.07	55.26
	q_2	53.49	53.49	53.49	52.91
	q_3	51.58	50.47	50.59	50.47
	q_5	49.63	49.21	49.21	48.98

Source: Ewbank, D.C., Age misreporting and age selective underenumeration: Sources, patterns and consequences for demographic analysis, National Academy Press, Washington, D.C., 1981.

Table 2.22 Percentage of children dead by single years of age of mother, Ghana, 1971

Age	16	17	18	19	20	21	22	23	24	25	26	27	28	29
% dead	10.0	6.8	8.1	8.6	10.7	12.7	11.1	12.7	12.7	13.3	13.8	15.9	15.6	15.7
Age	30	31	32	33	34	35	36	37	38	39	40	41	42	43
% dead	17.3	15.7	17.8	18.6	19.3	19.7	18.6	21.0	20.3	19.0	22.2	22.4	24.0	25.4

Source: Ewbank D.C., Age misreporting and age selective underenumeration: Sources, patterns and consequences for demographic analysis, National Academy Press, Washington, D.C., 1981.

191. The effect of age misstatement on Brass estimates of child mortality was studied by Santow through a micro simulation model and following Table 2.23 presents the results. In one simulation, the age of half of the women with parity above average was shifted upward and the other simulation shifted all women with above average parity. For moderate level of mortality, the difference between no shift and half shift was on the average one year in life expectation, and was same for high mortality as well. For the full shift, the differences were much larger-about 3 years on the average. Interestingly, the life expectation shown by the successive q_x values were declining in all the cases - no, half and full shift and high and moderate mortality.

192. Regarding data on parent survival, exaggeration of age is a serious one. This would result in estimates of life expectation much higher than the true value. For example, an exaggeration of 2.5 years would result in adding 2.5 years to estimated life expectation. The bias may be more, if reporting of age is different among those with parents alive and those who have lost their parents. An important variable used in adult mortality estimation is the mean age at child birth. Age exaggeration certainly would bloat this up. Again, in societies where male remarriage is not infrequent, there could be large difference between male and female age at marriage and hence in age at child birth. For example, one study with data from the Sudan 1973 census indicated that an increase of 3 years in the age difference between husbands and wives would result in an increase in male life expectation by 1.5 years. A similar attempt using WFS data from Lesotho showed that an assumption of an exaggeration of 2 years in age reporting but with the mean age at birth not changed, reduced life expectation of males only by half a year. But a similar exaggeration both in age reporting and in mean age at birth for females showed a decrease of about 2.4 years.

Table 2.23 Simulations of the effect of age misreporting on Brass child mortality estimates

Parameter	Female life expectancy at birth with no age misreporting	half shift	full shift	mixed
Moderate mortality				
q_1	61.4	63.5	66.2	65.9
q_2	60.6	61.4	64.2	60.3
q_3	60.3	60.3	60.7	59.4
High mortality				
q_1	45.3	44.1	50.4	
q_2	38.2	40.9	46.6	
q_3	37.3	38.8	39.5	

Source: Santow G., Microsimulation test of the effect of age misstatement on estimates of childhood mortality, Working paper No. 8, Voorberg, Netherlands, 1977.

193. Brass parental survival method may be potentially biased not only by age misreporting but also by the effects of age misreporting on the standard life table used to smooth the survivorship estimates obtained.

194. The effect of age misstatement on widowhood/widowerhood also is important in estimation of adult mortality. The proportions with spouses still alive as obtained from the 1977 fertility survey of Lesotho is presented below. (Table 2.24) Even though the proportion is declining smoothly, still the values look large especially at older ages. Using these values, life tables were constructed by splicing child and adult mortality through the method suggested by Brass and Hill and the life expectations were too high.

195. Rashed tested the sensitivity of the Brass death distribution technique to age misreporting under 2 models - one with the same pattern of age misreporting for the population and the deaths and the other with two different patterns for population and deaths. The result showed that death rate could be off by 1 to 2 points per 1000. Smoothing the age distribution of total population and of deaths separately tends to increase the error in estimated death rate.

Table 2.24 Proportion of ever married persons with first spouse still alive, Lesotho, 1977

Age group	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59
M	.9942	.9904	.9765	.9704	.9399	.9143	.8891	.8632
Prop. not widowed								
F	.9793	.9076	.9311	.8981	.8154	.7298	.6440	.5208

Source: Lesotho Fertility Survey, 1977.

2.7 Conclusion

196. It has to be stressed that data correction will not substitute for care and caution in data collection and in fact it would be detrimental to give the impression that improvement of data collection is not necessary because correction can always be done. Methods and techniques would be harmful if they were used as a pretext to ignore the importance of a deep quantitative knowledge of a population's characteristics, geography, economic life, sociology, customs and history.

197. This does not imply that data need not be adjusted, but what is needed is that in every case, adjustment should be specific to the situation and care should be taken not to over correct data, so that the so called 'corrected data' does not lead away from reality.

198. A battery of tests and methods are needed before an assessment of data can be made. A cafeteria approach subjecting the data to several types of scrutiny might reveal

the patterns and magnitudes of errors, biases and other deficiencies and one may be not only able to arrive at plausible explanations for these so that future data collection efforts could be aware of these and avoid them, they could also be guidelights for adjustments.

199. The types and varieties of errors vary from one country to another and perhaps over time. Generalisation from one experience is risky. Evaluation and adjustment is still an art and needs a lot of experience in handling data. In developing countries with many unknown facts about population, evaluation is like completing a jig-saw puzzle with some pieces missing and a few distorted. What one tries is to disentangle the pieces and see what emerges.

200. It is useful if evaluation studies are published in detail so that lessons could be learnt from the experiences. There is clear need for not only detailed studies on data quality but also their publication and wider dissemination.

201. Evaluation and analysis of data should be planned in advance and continuous quality control must be ensured. A post mortem examination might not reveal significant factors and facts and much of the vital information collected might become either less valuable or sometimes useless.

REFERENCES

Banister. J.,
Use and abuse of census editing and imputation, Asian and Pacific census forum, Vol. 6. No. 1, East West Population Institute, Honolulu, 1980.

Botswana
Census administrative/technical report, 1981 Population and housing census, Gaborone, 1983.

Brass. W., et al
The demography of tropical Africa, Princeton University Press, New Jersey, 1968.

Caldwell. J.C., and Igun. A.A.,
An experiment with census type age enumeration in Nigeria, Population Studies Vol. 25(2), 1971.

Carrier. N.H., and Hobcraft J.
Demographic estimation for developing societies, Population investigation committee, London, 1971.

Coale. A.J., and Demeny, P.,
Regional model life tables and stable populations, Princeton University Press, New Jersey, 1966.

Ewbank. D.C.,
Age misreporting and age selective underenumeration: Sources, patterns and consequences for demographic analysis, Committee on population and demography, Report No. 4, National Academy Press, Washington, D.C., 1981.

Ghana
1984 population census of Ghana, Preliminary report, Accra, 1984.

Mauritius
1983 housing and population census of Mauritius, Analysis report, Vol. 1. Rose Hill, 1985.

Ramachandran. K.V., and Nair., P.S.G.,
The overall survival ratio method of evaluating defective and incomplete data and estimating mortality, Studies in demography, edited by A. Bose et al, George Allen and Unwin, London, 1970.

United Nations (ECA)
Study on methods and problems of the 1970 round of African population and housing censuses, Tenth Session of the Conference of African Statisticians, E/CN. 14/CAS. 10/15, Addis Ababa, 1977.

United Nations (ECA)
Report of the working group on recommendations for the 1980 population and housing census in Africa, E/CN. 14/CPH/47, Addis Ababa, 1978.

United Nations (ECA)
Evaluation of the 1980 round of population and housing census in Africa, E/ECA/PSD/4/40, Addis Ababa, 1986.

United Nations (ECA)
Working group on recommendations for the 1990 round of population and housing censuses in Africa, E/ECA/CPH/4, Addis Ababa, 1987.

United Nations
Manual X, Indirect techniques for demographic esimaion, ST/SOA/SER.A/81, New York, 1983.

United States (Bureau of the Census)
Popstan- A case study for the 1980 censuses of population and housing, US Dept. of Commerce, Washington D.C, 1979.

Zambia
1980 population and housing census of Zambia, Analytical report, Vol. II, CSO, Lusaka, 1987.

CHAPTER 3

ESTIMATION OF MORTALITY

3.1 Introduction

1. The importance of estimating current levels and future trend of mortality in Africa for planning and policy formulation can not be over-emphasized. Such an exercise will also help in evaluating the quality of data collection which will consequently throw light on the types of data required. However, the statistical systems in Africa have not always produced reliable data through the traditional means of civil registeration. As a result, demographers have increasingly depended on "indirect" techniques of estimating vital parameters.

2. Over the years, the amount of data collected has increased in volume. Although the improvement in the quality of data might not be as much as the increase in quantity, it is safe to say that there has also been an improvement in the quality of data since the early 1960s. On the other hand, there have been a proliferation of techniques of data evaluation and analysis aspecially mortality and fertility in the last two decades. This section will review some of the techniques of mortality analysis and illustrate methods for estimating levels and some trends of mortality.

3.2 Data Sources

3. Except for some off-shore Islands and some of the countries bordering the Mediterranean Sea, most Afican countries, éither do not have Vital Registeration Systems or where they exist they are neither representative nor complete. The result is that mortality levels and trends can not be estimated using conventional methods. Indices such as crude death rates, standardized crude death rates and expectation of life at birth can not be computed directly in the absence of data on deaths in a year by age and sex. Consequently, estimation of mortality has concentrated on data collected in censuses and sample surveys.

3.2.1 Censuses

4. "Census" taking in Africa started during the early days of colonialism when the then admistrators saw the need to take population counts to assess the magnitude of labour reserves or to levy hut tax. Most of these "censuses" were head counts which did not collect information useful for vital rate estimation. After Independence, most of the English speaking African countries undertook censuses which provided age sex distributions that were used to estimate the level of fertility and mortality under certain assumptions. The development of the Brass technique of estimating infant and child mortatility in the early 1960s from data on children ever born and children surviving to all women in the reproductive age group led to the inclusion of questions on survival of children ever-born to women 15 years of age and over. The questions were to be

answered by the head of the household on behalf of each woman in the Household who is 15 years and over. The questions were: How many children has she ever born? and How many of these are now still living? With the inclusion of these questions in census schedules and their increasing promise in providing somewhat robust estimates, the questions were modified to increase their accuracy. They were formulated to read: Of the children ever born to the woman, how many are still living with her? How many are still living elsewhere? and how many are dead? These information were preferably to be collected by sex of children.

5. The apparent success of the children ever born-children surviving (or children dead) data to estimate Infant and Child Mortality (ICM) statistics led to the investigation of the relevance of data on the survival of parents and the further recommendation by the U.N. to the inclusion of questions on survival of parents in Census Schedules. Although, the results of using data on survival of parents has not been as successful as the survival of children, it still forms an independent source of estimating adult mortality especially for females. Data on whether the first spouse of an ever married member of a household is still alive or not as well as data on survival of siblings are also suggested as potential sources of female adult and maternal mortality respectively. We do not envisage to expatiate on the last two ie. the survival of spouse as well as siblings, mainly because the data required for the application of these techniques are not collected by most of the African countries.

3.2.2 Sample Surveys

6. Sample surveys are second to censuses in providing information on mortality in Africa. Apart from the post-Enumeration surveys that accompanied censuses, ad-hoc surveys have been going on in the continent which have been collecting data on children ever-born and children surviving. In fact, prior to the 1980 round of censuses, most of the French-speaking African countries collected their demographic data from Sample Surveys. The data collected from such surveys have been: children ever-born and children surviving and the survival status of the most recent birth or status of the birth in the last 12 months. In terms of data on mortality from Sample Surveys, the World Fertility survey deserves a special mention. It is a survey which collected a very high quality data on fertility which has also been used to derive mortality indices from a small sample with intensive supervision though it was not meant for mortality studies. The Maternity history data which records the date of birth, and the date of death of each child born to a woman in a reproductive age group provides the source for estimating the level of infant and early childhood mortality both directly and indirectly, thereby allowing one to assess the reliability of the indirect methods of estimating infant and child mortality.

3.3 Methods of Estimating Mortality

The methods available to estimate mortality indices vary with the types of data on hand. They can broadly be classified into two parts. These are those that depend on conventional data and those that are based on unconventional data. The methods can also be classified by whether they estimate adult or child mortality. Except for a brief mention of the available methods, the paper will concentrate more on those methods that are based on unconventional data. Classification will also be made on whether they estimate childhood mortality or adult mortality.

3.3.1 Child Mortality Estimation

The conventional data on child mortality are: births, deaths in infancy and childhood, and child population by age. Each of these data are potentially subject to large errors. Evaluation of the data before computing infant mortality rate can be made using the balancing equation for consistency of the two sources of data; ie child population from census and births and deaths from vital registeration. Such data as stated above are not available in most sub-Saharan countries.

Child Mortality Estimation from Unconventional Data

9. Data on reports of women concerning their live births and the subsequent survival of those children are utilised to estimate mortality rates. In its simplest form, the technique requires the total number of children ever born and the total number surviving among the ever born classified by age of mother at the time of census or survey. Methods have been developed, first by Brass (1964) and subsequently by Sullivan (1972) and Trussell (1975) to convert the proportion of children dead reported by women of each age group into probabilities of dying, q(i) before childhood age (i). Ideally, to estimate q(i), one would like to identify a group of children at birth, follow them for i years, and see how many fail to survive. But, this direct method is impossible to achieve in Africa south of the Sahara because of the reasons given earlier. Therefore, the Brass technique circumvents the problem of following a cohort of births by developing an ingenious method of converting a mortality statistics D(x), the proportion of children dead among ever born to women of age x, into a probability of dying before age i, q(i). Since the children of mothers who are aged x are not of the same age i, the proportion dead, D(x), is a composite of child-mortality levels. Brass's technique is therefore essentially the development of a multiplier that will convert D(x) to q(i).

10. For a clear understanding of the underlying principles of the method, a brief expositions will be made below.

Let $D(x)$ = the ratio of children dead to all children ever born to women of age x in completed years.

$C_x(t)$ = the proportion of children born t years before the census or survey (whether alive or not) to mothers of age x

$q(t)$ = the proportion of children who died among those born t-years before the census.

If s represents the earliest age of child-bearing,

$$D(x) = \sum_{t=1}^{x-s} C_x(t)q(t) \qquad (1)$$

Under assumption of constant fertility and mortality and given that the information on number of children ever-born and children surviving were accurately reported, a life table satisfying equation (1) can be determined for a given $C_x(t)$ distribution. Let $q'(t)$ be such a life table.

Equation (1), then becomes

$$D(x) = \sum_{t=1}^{x-s.} C_x(t)q'(t) \qquad (2)$$

From equation (2), it follows that the proportion of children dead is a weighted average of the probability of dying with $C_x(t)$ as weights. By mean value theorem, there exists a number u where $1 < u < x-s$ such that

$$D(x) = q'(u) \qquad (3)$$

Further, let us assume that the pattern of mortality is known. In such a life table system $q'(t) = K.\ q'_s(t)$ where $q'_s(t)$ is the survival probability which can be picked from among the one parameter life table systems which is considered here as a standard. K is independent of age and measures the level of mortality.

12. In such a situation, equation (3) becomes

$$D(x)=K\ .\ q'_s(u) \qquad (4)$$

Since u can assume integal as well as non-integral values depending on the distribution of children by age, Brass developed a set of conversion factors to transform the proportions dead, $D(x)$, to probabilities of dying to specific ages, $q'_s(i)$.

The formula used to develop the conversion factors for age group u is:

$$K(i) = \frac{q'(i)}{D(x)} = \frac{k.q'_s(i)}{k.q'_s(u)} = \frac{q'_s(i)}{q'_s(u)} \qquad (5)$$

Similar argument can also be used to determine the conversion factors for data tabulated by duration since first marriage.

The correspondence between u and i in that case will remain the same, the only change will be the age group of mothers being replaced by duration since marriage. The correspondence is that duration 0-4, 5-9, 10-14, 15-19, 20-24,25-29 and 30-34 replace the five year age groups 20-24, 25-29, 30-34 upto 50-54 respectively.

13. The development of equation (5) assumes that the country under study has the relevant data to determine the $C_x(t)$ correctly. But, since the state of demographic data collection in Africa has not reached such a level, the need for indirect methods of estimating mortality would be in order. Lack of reliable data to estimate $C_x(t)$ led Brass to use a model fertility function or artificially constructed population which was assumed to broadly define the fertility pattern in Africa. The fertility model was a third degree polynomial in age expressed as $f(x) = C(X-S)(S+33-X)^2$ where x is age of woman, S is the earliest age of child bearing and C is a constant dependent on the level of fertility. Note that the fertility function is a rigid curve that changes the level but not the structure. The curve slides forward and backwards along the age axis with changes in the value of S, the earliest age of childbearing.

14. A panel of multipliers is therefore developed for various values of S, but since measuring S in real situations will be difficult, an entry parameter which is related to S was used instead. Such a parameter was a ratio of the mean number of children ever born to women 15-19 to the mean number of children ever born to women 20-24 (P1/P2) hereinafter called PAR1. Other entry parameters such as the mean and the median of the age specific fertility rates are also used to select multipliers for the older age cohorts (women 30 years and over). Later, P2/P3, (PAR2), was also included as another parameter for selecting the multipliers for age groups under 30 years. The logit system generated by the general standard was the mortality pattern used in the development of the multipliers (UN. 1983).

15. Later studies by Sullivan (1972) and Trussell (1975) examined the estimation problem in a different way. They studied the relations between D(i) and q(j) using regression analysis. For these analysis, observed fertility distributions of European experience were used by the former while the latter used data generated from model fertility schedules developed by Coale and Trussell (1974). The entry parameters were PAR1 for the former and PAR1 and PAR2 for the latter. Sullivan developed multipliers for three age groups ie 20-24, 25-29 and 30-34, while Trussell developed multipliers for all seven (15-19, 20-24, ..., 45-49) age groups. Unlike Brass, both Sullivan and Trussell developed their coefficients for each of the Four Coale and Demeny (1966) Regional Model life tables. Based on the New United Nations Model life tables and fertility schedules generated by Coale-Trussel, Palloni and Heligman (1986) have also developed coefficients that help investigators

to compute mutipliers that help to convert proportion dead among children ever born by age of mother to probability of dying before exact ages 1, 2, 3, 5, 10, etc. For a detailed exposition of the various variants, the origional papers should be consulted; though in the case of Trussel's variant a modified and simpler version is given in a recent publication of the United Nations(1983) on indirect techniques for demographic estimation.

16. Here, the Trussell variants will be applied. Tables 3.1 and 3.2 give the multipliers. Experience in RIPS and elsewhere indicate that in practice, the various procedures for analysing data on proportions dead among children ever born give similar results. For this reason, the Trussell variant will be demonstrated.

<u>Application of the Children Ever Born/Children Surviving to estimate Child Mortality</u>

17. The data required for the application of the Trussell variant of estimating child mortality are:

(A) Women in the reproductive age group by age or married women by duration since first marriage.

(B) Number of children ever born to women in each age group or duration group.

(C) Number of children dead or surviving among those ever born in each age or duration group.

The mechanics of the technique:

(1) Let CEB(i), CD(i) and W(i) denote respectively, the number of children ever born, the number dead among the ever born and the number of women in age group i where i=1,2,3,4, ..., 7, corresponding respectively to age groups 15-19, 20-24, ... and 45-49.

<u>Step 1</u>: Calculate the mean Parity, denoted as P(i), using

$$P(i) = \frac{CEB(i)}{W(i)} \quad \text{for each of the i's}$$

For instance, in the case of the Kenyan example, Table 3.3; for i=4, CEB(4) = 2,223,620; CD(4) = 368,676 and W(4) = 412,691

$$P(4) = \frac{2223620}{412691} = 5.3881$$

Table 3.1 Coefficients for Estimation of Child Mortality Multipliers, Trussell Variant, When Data Are Classified by Age of Mother

Mortality model (1)	Age group (2)	Index i (3)	Mortality ratio[a] q(x)/D(i) (4)	Coefficients a(i) (5)	b(i) (6)	c(i) (7)
North........	15-19	1	q(1)/D(1)	1.119	-2.9287	0.8507
	20-24	2	q(2)/D(2)	1.2390	-0.6865	-0.2745
	25-29	3	q(3)/D(3)	1.1884	0.0421	-0.5156
	30-34	4	q(5)/D(4)	1.2046	0.3037	-0.5656
	35-39	5	q(10)/D(5)	1.2586	0.4236	-0.5898
	40-44	6	q(15)/D(6)	1.2240	0.4222	-0.5456
	45-49	7	q(20)/D(7)	1.1772	0.3486	-0.4624
South........	15-19	1	q(1)/D(1)	1.0819	-3.0005	-0.8689
	20-24	2	q(2)/D(2)	1.2846	-0.6181	-0.3024
	25-29	3	q(3)/D(3)	1.1223	0.0851	-0.4704
	30-34	4	q(5)/D(4)	1.1905	0.2631	-0.4487
	35-39	5	q(10)/D(5)	1.1911	0.3152	-0.4291
	40-44	6	q(15)/D(6)	1.1564	0.3017	-0.3958
	45-49	7	q(20)/D(7)	1.1307	0.2596	-0.3538
East........	15-19	1	q(1)/D(1)	1.1461	-2.2536	0.6259
	20-24	2	q(2)/D(2)	1.2231	-0.4301	-0.2245
	25-29	3	q(3)/D(3)	1.1593	0.0581	-0.3479
	30-34	4	q(5)/D(4)	1.1404	0.1991	-0.3479
	35-39	5	q(10)/D(5)	1.1540	0.2511	-0.3506
	40-44	6	q(15)/D(6)	1.1336	0.2556	-0.3428
	45-49	7	q(20)/D(7)	1.1201	0.2362	-0.3268
West	15-19	1	q(1)/D(1)	1.1415	-2.7070	0.7663
	20-24	2	q(2)/D(2)	1.2563	-0.5381	-0.2637
	25-29	3	q(3)/D(3)	1.1851	0.0633	-0.4177
	30-34	4	q(5)/D(4)	1.1720	0.2341	-0.4272
	35-39	5	q(10)/D(5)	1.1865	0.3080	-0.4452
	40-44	6	q(15)/D(6)	1.1746	0.3314	-0.4537
	45-49	7	q(20)/D(7)	1.1639	0.3190	-0.4435

Estimation equations:

k(i)=a(i)+b(i) PAR1+c(i)PAR2

q(x)=k(i)D(i)

[a] Ratio of probability of dying to proportion of children dead. This ratio is set equal to the multiplier k(i).

Source: United Nations. (Manual X). Indirect Techniques for Demographic Estimation. Population Studies of the United Nations. No.81. ST/ESA/SER.A/81. New York. P.77.

Table 3.2 Coefficients for Estimation of the Reference Period, t(x)[a], to which the Values of q(x) Estimated from Data Classified by Age Refer

Mortality model (1)	Age group (2)	Index i (3)	Age x (4)	Parameter estimate (5)	Coefficients a(i) (6)	b(i) (7)	c(i) (8)
North........	15-19	1	1	q(1)	1.0921	5.4732	-1.9672
	20-24	2	2	q(2)	1.3207	5.3751	0.2133
	25-29	3	3	q(3)	1.5996	2.6268	4.3701
	30-34	4	5	q(5)	2.0779	- 1.7908	9.4126
	35-39	5	10	q(10)	2.7705	- 7.3403	14.9352
	40-44	6	15	q(15)	4.1520	-12.2448	19.2349
	45-49	7	20	q(20)	6.9650	-13.9160	19.9542
South........	15-19	1	1	q(1)	1.0900	5.4443	-1.9721
	20-24	2	2	q(2)	1.3079	5.5568	0.2021
	25-29	3	3	q(3)	1.5173	2.6755	4.7471
	30-34	4	5	q(5)	1.9399	- 2.2739	10.3876
	35-39	5	10	q(10)	2.6157	- 8.4819	16.5153
	40-44	6	15	q(15)	4.0794	-13.8308	21.1866
	45-49	7	20	q(20)	7.1796	-15.3880	21.7892
East	15-19	1	1	q(1)	1.0959	5.5864	-1.9949
	20-24	2	2	q(2)	1.2921	5.5897	0.3631
	25-29	3	3	q(3)	1.5021	2.4692	5.0927
	30-34	4	5	q(5)	1.9347	- 2.6419	10.8533
	35-39	5	10	q(10)	2.6197	- 8.9693	17.0981
	40-44	6	15	q(15)	4.1317	-14.3550	21.8247
	45-49	7	20	q(20)	7.3657	-15.8083	22.3005
West	15-19	1	1	q(1)	1.0970	5.5628	-1.9956
	20-24	2	2	q(2)	1.3062	5.5677	0.2962
	25-29	3	3	q(3)	1.5305	2.5528	4.8962
	30-34	4	5	q(5)	1.9991	- 2.4261	10.4282
	35-39	5	10	q(10)	2.7632	- 8.4065	16.1787
	40-44	6	15	q(15)	4.3468	-13.2436	20.1990
	45-49	7	20	q(20)	7.5242	-14.2013	20.0162

Estimation equations:

t(x)=a(i)+b(i) PAR1+c(i) PAR2

[a] Number of years prior to the survey.

Source: United Nations. (Manual X). Indirect techniques for Demographic Estimation. Population Studies of the United Nations. No.81. ST/ESA/SER.A/81. New York. P.78.

Using the same formula, the mean parities for all the other age groups are calculated and the results entered in column 6.
Column seven contains the proportion dead for each age group. It is calculated by the formula

$$d(i) = \frac{CD(i)}{CEB(i)}$$

For example,

$$d(4) = \frac{368676}{2223620} = 0.1658.$$

The rest of the figures are similarly computed.
To determine the appropriate multipliers from the table of coefficients developed for this purpose by Trussell, PAR1 and PAR2 would have to be computed.

18. For Kenya, PAR1 = 0.3206/1.8529 = 0.1730 and PAR2 = 1.8529/3.6521 = 0.5074. For each age i, the conversion factors are computed from Trussell coefficient given in Table 3.2 using the relationship

$$K(i) = a(i) + b(i)^{*}\ PAR1 + c(i)^{*}\ PAR2.$$

Once again for i=4, and pattern of mortality depicted by North model of Coale and Demeny, K(4) = 1.2046+(0.3037x0.1730)-(0.5656x0.5074) = 0.970.

19. For fixed values of PAR1 and PAR2 and varying a(i), b(i) and c(i), the set of multipliers for all the seven age-groups were computed and entered in column 8.

20. Column 9 contains the exact ages to which the probabilities of dying refer. The proportion dead in column seven are then multiplied by the multipliers in column 8 and the resulting figure entered in column 10.

21. In theory the first figure is a measure of infant mortality and the last a cumulative mortality from infancy to age 20. For evaluative purposes, it will be necessary to convert them to same unit of measure such as implied levels, expectation of life or probability of dying before age five, etc. before one can make a judgement on the estimated infant and child mortality. Here the implied levels were determined and provided in the last but one column.

22. One way of computing the implied levels of mortality will be to determine the proportion of those who survive a particular age from the probability of dying by that age, i.e.

$$l(x) = 1 - {}_{x}q_{o}.$$

If the data were given by sex and the probability of dying estimated for one sex only, the problem could be solved by matching the estimated l(x) against the appropriate model life table l(x) values which are given in the Coale and Demeny Model life tables for each level of mortality. But, since the data on children ever born and children dead are given for both sexes combined, the estimated probabilities of dying or probabilities of surviving from birth are also for both sexes combined. Therefore, the model l(x) values will have to be combined by a sex ratio before the matching process begins. Fortunately the combined probabilities of survival from birth to age x for x = 1,2,3,5,10, and 15 are available in Manual X pages 272-283 (UN,1983) for various sex ratios at birth. Hence they can conveniently be used. In the case of Kenya, the North model levels with sex ratio of 103 males for every 100 females was used.

23. The last column refers to the calendar periods to which the estimated mortality levels apply. They are determined by subtracting the reference period, t(x), from the date of the census which was on 24/08/1979, ie.(1979.73).

24. The t(i) are estimated from Table of Trussell coefficients (Table 3.1) for given PAR1 and PAR2 (See Table 3.3)

For i = 1, $t(1) = 1.0921+(5.4732x0.123)+(1.9672x0.5074)$
$= 1.04$ years before census.

Therefore, the calendar year to which $_1q_0$ refers will be 1979.73-1.04 = 1978.69 (i.e. July-August 1978).
From column 12, it appears that mortality has been declining from the late sixties upto the middle seventies. However, the $_2q_0$ and $_1q_0$ implied levels suggest that the declining trend in mortality were reversed since 1975 and infact if the figures were to be taken at face value, they suggest that the deterioration was on the increase upto the late 1970s. Though, the quality of life in the country in the 1970-80 decade might not have been as good as the decade preceeding it, the implied levels by the first two age groups have usually been depicting unusually high mortality level. It is likely that age errors, selectivity as well as differential mortality by birth order (higher amongst first born) are to explain the apparent inconsistency. It is also possible that small number of events at the early ages especially of births and deaths might also introduce fluctuations.

Table 3.3: Application of Trussell Method of Estimating Infant and Child Mortality From Children Ever Born and Children Dead Tabulated by Age Group of Women Kenya 1979

Age Group x-x+4	Index Age Group i	Number of Women W(i)	Number of Children Ever Born CEB(i)	Number of Children Dead CD(i)	Parity P(i)	Proportion Children Dead d(i)	Trussell Multipliers M(i) M(i)	Exact Age j	Probability of Dying Before Age j q(j)	Probability of Surviving from Birth to Age j, l(j)	Corresponding Level of Mortality	Calendar Year to Which Level Refers
1	2	3	4	5	6	7	8	9	10	11	12	13
15-19	1	887722	284604	33042	0.3206	0.1161	1.0368	1	0.120	0.880	12.6	1978.69
20-24	2	686003	1271095	158633	1.8529	0.1248	0.9810	2	0.123	0.877	14.6	1977.37
25-29	3	541261	1976739	278918	3.6521	0.1411	0.9341	3	0.132	0.868	15.0	1975.46
30-34	4	412691	2223620	368676	5.3881	0.1658	0.9702	5	0.161	0.839	14.5	1973.19
35-39	5	325367	2105222	388413	6.4703	0.1845	1.0326	10	0.191	0.809	14.3	1970.65
40-44	6	273702	1921799	417799	7.0215	0.2174	1.020	15	0.221	0.779	13.7	1967.94
45-49	7	221965	1592266	403162	7.1735	0.2532	1.0029	20	0.254	0.746	13.1	1965.05

Problems in the Implementation of the Methods of Mortality Estimation On Children Ever Born - Children Dead

25. The first problem in implementing the technique arises when the age pattern of mortality is unknown. In the early development of the technique, Brass gave his multipliers for one parameter model. Later, however, multipliers were given for the four families of Coale-Demeny Model Life Tables (Coale and Demeny,1966). Multipliers that can be used to convert proportion of children dead among ever born are also available for the new UN model life tables. Since the assumed age pattern of mortality influences the probability of dying before age x, xqo, it is helpful to have some evidence on the pattern of mortality in order to choose the appropriate multipliers. The effect of pattern is more serious for the youngest age group than for the older ones. The implied infant mortality varies widely with pattern of mortality. It is suggested that $_5q_0$ is less affected by pattern than $_1q_0$, $_2q_0$ and $_3q_0$ (Hill, 1986). In an event of lack of information on the true pattern of mortality in a population under study, it is suggested that the multipliers for the West Model life tables of the Coale-Demeny system be used since this choice minimizes error.

26. In sub-Saharan Africa, the North family has been considered the most appropriate and subsequently extensively used because hospital statistics on Infant and child mortality in this area indicated very high infant and early childhood mortality compared to late childhood mortality. However, a number of studies suggest that the pattern in Africa south of the Sahara is unlikely to be a single pattern (Ekanem and Som, 1984; Blacker and Hill, 1985; Tesfay (Forthcoming); Tesfay and Venkatacharya 1986).

27. A second problem is the assumption that fertility has remained constant for quite a long time at least for the young age group of women. Such an assumption is likely to bias the estimated probability of dying. If we assume mistakenly that the present low fertility rates obtained in the past as well, and calculate C_x, the percentage distribution of children of women aged x, based on the observed PAR1 and/or PAR2 we would under-calculate the frequency of children at higher ages (when fertility was high) and over calculate it at younger ages. Since q(t) increases with age, the model proportion dead $C_x(t).q(t)dt$ will be smaller than it should be and q(i) will over estimate the true q(j). Fortunately, since declines in fertility can be achieved by various methods i.e. postponement of age at first marriage, age at first birth, use of contraception to space births, etc, their effects may sometimes cancel out and lead only to a moderate bias. The error is usually not too serious for younger ages but increases with age. Various trials with some declines of fertility under different levels of mortality did not result in an error in q(i) more than 8% (Venkatacharya, 1979). The effect of changes in fertility can also be reduced by using Preston-Palloni's

technique which attempts to determine the distribution of children ever born from the distribution of surviving children without recourse to fertility models. One problem of the Preston-Palloni method is the need for matching of children surviving by age to their mothers in the household schedule. In most censuses in Africa, matching will be a problem because of polygyny, fosterage, and complex household structures. Additional question on serial number of mother for surviving children will have to be added to the census or survey schedules. But, whether an addition of a question on census schedule is cost-effective in view of the problems of age-reporting and poor completion of enumeration in the young ages which are likely to undermine the advantages of the technique, is questionable. Another solution could also be to list children in a household immediately after the mother.

28. A serious assumption in the original formulation to estimate infant and childhood mortality from information on children ever-born/ children dead is that of constant mortality. The bias introduced by the violation of the assumption i.e. declining death rates over the period before the census or survey has been investigated by Kraly and Norris (1978), Sullivan and Udofia (1979) and Palloni (1979, 1980). Under the assumption that child mortality changes regularly over time, model calculations showed that the derived probabilities of dying from proportion dead among children ever born closely approximate the period life table for a particular point in time prior to the survey or census. This observation was noted regardless of the pace of mortality change. Feeney(1980) was the first to use this relationship to provide a table which will help investigators to estimate infant mortality rates to each proportion dead by age of woman and the time to which these estimated infant mortality rates refer. Since the tables were given for the Brass general standard, it was not flexible enough to accommodate varying patterns of mortality. Coale and Trussell (1978) also developed a procedure for estimating the time location of each of the $_jq_o$ estimated from children ever born and children dead. In this paper the variant given in Manual X, will be used because of its simplicity in its application and its capability to be used for varying patterns of mortality. Its formulation is:

$$t(j) = a(i) + b(i)(PAR1) + c(i)(PAR2)$$

where t(j) is the number of years before the census or survey to which q(j) the period life table with q(j) as probability of dying was operational, a(i), b(i) and c(i) are coefficients provided for each age group indexed by i=1 (15-19), i=2 (20-24), etc, and PAR1 and PAR2 are as defined earlier.

29. The Brass procedure for estimting child mortality is vulnerable to gross errors in age reporting and differential reporting of dead and surviving children. In order to improve the quality of data on children ever born and children dead, the earlier formulation of the census question ie. How many children have you born? How many are now living was changed to: Of the number of children ever born to you, how many are living with you now? How many are living elsewhere? How many are dead? An addition of the three gave the number of children ever born unlike in the previous census/surveys which asked the total childrn ever born.

30. In censuses and surveys, two types of errors are easily noticed. In the first case a number of women of zero parity in the young age groups are misclassified as women whose parity is not stated because enumerators sometimes equate a blank for a zero. The problem is serious for the younger women. Depending on the number involved, the mean parities will either be increased (if excluded) or decreased (if included) thereby biasing the multipliers. In such situations, the El-Badry (1961) correction factor should be applied before computing the multipliers.

31. The other obvious error is the effect of omission of dead children mostly those who died shortly after birth in the past by old women. Observation of mean parities by age of women usually increases by age upto about age 40 but abruptly declines thereafter indicating either recall lapse or selectivity or both.

32. What ever may be the reason, it affects the estimated mortality rates. For this reason, it is often suggested that the experience of young women should only be utilised. Unfortunately, the age group 15-29 in Africa is affected by gross age errors. If age errors were random, it is unlikely they would bias the estimated childmortality rates seriously because multipliers developed for non-conventional age-groups could be used after grouping the data in appropriate ages. otherwise, the direction and magnitude of the error will depend on the direction and magnitude of the pattern of reporting. The effect of age errors will enter the computation through PAR1 and PAR2 ratios as well as through the reported proportions of children dead by age of mother. It is possible that some of the errors cancel out.

33. The development of the Brass-type technique of estimating infant and child mortality assumes that child mortality risks do not vary by age of mother. Lack of data have hampered the assessment of such assumption in Africa; nevertheless practical experience shows that mortality estimates based on reports of women aged 15-19 and sometimes 20-24 are usually out of line with the other estimates (usually on the high side). Among the reasons given are that these group of women are selective (i.e. the women marry young and are of low socio-economic status) or suffer from differential mortality by birth order (higher amongst first born).

34. Mothers in old ages are likely to be non-representative of their birth cohorts if the mortality situation has been high for them. The child mortality experience of mothers who are dead and those alive might not be the same. Infact the children of dead mothers might be subjected to higher mortality risks than those whose mothers were alive at the time of the survey.

<u>Survival of most recent live birth</u>

35. In Censuses and Sample Surveys, data of each woman's most recent live birth or sometimes questions as to whether a birth occured in the 12 months preceding the survey is asked for current fertility estimation. In the same schedule a question on the survival status of the child is also included. From such data, estimation of recent child mortality can be made. In addition, if the data are tabulated by age of women, child mortality differentials by age of mother could be potentially studied which is not possible from data on children ever born and children dead.

36. In this technique, the proportion surviving among births in the 12 months preceeding the survey would provide an estimate of L(0,1)/l(0), where L(0,1) is the persons-years-lived under age 1 for a radix of l(o) = 1.00. Implied levels of mortality and expectation of life can be estimated once the pattern of mortality is determined. The technique is inherently sensitive to the pattern of mortality used.

37. In the 1976 Census of Swaziland the resident African women 15 years and over were asked to state the number of children born to them in the 12 months preceeding the census and to indicate whether the child was alive or dead by the time of the census. Table 3.4, demonstrates how such data can be used to estimate the level of mortality once the pattern of mortality is known.

<u>Step One</u>

Let B(i) and S(i) represent the number of babies born and those surviving among them in the last 12 months preceeding the census or survey to women aged i at the time of census or survey. These are the figures in column 2 and 3 respectively in Table 3.4

For each i, calculate the proportion surviving, PS(i), as

$$PS(i) = \frac{S(i)}{B(i)}$$

For i = 5, the 1976 Swaziland data recorded B(5) = 2059; S(5) = 1862

and therefore

$$PS(5) = \frac{1862}{2059} = 0.904$$

Using the same procedure, the column 4 was computed from columns 2 and 3. These proportion surviving are similar to the ratio of person-years lived under age 1 to the radix in a life table i.e., L(0,1)/l(0). For each ratio, therefore, an estimate of the level of mortality implied by each figure can be determined by a process of matching. However, in this case the matching cannot be directly effected because the proportions surviving are for both sexes and the life table ratios are for single sex. To make computation easier, two levels can be estimated, one based on male life table and another on females and the result combined by an assumed sex ratio. The correct procedure should have been to generate L(0,1) for males and females combined and the level determined by matching. The earlier one, though crude, will be satisfactory for our purpose. The estimated levels are given in column 6.

38. Table 3.4 shows that babies born to old women experience heavier mortality as compared to those in the age group 20-34. Though slightly better, the youngest women also appear to experience higher infant mortality than those for the 20-34 age group. Whether this is an indication of higher risks of mortality of single parity young women as well as higher order parities of old women, is difficult to assess because data on births in the last 12 months are usually under reported and certainly those who died will more frequently be omitted than the surviving ones. Infact, with a population of 494,000, the reported births produce only a birth rate of about 41 which is low and indicates some omission of births. Indications are that omission of births that occur in a specified reference period are not invariant by age. Therefore, the differential in implied mortality rates between the older age groups as compared to those depicted by age group 20-34 might be explained by age misstatement errors and differential perception in the 12 months period. Apart from problems of errors, the method also suffers from selectivity. In situations where the birth and instantaneous death of a baby occur at the beginning of the year followed by subsequent pregnancy, and another birth in the same year, the earlier event will escape recording and thus bias the estimate.

39. Generally, an over all level of mortality of 11.8 in the North family of Coale-Demeny set was found to be consistent with the implied level estimated from indirect techniques using the Trussell variant (Mabuza, 1981).

Table 3.4 Estimation of Mortality from Survival Status of Births in the Last 12 Months - Swaziland 1976

Age Group x-x+4	Age Index i	Birth Last 12 Months B(i)*	Those Surviving S(i)*	Proportion Surviving PS(i)	Corres-ponding Level
1	2	3	4	5	6
15-19	1	3322	3014	0.907	11.0
20-24	2	5866	5382	0.917	12.1
25-29	3	4749	4377	0.922	12.7
30-34	4	2876	2654	0.923	12.9
35-39	5	2059	1862	0.904	10.6
40-44	6	873	783	0.897	9.9
45-49	7	481	421	0.875	7.8
Total		20226	18493	0.914	11.8

*Source: Mabuza, S.E. "Some Aspects of Mortality in Swaziland" Unpublished M.A. Thesis, held by the University of Ghana, Legon, 1981, p.49.

Child survival data from fertility history

40. The World Fertility Survey (WFS) has introduced a significant improvement in the art of fertility data collection. Its schedule contains a fertilty history that collects the date of birth of each child and the date of death of each child that has died by the survey date. Such information makes itself amenable to the measurement of mortality using the life table technique. One advantage of such data is that they help investigators to calculate probabilities of dying that do not require reference tables (models). In addition, studies can be made for birth cohorts, time periods, as well as studies by birth order, etc. The process of computation is given in a Technical Bulletin of the WFS (Smith,1980). A brief description of the technique as applied to all children of women 15-49 is given in Table 3.5.

41. Since the aim is to estimate infant and child mortality, all children born to women 15-49 are grouped into two, ie. those who died and those who are alive.

Let x = age in years; n = width of intervals in years.
$_nD_x$ = number of children dead between age x-x+n years
$_nA_x$ = number of children alive at the time of survey and who are aged x-x+n years.
w = the highest age of the eldest child.

42. The number of children who start age x alive will be given

by $= \sum_{i=x}^{w} ({}_nD_i + {}_nA_i)$

In the absence of death, each of these peoplė will live a duration of life equal to the size of the interval, n. All of them will thus live

$$n. \sum_{l=x}^{w} ({}_nD_i + {}_nA_i)$$

43. But since some of them would have died in the interval and others were born in the interval, their contribution to the total person years lived will be shorter than n. Therefore the total person years lived will be estimated as

$$_nE_x = n \sum_{i=x}^{w} ({}_nD_i + {}_nA_i) - n/2[({}_nD_x + {}_nA_x)] \quad (6)$$

In equation 6, an assumption is made that those who died and those who are still alive but did not yet complete the age interval x to x+n are uniformly distributed. Once the person years-lived are estimated, the central death rates are computed by

$$_nM_x = \frac{_nD_x}{_nE_x} \quad (7)$$

The death rates can then be converted into probability of dying from age x to x+n, $_nq_x$, by the relation

$$_nq_x = \frac{2.n._nM_x}{2+n._nM_x} \quad (8)$$

The relationship given in equation (8) represents an approximation; It is based on the assumption that the life table survivalship function, l(x), is linear between ages x and x+n. Though the assumption is not correct, its effects can be reduced by considering shorter age intervals especially in age group 0-1. Neo-natal and post-neo-natal mortality rates can also be studied by splitting the 0-1 interval to under one month and one month to eleven months using the same method. The $_nq_x$ can then be converted into probability of surviving from age x to x+n, by the relationships

$$l(x+n) = (1-{}_nq_x)l(x) \quad (10)$$

for a radix of l(o) = 1.

Table 3.5 Illustration of the Calculation of Time Lived, Mortality Rates and Survival Probabilities, by Age Nigeria 1981-82 Fertility Survey

Age Group x-x+n	Age Index i	Length of Interval in Years	Deaths in Interval Di	Living at Time of Survey and age i	Number of People Who Started Interval i Alive (Di+Si)	Time Lived $_nE_x$ in Years	Annual Mortality Rates $_nM_x$	Probability of Dying Between Age x and x+n, $_nq_x$	Survival to Age x l_x	Probability of Dying Between Birth and x, $_xq_o$
1	2	3	4	5	6	7	8	9	10	11
0-1/12	1	1/12	1462	84	29989	2435	0.6004	0.0244	1.000	0.0244
1/12-1	2	11/12	1422	1621	28443	24678	0.0576	0.0257	0.9756	0.0495
1 - 2	3	1	957	1605	25400	24119	0.0397	0.0195	0.9505	0.0680
2 - 5	4	3	1049	5103	22838	59286	0.0177	0.0259	0.9320	0.0821
5+	5	-	449	16237	16686	x	x		0.9079	-

Problems With Fertility History Data

44. The problem with fertility history data as a source of mortality estimation arises from a conflict between the need to measure mortality which requires a large sample and the need for a careful field control, which can only effectively be done on a small-scale. For mortality studies, large sample sizes are required because deaths are usually rare events compared to births. So, samples intended to measure fertility are unlikely to give enough cases to study mortality. Besides, fertility history data suffer from truncation which complicates their interpretation. For instance, computation of the level of mortality some 15 years before the survey would only be based on the experience of women under 35 years then because the fertility history data limited its questions only to those under 50 years of age. Besides, it is the mortality experience of those whose mothers were alive that the estimates will be based on. The other problems which are specific to these type of data are omission and misplacement of dates of birth and death. Potter (1977) and Brass (1977) amongst others have identified systematic biases arising from respondents' over or under-estimation of the length of time periods either quite close or quite remote from the date of survey. The effect of the latter will be mainly to distort the trend or give a wrong impression of a trend even where there is none, the former affects the level of mortality. Evidence exists that women omit living children. It is therefore most likely that the same women will omit dead children even more than the living in the reporting of the dead children. The problem of omission increases with age of mother.

45. Another problem whose effect has not been studied carefully, at least in the African context, is the effect of imputation of the dates of birth and death of children of mothers who are ignorant about the exact dates of birth and death of their various children. In situations where birth certificates, weighing cards, etc are not issued to the children at birth it will be too much to expect the month and year of birth of the dead and living children especially of those events that occured long time in the past to be remembered by the mothers.

Inter Survey Mortality Estimation Using Hypothetical Cohorts

46. One other method that will be used to estimate mortality levels for those countries that have collected data on children ever born and children surviving or dead from two censuses and/or surveys separated by five or ten years is that of Zlotnik and Hill (1981). The method helps to determine the level of mortality of the interval between two censuses and/or surveys. After calculating the average number of children ever born and children dead by age group of women for the two periods, cohort changes are obtained by subtracting the initial values for each cohort from the final values of the same

cohorts (Zlotnik and Hill, 1981). The assumptions underlying the technique are that fertility is not changing fast and that the female population aged 15-19 is sectionally closed. In addition, it is assumed that the completness of reporting of births and deaths in both censuses and/or surveys were similar and that ages were reported to the same degree of accuracy.

47. For data given at two points of time, the technique is to generate a mean number of children ever born and children dead for a hypothetical cohort which is assumed to be exposed to the inter-survey vital rates. The parities and the mean number of children dead are calculated by cumulating the cohort inter-survey increments by age. Specifically,

48. Let P1(i), P2(i) and Ph(i) be the mean number of children ever born to age group i, in the first, second censuses (surveys) and hypothetical cohort respectively. Again let, D1(i), D2(i) and Dh(i) be the mean number of children dead to the same age group and for the same populations. The mean number of children ever born to the hypothetical cohort will be defined as:

$$Ph(i) = \begin{cases} P2(i) \text{ for } i = 1,2 \\ Ph(i-2) + [P2(i) - P1(i-2)] \text{ for } 3 \leq i \leq 7 \end{cases}$$

Similarly, the mean number of children dead to the hypothetical cohort is defined as

$$Dh(i) = \begin{cases} D2(i) \text{ for } i = 1,2 \\ Dh(i-2) + D2(i) - D1(i-2) \text{ for } 3 \leq i \leq 7 \end{cases}$$

Once the mean parities and the mean number of children ever born for the hypothetical cohort* are calculated the proportion dead are estimated and the Trussell variant applied.

---

* The above equations are valid for ten year intercensal interval. For five year interval, the equation is :

$$Ph(i) = \begin{cases} P2(i) \text{ for } i=1 \\ Ph(i-1) + [P2(i) - P1(i-1)] \text{ for } 2 \leq i \leq 7 \end{cases}$$

Dh(i) will also be similarly modified.

Application of the Zlotnik and Hill Method to Kenya Data

49. The data required consists of mean parities and average number of children dead by age of women for two points in time preferably separated by five or ten years.

50. In both the 1969 and 1979 censuses of Kenya, data on the survival status of children ever born to women in the reproductive ages were collected. The technique by Zlotnik and Hill will be used to estimate the level of infant and child mortality for the inter censal period.

51. From Table 3.6, figures in columns 2, and 3 are the mean number of children ever born and the mean number of children dead for 1969 and those for 1979 are given in columns 4 and 5.

If P1(i), D1(i), P2(i), D2(i), Ph(i) and Dh(i) are as defined above

Ph(1) = P2(1)=0.3206; Ph(2)=P2(2)=1.8529; Dh(1)=D2(1)=0.0372 and Dh(2)=D2(2)=0.2316.

For i=4, Ph(4) = Ph(2)+P2(4)-P1(2) = 1.8529+(5.3881-1.3705)
= 1.8529+4.0176 = 5.8705

Dh(4) = Dh(2)+D2(4)-D1(2) = 0.2316+(0.8944-0.2008)
= 0.2316+0.6936 = 0.9252.

Similarly, the mean parities and mean children dead were estimated for the hypothetical cohort and the figures inserted in column 6 and 7. The proportion dead for the hypothetical cohort are then computed as in the example for 1979 above. From the mean parities of hypothetical cohorts, PAR1 and PAR2 are also calculated and the multipliers determined in the same way as in the case of the actual population. The results are provided in column 9. The product of the figures in column 8 and 9 are shown in column 10 which are the probabilities of dying before exact ages 1,2,3,5, ..., 20 in the intersurvey period. The probabilities of surviving from birth to exact ages and the corresponding levels in the North Model life table of Coale and Demeny are given in the last two columns.

52. The figures indicate that Kenya had a level of mortality of about 15 in the period 1969 to 1979. This implies an expectation of life at birth of 55 years for females and 51.5 for males. As expected, the ${}_1q_0$ implied level indicated much heavier mortality than the others. The ${}_2q_0$ implied mortality was also slightly out of place than the others though not as bad as the ${}_1q_0$. Apart from these two, one other feature worth noting is the alternating nature of the estimated levels - the age groups starting in digit five implying a higher level than those starting with digit zero. The fluctuation is likely

Table 3.6 Application of Zlotnik and Hill Method to Estimate Inter-Census Infant and Child Mortality - Kenya 1969-1979

Age Group	Mean Number of Children Ever Born and Children Dead for for 1969		Mean Number of Children Ever Born and Dead for 1979		Mean Number of Children Ever Born and Dead to Hypothetical Cohort		Proportion Dead for Hypothetical Cohort	Trussell Multipliers M(j)	Index to which Probability of dying refers, j	Probability of Dying $_jq_0$	Probability of Surviving to Exact Age j	Corresponding Levels
	P1(i)	D1(i)	P2(i)	D2(i)	Ph(i)	Dh(i)	dh(i)					
1	2	3	4	5	6	7	8	9	10	11	12	13
15-19	0.2357	0.0301	0.3206	0.0372	0.3206	0.0372	0.116	1.0270	1	0.119	0.881	12.6
20-24	1.3705	0.2008	1.8529	0.2316	1.8529	0.2316	0.125	0.9841	2	0.123	0.877	14.6
25-29	2.8468	0.4945	3.6521	0.5149	3.737	0.5220	0.140	0.9400	3	0.132	0.868	15.0
30-34	4.2090	0.8515	5.3881	0.8944	5.8705	0.9252	0.158	0.9767	5	0.154	0.846	14.8
35-39	5.3354	1.2319	6.4703	1.1970	7.3605	1.2245	0.166	1.0395	10	0.154	0.827	15.1
40-44	6.1091	1.6061	7.0215	1.5237	8.683	1.5974	0.184	1.0265	15	0.189	0.811	14.9
45-49	6.5487	1.9882	7.1735	1.8149	9.1986	1.8075	0.196	1.0082	20	0.198	0.802	15.2

to be a reflection of the nature of age misstatement or an error carried over from the age groups 15-19 and 20-24 which form the base for the generation of the hypothetical average number of children ever born and children dead for those above age 25.

53. Further, it is evident from the comparison of the estimates based on the Zlotnik and Hill method and the Trussell methods based on one census of 1979, that except for the results of the first age groups ie., 15-19 and 20-29, the Zlotnik and Hill method implied a lower mortality than the one census method because the implied levels refer to the mid-point of the inter-censal interval, ie. 1974, while the 1979 results refer to various dates prior to the census which are indicated in column 13, Table 3.3. The results are consistent with the claim that mortality has been on the decline in Kenya since middle of the 1960s.

54. The utility of this technique will be limited where censuses are not separated by five or ten years as has been the case in most sub-saharan African countries.

55. Estimating inter-survey mortality rates based on two dates separated by five or ten years can serve a number of purposes. In the first instance it can provide a sensitive check on the consistency of the two sets of data. Scondly, it can also help avoid or reduce the effect of declining fertility on the estimated child mortality rates (U.N., 1983). Since the mean parities and mean children dead for the hypothetical population beyond age 20 or 25 are generated by cumulating the experience of those aged 15-19 and/or 20-24 in the second census or survey, the errors or biases in the two groups are carried over alternatively in all the age groups. In view of the fact that usually women aged 15-19 and 20-24 indicate higher than average mortality experience, the $_jq_0$ for all $j \geq 3$ will be subject to these biases. Errors brought in by younger cohorts may also bias the estimates based on inter survey method for women 25 years and over.

Other Methods

56. Various other methods of estimating infant and child mortality under varying assumptions are available. These are by Kraly and Norris (1978), Preston and Palloni (1978), Palloni (1980) and Feeney (1980). Most of these have not been evaluated under African situations. Nevertheless, most of the indirect techniques give similar results (Hill, 1984) and therefore no attempt is made to either evaluate them or apply them here. For those who are interested to read about the techniques, a concise description of each of these techniques with illustration is given by Venkatacharya (Forthcoming). For further aditional information readers can refer to the authors' original papers as well.

Adult Mortality

57. The most direct method of estimating mortality is from data on deaths in vital registration systems. Unfortunately such a system is generally faulty and inadequate. In the absence of such data, questions on deaths in a household in the year preceeding a survey or census would have provided the necessary information for the estimation of adult mortality. Infact, in the early African censuses and sample surveys such data were collected. However, experience on asking direct questions on deaths of people in a specified reference period were useless because they usually give extremely very low mortality rates. In fact, most censuses have dropped including the question on deaths in the 12 months preceeding the census or survey for sometime now. Some of the reasons alluded for gross omission of the deceased in a given reference period preceding the census/survey are: failure on the part of heads of households to delimit clearly the reference period, dissolution or reformation of households after the death of an older adult, the unpalatability of the event, etc. Recently, however, various methods of correcting for under enumeration of deaths have been developed based on various assumptions. The methods are the Brass growth balance method (Brass,1975), the Preston-Coale death distribution method (Preston et al., 1980), the Courbage-Fargues (1979) method, the Preston and Hill(1980) method, and the Bennett and Horuichi (1981) methods. The paper will discuss the Brass growth balance and the Preston and Coale methods. For a review of all the other techniques, readers are advised to refer to an article by Preston (1984) on "Use of Direct and Indirect Techniques for Estimating the Completeness of Death Registration Systems".

Direct Methods of Estimating Adult Mortality

Brass Growth Balance Method

58. In a closed population where the data on age and death distribution are correctly recorded

$$N(x) = r(x+).N(x+) + D(x+) \qquad (10)$$

where N(x) is the number of people at age x, r(x+) is the rate of growth of the population aged x and over, N(x+) and D(x+) are the population and the deaths in the population aged x and over. Dividing both sides of equation (10) by N(x+) gives

$$\frac{N(x)}{N(x+)} = r(x+) + \frac{D(x+)}{N(x+)}$$

In a stable situation, the rate of growth does not vary from age to age and since Brass assumes stability to derive his equation, the r(x+)is replaced by r and the equation becomes:

$$\frac{N(x)}{N(x+)} = r + \frac{D(x+)}{N(x+)} \qquad (11)$$

If registered or reported deaths represent a constant proportion of true deaths at each age,

ie., $D'(x+) = C.D(x+)$ (12)

$D'(x+)$ is the reported number of deaths to people x years and over; $D(x+)$ is the true number of deaths to the same category of people and C is the degree of completeness of death recording. Substituting (12) into (11) gives:

$$\frac{N(x)}{N(x+)} = r + \frac{1}{C}\frac{D'(x+)}{N(x+)} \qquad (13)$$

Equation (13) is a straight line where the intercept is the rate of growth and the slope is the reciprocal of the completeness of death registration. The values of r and C can be estimated by a number of methods such as the group mean method, least squares method, etc., but before applying these methods the N(x)'s would have to be estimated. Though, again, there are a number of ways of obtaining the values of N(x) , a simple but effective method will be to use the formula:

$$N(x) = [N(x-5,x-1) + N(x,x+4)]/10 \qquad (14)$$

Taking $V = D'(x+)/N(x+)$ and $W = N(x)/N(x+)$, a plot of the points (V,W) can be made on a graph paper and the scatter studied. The linear fitting can then be made after the outliers are discarded.

59. After the reciprocal of the slope of equation (13) is estimated, adjuted age specific death rates can be computed by inflating the observed deaths by a factor of C.

60. Table 3.7 shows the application of this method to 1967 Tanzanian census data (Mainland) where the number of deaths in the 12 months preceeding the census were collected.

61. Columns 2 and 3 give respectively the age distribution of females and the number of female deaths to each age group in the year preceeding the census. Let these be denoted by N(x, x+4) and D(x, x+4) for any age group x to x+4. Column 4 gives the estimated populations at exact age x using equation (14).

For example, N(10,14) = 579354; N(15,19) = 558899 and hence

N(15) = (579354 + 558899)/10 = 113825

The estimated population at exact ages are entered in Column 4. Columns 5 and 6 are cumulated populations and deaths from the oldest to the youngest respectively and are denoted by N(x+) and D(x+). For x = 60+,

$$N(60+) = N(60,64)+N(65,69)+N(70,74)+N(75+)$$
$$= 111865 + 72191 + 60753 + 179961 = 424770$$

$$D(60+) = D(60,64)+D(65,69)+D(70,74)+D(75+)$$
$$= 1986 + 1044 + 1704 + 5979 = 10713.$$

The "birth day" rates and the death rates to age x and over are then calculated using:

$N(x)/N(x+)$ and $D(x+)/N(x+)$ and the results entered in columns 7 and 8.

For example, for x = 60, N(60) = (100492 + 111865)/10 = 21236 Therefore, since N(60+) and D(60+) are estimated above, the "birth day" rate = N(60)/N(60+) = 21236/424770 = 0.050 and the death rate to those 60 years and over = D(60+)/N(60+) = 10713/424770 = 0.022.

The set of ratios given for $5 \leq x \leq 70$ can then be plotted on a graph and the scatter studied in order to purge the outliers. A straight line curve of the form W = a+bV with the ratios D(x+)/N(x+) and N(x)/N(x+) forming the V- and W- coordinates is fitted. As mentioned earlier, a number of curve fitting procedures can be used but here the group mean method was prefered because it is less sensitive to errors at the extremes. The entire age range is divided into three groups; the first group contains ages 5 to 29, the second, ageas 30 to 49 and the third, ages 50 to 74.

62. The straight line was fitted on the mean points of the first and last groups. If (V1, W1) and (V2, W2) are the mid-points of the two groups, the slope is estimated as:

(Slope) = (W2-W1)/(V2-V1) and

Intercept = W2 - V2.(Slope).

For Tanzania 1967, V1 = (0.009+0.009+0.009+0.009+0.01)/5
= 0.0092
V2 = (0.019+0.022+0.0250+0.028+0.032)/5
= 0.0252
W1 = (0.040+0.037+0.032+0.037+0.045)/5
= 0.0382
W2 = (0.057+0.053+0.050+0.059+0.055)/5
= 0.0548
Slope = (0.0548 - 0.0382)/(0.0252 - 0.0092)
= 1.0375
Intercept = 0.0548 - 1.0375 x 0.0252 = 0.0287

$$\text{The degree of completeness} = \frac{1}{\text{Slope}} = \frac{1}{1.0375} = 0.96 \text{ and}$$

the rate of growth = 2.87%.

The results indicate that 96% of the deaths in the 12 months preceeding the census were reported. This is quite high but appears to be reasonable considering that the implied rate of

Table 3.7 Application of the Brass Growth Balance Method to Tanzania "(Mainland) Females, 1967 Census Age Distribution and Reported Deaths in the 12 Months Preceeding the Census

AGE GROUP x-x+4	REPORTED FEMALE Population	Deaths	Population at Exact Age x, N(x)	Cumulated Population N(x+)	Deaths D(x+)	RATIO OF N(x)/N(x+)	D(x+)/N(x+)
1	2	3	4	5	6	7	8
0-4	1082729	78481	-	6113856	124737		
5-9	937761	10780	202049	5031127	46256	0.040	0.009
10-14	579354	4610	151712	4093366	35476	0.037	0.009
15-19	558899	3928	113825	3514012	30866	0.032	0.009
20-24	528402	3499	108730	2955113	26938	0.037	0.009
25-29	556966	3067	108537	2426711	23439	0.045	0.010
30-34	387987	2127	94495	1869745	20372	0.051	0.011
35-39	326671	1568	71466	1481758	18245	0.048	0.012
40-44	226752	2053	55342	1155087	16677	0.048	0.014
45-49	226845	1313	45358	928335	14624	0.049	0.016
50-54	176228	1597	40307	701490	13311	0.057	0.019
55-59	100492	1001	27672	525262	11714	0.053	0.022
60-64	111865	1986	21236	424770	10713	0.050	0.025
65-69	72191	1044	18406	312905	8727	0.059	0.028
70-74	60753	1704	13296	240714	7683	0.055	0.032
75+	179961	5979	-	179961	5979		

increase was 2.87% which is consistent with the estimated rate of growth for that time. The analysis of the 1967 census data on fertility and infant and child mortality indicated a rate of growth of 2.6-2.7% for the 1957 to 1967 period. However, since the technique is sensitive to the method of fitting and the age range of fitting, the estimates of the degree of completeness and the rate of growth may be in error. Once the degree of completeness is estimated, the actual number of deaths by age are estimated by dividing the reported deaths by age by the degree of completeness i.e. $D(x,x+4) = D'(x,x+4)/C$. For Tanzania deaths in each age group should be divided by 0.96 before age specific death rates are calculated for the construction of a life table. To show the degree of arbitrariness in the method, the group mean method was applied to the age range 15-64 with 15-39 and 40-64 as the two groups. The fitting gave a slope of 0.978 and a rate of growth of 3.26%. The implication is that deaths were more than complete by 2% which seems ridiculous unless we assume that deaths that occured more than 12 months period were also included in the reports. It is worth noting that, fitting should avoid the age range that are affected by net migration.

63. The Brass growth balance method suffers from some basic problems. In the first instance the choice of the age range for fitting seriously affects the method. Discarding outlying point introduces an unnecessary arbitrariness. However, avoiding these points is also necessary if one wants to avoid a ridiculous result. Another problem with the technique is age-misstatement. Exaggeration of age by the dead more than the living will increase the segmented death rate which in turn will increase the estimate of C. The method is also sensitive to the violations of the assumption of stability. The effect of declining fertility in sub-saharan African countries might not be very serious because changes in fertility are assumed to be slow and perhaps it is likely that fertility changes must have taken place not for a long time in most of these countries. But, in areas where fertility has changed drastically, dramatic effects on the age distribution will bias the estimation of the degree of completeness. In those countries where the destabilization has started gradually and not long ago, the initial age at which estimation begins can be chosen in such a way that it avoides cohorts that have been reduced in size by fertility decline. Declining mortality, which most African countries have been going through, leads to increasing rates of growth by age which in turn increases the slope but reduces the completeness.

Preston-Coale death distribution method

64. In this method, the degree of completeness of death recording/reporting, C, is measured as

$$C = \sum_{x=a}^{w} N'(x,x+4) / \sum_{x=a}^{w} N(x,x+4) \qquad (14)$$

where N'(x,x+4) is the estimated population based on reported deaths and N(x,x+4) is the enumerated population in the age group x to x+4. The rationale of this method is that in a stable population, the number of deaths that will occur in a cohort aged a in the future is related to current distribution by age after the appropriate inflation has been made for the differences in their cohort sizes at birth. In other words, the number of people who died at age x this year is greater than those who died to the same age x, a year ago by, EXP(rx).

65. In a stable situation, the annual number of births are assumed to increase from year to year by an exponential growth law. For instance, this year's births will be larger than last year's by e^r where r is the constant rate of growth. Similarly, this years births will be larger than births t years ago by e^{rt}. Since mortality is also constant from year to year, the number of deaths this year to those who are aged x will be smaller than the number of deaths to a cohort born t-years later by e^{-rt}. Therefore, the number of deaths to a population aged x now can be estimated from reported distribution of deaths by age. This is done by the equation

$$N'(a) = \sum_{j=a}^{w} D(j)\ \exp(r(j-x)) \qquad (15)$$

where N'(a) is the estimated population at age a; r is the rate of growth and D(j) is the number of deaths at age j and w the upper limit of age. In the absence of errors, the estimated N'(a) will be exactly the same as the actual or true N(a). A ratio of the estimated to the reported population, N'(a)/N(a), will therefore give an estimate of the degree of completeness.

66. The estimated five years age group populations can be determined by using

$$N'(x,x+4) = 2.5.(N'(x)+N(x+5)) \qquad (16)$$

Once the five year age groups are estimated, the degree of completeness can also be estimated as

$$C = \sum_{j=x}^{w} N'(j,j+5) / \sum_{j=x}^{w} N(j,j+5) \qquad (17)$$

This is a comparison of the populations age x and over in the estimated and reported groups. By varying the value of j, a series of estimates can be determined and a median or a mean selected.

67. In the Preston and Coale method, the estimation of N'(j) for the open ended age group gives a problem. To estimate this population with ease, the Manual x has given a table of coefficients. The coefficients which are reproduced in Table 3.8 are given for the four patterns of mortality of Coale and Demeny (1966) for estimating the length of time for the open interval Z(A) such that

$$D(A+)Exp(r.Z(A)) = \sum_{j=A} D(j)Exp(r(j-A)) \qquad (18)$$

It is based on the ratio of deaths over age 45 to deaths over age 10 and the population growth rate. The estimating equation is

$$Z(A) = a(A)+b(A).r+c(A)Exp(D(45+)/D(10+)) \qquad (19)$$

The values of A range from 45 to 85 in steps of 5. Once Z(A) is estimated

$$N'(A) = D(A+)Exp(r.Z(A)) \qquad (20)$$

Once the population at the lower boundery of the open age interval, N'(A), is estimated using equation (20), a recursive equation of the form,

$$N'(a) = N'(a+5).Exp(5r) + D(a,a+4).Exp(2.5r) \qquad (21)$$

can be derived from equation (15) which will be used in the computation of the estimated population for any age a, $0 \leq a \leq A$.

68. One advantage of the method is the process of cumulation which will tend to minimize the effect of age misstatement errors. The technique is however sensitive to the rate of growth. In general, sensitivity analysis of the procedure on simulated data showed that the technique is not robust to violations of the underlying assumptions though it is less sensitive than the Brass Growth balance method (Preston 1984). Infact, the Preston-Coale method produced errors less than half as large as the Brass procedure under declining mortality (Preston, 1984).

69. Application of the technique to Tanzanian data is given in Table 3.9. The various steps to be followed are:

<u>Step 1</u>:

An assumption is made that the deaths and population refer to the same point of time when infact that is not the case. The deaths refer to the mid-point of the year preceding the census while the population distribution refers to the census night. Adjustment for six months difference is ignored. The rate of growth of the female population between

Table 3.8 Coefficients for Estimation of the Age Factor for the Open Interval, Z(A), From the Ratio of Deaths Over Age 45 to Deaths over age 10 and the Population Growth Rate

AGE A	NORTH a(A)	NORTH b(A)	NORTH c(A)	EAST a(A)	EAST b(A)	EAST c(A)
45	-11.42	185.2	17.02	-15.87	174.3	18.06
50	-10.63	167.2	14.99	-15.14	158.5	16.06
55	- 9.78	147.8	12.96	-13.97	140.4	13.93
60	- 8.57	126.1	10.85	-12.10	118.8	11.60
65	- 6.83	101.6	8.62	- 9.43	93.9	9.05
70	- 4.53	74.6	6.28	- 6.07	66.5	6.38
75	- 1.91	47.1	3.98	- 2.52	39.3	3.81
80	0.46	22.7	2.00	0.37	16.8	1.73
85	1.82	6.4	0.67	1.79	3.5	0.48
	WEST			**SOUTH**		
45	-13.43	181.4	17.57	-15.26	183.4	18.23
50	-12.49	163.6	15.49	-14.91	168.4	16.36
55	-11.24	143.7	13.34	-14.22	151.2	14.38
60	- 9.50	121.2	11.07	-12.89	130.8	12.22
65	- 7.21	96.1	8.67	-10.67	106.4	9.80
70	- 4.48	69.2	6.23	- 7.53	78.4	7.15
75	- 1.64	42.9	3.91	- 3.84	48.8	4.47
80	0.72	20.5	1.98	- 0.47	22.6	2.14
85	2.03	5.9	0.70	1.47	5.6	0.63

Source United Nations (Manual X) Indirect Techniques for Demographic Estimations. Population Studies, NO. 81. ST/ESA/SER.A/81. Pg. 134.

1957 and 1967 is considered to be 2.7%. Generally, for a country that has two censuses, the intercensal rate of growth can be estimated using the exponential rate of growth, ie.,

$$r=\ln(N2/N1)/n$$

where N1 and N2 are the total female populations in the previous and most recent censuses. n is the intercensal interval in years.

Step 2:

The number of people at exact age x are estimated using equations (19), (20) and (21). Using equation (19), the average length of time (A) over which reported deaths in the open interval need to be inflated by the growth rate inorder to give the number of people who are alive at the beginning of the lower boundary of the open interval is estimated. For Tanzania 1967, the lower boundary of the open age interval, A, was 75 years.

Z(75) = a(75) + b(75).r + C(75).EXP (D(45+)/D(10+)
For North Model, Using Table 3.8,

a(75) = -1.91; b(75) = 47.1 and C(75) = 3.98
Since from Table 3.9, column 3, D(45+) = 1313 + 1597 + ...+ 1704 + 5979 = 14624 and D(10+) = 4610 + 3928 + 3499+...+1704 + 5979 = 35476,

Z(75) = 1.91+47.1(0.027) + 3.98.EXP (14624/35476)
= 5.37218

Using equation (20), the estimated number of people at age 75 years equals,

N'(75) = D(75+) EXP [r.Z(75)]
For D(75+) = 5979 and r = 0.027
N'(75) = 5979 EXP (0.027X5.37218) = 6912.3

Values of N'(70), N'(65), N'(60), ..., N'(5) can be estimated Successively using the recursive equation (21).
For example, for a = 70,

N'(70) = N'(75) Exp (5 x 0.027) + D(70, 74).Exp (2.5 x 0.027)
= 6912.3 x 1.14454 + D(70,74) x 1.06983

Since D(70,74) = 1704,

N'(70) = 6912.3 x 1.14454 + 1704 x 1.06983 = 9734.4

for a = 65,
N'(65) = N'(70) x 1.14454 + D (65,69) x 1.06983
= 9734.4 x 1.14454 + 1044 x 1.06983

Similar application of equation (21) to deaths in each age group in column 3 gives the estimated female population at exact ages x and the results are entered in column 4.

Step 3:

Female population for five year age groups are estimated using equation (16)

For instance, for age group 5-9,
N'(5,9) = 2.5 (N'(5) + N'(10)) = 2.5 (134059.4 + 107053.4) = 602782.

Repeated similar application of the formula will give results shown in column 5.

From the table, the estimated population for 0-4 and 75+ are not shown. In the case of the 0-4 age group, estimation of N'(0,4) is avoided because N'(0) depends on infant and child deaths which are usually subjected to a completeness of reporting/registration quite different from that of deaths at older ages. In the case of the open interval, estimation of N'(A+) will be subject to errors and thus, it is also avoided.

Table 3.9 :

An Illustrative Example of the Application of Preston-Coale Method of Estimating Completeness of Deaths, Tanzania 1967

Age Group x-x+4	Reported Female Popula-tion N(x,x+4)	Deaths D(x,x+4)	Estimated Population at age x N'(x) r = .027	Estimated Poulation N'(x,x+4)	Cumulated Population Reported N(x,74)	Estimated N'(x,74)	Ratio of N'(x,x+4) / N(x,x+4) Col.5 / Col.2 =	N'(x-75) / N(x-75) Col.7 / Col. 6 =	Estimated Popula-tion for r= .03	Estimated Popula-tion N(x,x+4) for r=0.03	Col.11/ Col.2
1	2	3	4	5	6	7	8	9	10	11	12
0-4	1082729	78481	237397.2						264709.8		
5-9	937761	10780	134059.4	602782	4851166	3215180	0.643	0.663	155027.6	696150	0.742
10-14	599354	4610	107053.4	490697	3913405	2612398	0.847	0.668	123432.4	563487	0.973
15-19	558899	3928	89225.2	480778	3334051	2121701	0.731	0.636	101962.3	465195	0.832
20-24	528402	3499	74285.9	339800	2775152	1712923	0.643	0.617	84115.6	383171	0.725
25-29	556966	3067	61634.1	281545	2246750	1373123	0.505	0.611	69152.8	314570	0.565
30-34	387987	2127	50983.9	233853	1689784	1091578	0.603	0.646	56675.0	258706	0.667
35-39	326671	1568	42557.3	195687	1301797	857725	0.599	0.659	46807.3	214100	0.655
40-44	226752	2053	35717.3	162513	975126	662038	0.717	0.679	38832.7	175879	0.776
45-49	226845	1313	29287.8	133461	748374	499525	0.588	0.667	31519.0	143574	0.633
50-54	176228	1597	24096.4	109143	521529	366064	0.619	0.702	25910.5	116826	0.663
55-59	100492	1001	19560.7	89289	345301	256921	0.889	0.744	20819.8	94527	0.941
60-64	111865	1986	16154.8	71033	244809	167632	0.635	0.685	16991.1	74432	0.665
64-69	72191	1044	12258.3	54982	132944	96599	0.762	0.727	12781.8	57036	0.790
70-74	60753	1704	9734.4	41617	60753	41617	0.685	0.685	10032.9	42719	0.703
75+	179961	5979	6912.3						7054.4		

Step 4:

Columns 6 and 7 show respectively the reported and estimated female populations cumulated from age x to age 74.

Step 5:

The ratios of the estimated female population to the reported population for five-year age groups, N'(x,x+4)/N(x,x+4), are shown in column 8, while column 9 shows the ratios of the estimated to the reported for age groups x to 74, ie., N'(x,74)/N(x,74).

For example, for x = 5
N(5,9) = 937761 and N'(5,9) = 602782
.·. C = 602782/937761 = 0.642
N(5,74) = 4851166 and N'(5,74) = 3215180
.·. C = 3215180/4851166 = 0.663

Therefore, by diving column 5 by column 2 and column 7 by column 6 the estimates of the degree of death reporting are made and results entered in columns 8 and 9 respectively.

70. From columns 8 and 9, it is clear that cumulation reduces the range of the estimates of the degree of completeness of death reporting by the two methods. However, if the median or the mean of the first nine figures is taken, the differences are minimal. Column 8 yields a mean of 0.653 and a median of 0.643 while column 9 gives a mean of 0.650 and a median of 0.659. In other words, the reported deaths would have to be inflated by about 50-55 per cent inorder to estimate the crude death rate or to construct an abridged life table.

71. Inflating, the reported female deaths by say 52% will lead to a female crude birth rate of 57.5 per 1000 which appears to be on the higher side for Tanzania 1967. Apart from age and other content and coverage errors of the population, the technique is sensitive to the rate of growth assumed.

72. To observe the effect of errors in estimating the rate of growth, the estimation of the female population was repeated for rate of growth of r = 0.03. The estimated female population for exact age x and for age group x to x+4 are given in columns 10 and 11 respectivcely. The computations are similar as in steps 2 and 3 described earlier. Column 12 refers to the ratio of the estimated to the reported, ie., column 11 divided by column 2.

73. From column 12, the median of the first nine estimated degrees of completeness is 0.742. This implies that only 74% of the actual death were reported in the 1967 census. The implied female crude birth rate will be about 54 per thousand which still appears to be on the high side.

Indirect Methods of Estimating Adult Mortality

74. Other methods of estimating adult mortality indirectly are:

1) Intercensal survival ratios if there are more than one census
2) Orphanhood method, and
3) Selecting the same level in a model life table system as that of childhood mortality.

75. The third option is more of an assumption rather than an estimation technique. The first and second options will be discussed in following sections. Under the first group of techniques are those that depend on the theory of stability as well as those that relax the stability assumption. In the former category, the Coale and Demeny age distribution and rate of growth or gross reproduction rate method (United Nations, 1967), the Arriaga(1968), Coale-Hoover(1958) and Stolnitz(1956) methods are those that assume stability. Among those that relax the stability assumption are the Preston-Bennette method (Preston and Bennett,1983), the Coale-Demeny projection and cumulation method (United Nations,1967) and the Carrier-Hobcraft methods (Carrier and Hobcraft,1971).

76. In the second group are included the survival of parents (orphanhood) as well as the survival status of first spouses (widowhood). Here the orphanhood method as developed by Brass and Hill will be discussed.

Methods Based on Stability of Age Distribution

77. In a closed population where age-specific fertility and mortality rates have remained constant for an extended period, a population with an unchanging age structure will be formed. The age structure will be typical of the vital rates. In otherwords two population with different age distributions will eventually have the same age structure if the populations are subjected to the same regime of fertility and mortality (Lotka and Sharpe, 1911). Such an age distribution will, among other things, have constant proportions at each age overtime, constant sex ratios and constant growth rates at all ages. Such a population is called a stable population.

Mathematically, a stable population is represented by:

$$C(x) = b\, e^{-rx} l(x) \qquad (22)$$

where C(x) is the proportion at exact age x; b is the birth rate and l(x) is the survivors from birth to age x, from a radix of l(o) = 1.0

From the above equation, it follows that

$$b = \frac{1}{e^{-rx} l(x)} \qquad (23)$$

and substituting this value of b in (18) gives

$$C(x) = \frac{e^{-rx} l(x)}{\sum e^{-rx} l(x)} \qquad (24)$$

From equation (20), it is clear that there are three unknowns, the rate of growth r, the probability of survival from birth to age x, l(x), and the proportion of the population at age x. Given any two of these unknowns, the third parameter can be estimated. Here the aim will be to estimate a mortality index given an accurate age distribution and a rate of growth. The index of mortality to be estimated can be the crude death rate or the expectation of life at birth.

Coale and Demeny Method

78. Censuses provide sex-age distribution of the population by single years or at least by five years of age. The application of the Coale-Demeny method requires an estimate of the rate of growth before the census. Where the country had a previous census which gives at least the total population, an estimate of r can be made using either the geometric or exponential growth law for the intercensal period. In situations where net migration, territorial changes and/or under or over enumeration are suspected, the data will have to be adjusted before the calculations are made.

79. The application of the Coale-Demeny method involves the matching of a proportion of the population in an age group in one of the Model Stable populations that are generated for the four families of Coale and Demeny tables for each level of mortality and for rates of growth ranging from -1.0% to 5.0% in steps of 0.5%.

80. Two problems arise in the application of the technique. The first one is the choice of the pattern of mortality and the other is the quality of age reporting. In the case of the former, the suggested solution will be to select population parameters that are invariant with the pattern of mortality. Such indices are the expectation of life at age 5 or age 10 (United Nations, 1967).

81. Infact the proportion under 15 for both sexes and the rate of growth are expected to give an estimate of expectation of life at age ten which is robust to the choice of mortality pattern (Hill, 1984). The second problem which relates to the quality of data can be minimized by considering grouped data under specified ages rather than the age-specific proportions. When considering the proportion under specified ages, net movement of people across that age will only distort the estimate. For instance, a systematic shifting of people towards younger age groups will create an impression of high fertility in the past and a heavy mortality for a fixed rate of growth. The bias will be reversed if there is a systematic exaggeration.

82. A third problem is the assumption of stability. In Africa, it is difficult to assume that mortality has remained constant for a long time. Infact there is ample evidence to show that in the 1960's mortality was on the decline. Fortunately, with gradual declines in mortality, the Coale-Demeny technique can be used to estimate the vital parameters of a country from age distribution. The current levels will be reflected on the age distribution. A population which has experienced a recent gradual change in mortality is called a quasi-stable population.

Estimation of Mortality From Age Distribution and Rate of Growth

83. The example that is given for illustrating the applications of stable or quasi-stable theory to estimate mortality in Africa is that of Kenya 1979. The process involves the computation of the proportion for a given age to the total for a specified sex group.

Let N(x,x+4) be the number of females at age x to x+4. The proportion at age x to x+4 will be:

$$C(x,x+4) = N(x,x+4)/\sum_{x=o}^{w} N(x,x+4) \text{ where w is the last age}$$

However, if the classification includes age not stated group, the proportion should be calculated after the non-stated group are subtracted from the total for that sex. In Kenya 1979 Census, out of 7,719,948 females, there were 13833 whose ages were not stated. 1,421,385 females were also reported at age 0-4. The proportion in age group 0-4 is therefore calculated as

$$C(0,4) = 1421385/(7719948 - 13833) = 0.1844$$

Similar computations were made up to age group 40-45 and the figures cumulated from age 0 to ages 5,10,15, ..., 45. For age 15,

$$C(15-) = C(0,4)+C(5,9)+C(10,14) = 0.1844+0.1615+0.1329 = 0.4788.$$

These cumulated proportions are expressed in percentages and the results entered in column 1 of Table 3.9.

84. The next step is to calculate the rate of growth of the population. The exponential growth law was applied to the female populations in 1969 and 1979. The total female populations for the two censuses were 5,460,324 for 1969 and 7719948 for 1979. The computed intercensal growth rate was 3.46.

Table 3.10 Application of Coale-Demeny method to Kenya Females (Age Distribution and Rate of Growth Method)

Age x	Proportion Under Age x C(x-)	Levels and the Corresponding Proportion that Bracket the Observed Proportion Under Age x 12	13	14	15	16	18	19	Level os Mortality Corresponding to Observed C(x-) and r = 0.0346
1	2	3	4	5	6	7	8	9	10
5	18.44						18.50	18.24	18.23
10	34.59				34.85	34.41			15.59
15	47.88			48.10	47.57				14.42
20	59.40	59.76	59.19						12.63
25	68.30	68.40	67.86						12.16
30	75.32	75.45	74.93						12.25
35	80.66		80.68	80.21					13.04
40	84.88			84.92	84.50				14.10
45	88.43			88.73	88.37				14.83

85. With r = 3.46% and C(5-) = 18.44, a search is made in the North Stable population of the Coale and Demeny set to locate the level of mortality with these characteristics. Since the tables of Coale and Demeny are tabulated for rates of growth of -1% to 5% in steps of 0.5%, the observed rate of growth of 3.46% is not one of the tabulated figures. But, it lies between tabulated rates of growth of 3.0% and 3.5%. By interpolation, the level is determined and the result recorded.

86. For the Kenyan example, a C(5-) of 18.44 with a rate of growth of 3.46% lies between levels 18 and 19 on the Coale and Demeny North model stable population with proportions under 5 of 18.50 and 18.24 respectively. The level of mortality corresponding to C(15-) of 18.44 and r = 3.46% is thus:

$$\text{level for } C(5-) = 18+(18.44-18.50)/(18.24-18.50)$$
$$= 18.23$$

Similar computations were made for the other cumulated proportions under ages 10, 15, 20, ..., 45 and the results entered in column 10 of Table 3.10. The estimated levels are ordered in ascending or descending order and the median selected. For Kenya females, the median levels was 14.10. The life expectation is thus estimated as 52.75 years corresponding to level 14.1. The crude birth rate is also estimated by first obtaining the female birth rate and converting this to crude birth rate (including males and females) by the formula:

$$CBR = (\text{Female Birth Rate}) \times \frac{\text{female Pop.}}{\text{Total Pop.}} (1 + SRB),$$

where CBR = Crude Birth Rate and SRB = Sex Ratio Birth

$$\text{hence } CBR = 48.739 \times \frac{7719948}{15327061} \times 2.03 = 49.83 \text{ per thousand.}$$

87. Given the same data situations, there are other methods of solving equation (18) to estimate the level of mortality. One such method is due to Arriaga (1968).

Arriaga Method

88. In the Arriaga method the equation is written as:

$$C(x,x+4) = b\, e^{r(x+2.5)} L(x,x+4) \qquad (25)$$

or

$$\frac{C(x,x+4)}{L(x,x+4} = b.e^{-r(x+2.5)} \qquad (26)$$

Taking natural logarithms on both sides and re-arranging the terms,

$$\ln \left(\frac{C(x,x+4)}{L(x,x+4)} \right) = (\ln b - 2.5r) - rx \qquad (27)$$

Equation (27) is a straight line with slope equal to the negative of the rate of growth and (lnb-2.5r) the intercept. From equation (27), both the rate of growth and the crude birth rate can be estimated simultaneously. The estimated rate of growth can then be compared against the computed intercensal rate of growth which will help to examine the quality of data. In such a case, the method appears to be more a technique of evaluation than a tool of mortality estimation. However, if the rate of growth is known, a life table among one of the models can be selected by either least squares or other methods that give consistent birth and death rates.

89. The method is sensitive to the rate of growth and age misreporting. Although the fitting can avoid errors at older and young ages by eliminating them from the range of fitting, the quality of data can have a major impact on the estimated level of mortality. An application of the technique to the 1979 female age distribution of Kenya is given in Table 3.11 For Kenya, the proportions are given in column 2 . From experience on the mortality situation of the country under study a level of mortality is selected and its stationary population expressed per unit radix and is given in column 3.

The natural logarithm of the quotient of column 2 and 3 are given in column 4. For example, for Kenya 1979, and x = 10,

$L(10,14) = 3.76916$ (level 12); $C(10,14) = 0.1329$ and
$\ln[C(10,14)/L(10,14)] = -3.3450$.

A curve of the form $Y = A+Bx$ is then fitted for age ranges 0-64 with a transformed x value. The formula for transformation is $X = (32.5-x)/5$ where x is the mid point of an age group and 32.5 is the mid-point of the central five years age group of the range of fitting, ie 0-64.

Equation (27) after substituting 32.5-5X for x gives

$$\ln \left(\frac{C(x,x+4)}{L(x,x+4)} \right) = \ln b - 2.5r - r(32.5-5X)$$

$$= \ln b - 35.0r + 5rX$$

Table 3.11 An Illustration of Arriaga method to Kenya Females 1979

Age Group x-x+4	Proportion at age x-x+4	L(x,x+4) for level 12	$\ln\left(\frac{C(x,x+4)}{L(x,x+)}\right)$	Transformed X $X=\frac{32.5-x}{5}$	$X\ln\left(\frac{C(x,x+4)}{L(x,x+4)}\right)$	L(x,x+4) for level 13	$\ln\left(\frac{C(x,x+4)}{L(x,x+4)}\right)$	$X\ln\left(\frac{C(x,x+4)}{L(x,x+4)}\right)$
1	2	3	4	5	6	7	8	9
0-4	0.1845	4.24534	-3.13593	+6	-18.9261	4.32429	-3.15435	-18.8156
5-9	0.1615	3.89719	-3.18351	+5	-16.9681	4.01632	-3.21362	-15.9175
10-14	0.1329	3.76916	-3.34501	+4	-13.5174	3.90085	-3.37935	-13.3800
15-19	0.1152	3.67933	-3.46382	+3	-10.5019	3.81734	-3.50064	-10.3915
20-24	0.0890	3.58077	-3.69470	+2	- 7.4679	3.72412	-3.73395	- 7.3894
25-29	0.0702	3.46922	-3.90034	+1	- 3.9423	3.61793	-3.94231	- 3.9003
30-34	0.0536	3.34427	-4.13345	0	0.0	3.49904	-4.17869	0.0
35-39	0.0422	3.20588	-4.33032	-1	4.33032	3.36729	-4.37944	4.3794
40-44	0.0355	3.05619	-4.45539	-2	8.91078	3.22340	-4.50866	9.0173
45-49	0.0288	2.89835	-4.61152	-3	13.83456	3.06980	-4.66899	14.0071
50-54	0.0248	2.72208	-4.69831	-4	18.79324	2.89620	-4.76031	19.0412
55-59	0.0175	2.50720	-4.96472	-5	24.8236	2.68298	-5.03248	25.1624
60-64	0.0142	2.23134	-5.05712	-6	30.3427	2.40679	-5.13281	30.7969

$$\ln \left(\frac{C(x,x+4)}{L(x,x+4)}\right) = \ln b - 35r + 5rX$$

which is a linear curve of the form

$$Y = A+Bx \text{ where } A = \ln b-35r,\ B=5r, \text{ and } Y = \ln\left(\frac{C(x,x+4)}{L(x,x+4)}\right)$$

For Kenya, the solution for 13 age groups with 32.5 as center gives

$$B = \frac{\sum YX}{\sum X^2} \text{ and } \qquad A = \frac{\sum Y}{n}$$

For level 12, $\sum X^2 = 182$, $\sum YX = 31.24083$

B = 31.24083/182 = 0.1711653 and r = B/5 = 0.171653/5 = 0.0343.

Since, the implied rate of growth of 0.0343 is lower than the observed rate of growth of 0.0346, it looks that level 12 is on the lowerside. Level 13, was then tried and an estimate of r = 0.0351 arrived. The level of mortality was then determined by interpolating for the observed rate of growth between levels 12 and 13. The level of mortality that will reproduce the rate of growth of 3.46% is

$$\frac{13(0.0346-0.0343) + 12(0.0351-0.0346}{(0.0351-0.0343)}$$

$$= \underline{12.4}.$$

The choice of 0-64 for fitting is arbitary. Higher ages were avoided because age mistatement errors in the form of shifting usually affect the estimate. Shorter ranges centred on 30-34 were attempted for fitting, the results implied higher risks of mortality or lower than level 12 on the Coale-Demeny set. It is obvious from the discussion above that the technique is sensitive to the rate of growth. Therefore, differential completeness in censuses will severely limit the utility of the technique as a method of estimating mortality. Besides, age mistatement also reduces its power of estimation. Destabilization caused by fertility and mortality changes also affects the method seriously.

90. Two other methods using the basic stable equation in order to provide formulae for estimating the mortality level of a country are those of Coale-Hoover (1958) and the variant of Stolnitz methods (Ramachandran, 1977).

Coale-Hoover Method

91. In the Coale-Hoover method, the stable equation is written as:

$$L(x,x+4) = \frac{1}{b} . C(x,x+4) . e^{r(x+2.5)} \quad (28)$$

Since $\frac{1}{b}$ is a constant, it can also be written as

$$L(x,x+4) \propto C(x,x+4) . e^{r(x+2.5)} \quad (29)$$

For x > 5, the proportional values of L(x,x+4) are computed by equation (29) upto the last closed age group. For the open ended age group, the same equation cannot be used because the mid-point of the age group is unknown. Experience from model relationships can be adopted. In this study, since the open ended age group considered is 65+, an average age of 72.5 years appears to give a reasonably good approximation of the mid point for the range of levels of mortality of African countries. So, once T(65) ie., L(65+) is calculated, T(10) can be estimated as a sum of the estimated L(x,x+4) values for x = 10, 15, ..., 60 and T(65). l(10) is estimated as an average of L(5,9) and L(10,14), ie. l(10) = (L(5,9) + L(10,14))/10. The expectation of life at age ten, e(10) is thus given by:

$$e(10) = \frac{T(10),}{l(10)} \quad (30)$$

where:

$$T(10) = \sum_{x=10}^{60} C(x,x+4) \, EXP[r(x+2.5)] + C(65+) \, EXP(72.5r) \quad (31)$$

$$\begin{aligned} l(10) &= (L(5,9)+L(10,14)/10 \\ &= [C(5,9) \, EXP(7.5r) + C(10,14)EXP(12.5r)]/10 \quad (32) \end{aligned}$$

Since T(10) amd l(10) contain EXP(2.5r) as common factor

$$e(10) = \frac{T(10)}{l(10)} = \frac{10. \sum_{x=10}^{60} [C(x,x+4)EXP(rx)] + C(65+)EXP(70r)}{[C(5,9)exp(5r) + C(10,14)(EXP \ 5r)^2]} \quad (33)$$

92. In Table 3.12, the second column indicates the proportion of the population in each age group. Column 3 shows the exponent to which exp(5r) should be raised to estimate the stationary populations needed for estimating T(10), and T(0) in Coale-Hoover and Stolnitz methods respectively. The constant $[exp(5r)]^V$ for inflating the C(x,x+4) and C(x,x+4)/C(0,4) figures for Kenya 1979 for rate of growth r = 0.0346 are given in column 4. The numbers that are proportional to stationary population, L(x,x+4), for Coale-Hoover Method are entered in column 5. Summation of these figures for all x > 10 gives:

$$\begin{aligned} T(10) &= L(10,14) + L(15,19) + \cdots + L(60,64) + L(65+) \\ &= 2.00689 \end{aligned}$$

$$l(10) = [L(5,9)+L(10,14)]/10 = (0.192+0.18784)/10 = 0.03797$$

Table 3.12 An Illustrative Example of Coale-Hoover and Stolnitz Methods to Kenya Females 1979

Age Group x-x+4	Kenya Female 1979 r=0.0346	COALE HOOVER V	COALE HOOVER $[Exp(5r)]^{V}$	$_5L_x$ Col.2*Col.4	STOLNITZ $\frac{C(x,x+4)}{C(0,4)}$	$_5L_x$ = Col.4*Col.6
1	2	3	4	5	6	7
0-4	0.1845	0	1.00000	0.18450	1.00	1.0000
5-9	0.1615	1	1.18887	0.19200	0.87534	1.04066
10-14	0.1329	2	1.41340	0.18784	0.72033	1.01811
15-19	0.1152	3	1.68035	0.19358	0.62439	1.04919
20-24	0.0890	4	1.99771	0.17780	0.48238	0.96366
25-29	0.0702	5	2.37501	0.16673	0.38049	0.90366
30-34	0.0536	6	2.82356	0.15134	0.29051	0.82029
35-39	0.0422	7	3.35684	0.14166	0.22873	0.76780
40-44	0.0355	8	3.99083	0.14167	0.19241	0.76788
45-49	0.0288	9	4.74457	0.13664	0.15610	0.74062
50-54	0.0248	10	5.64065	0.13989	0.13442	0.75820
55-59	0.0175	11	6.70598	0.11735	0.09485	0.63607
60-64	0.0142	12	7.97251	0.11321	0.07696	0.61360
65+	0.0301	14	11.26837	0.33918	0.16314	1.83836
T_0						12.91810
T_{10}				2.00689		

For Stolnitz - T(0) = 12.9181 L(0,4).

T(0)/L(0,4) = 12.9181

Estimated expectation of life at age ten by Coale-Hoover method

$$e(10) = 2.00689/0.03798 = 52.835$$

From the models the level of mortality with expectation of life at age ten equal to 52.835 is level 13. The technique could have also been used for any other age except zero. The reason for selecting ten is because it gives a level of adult mortality which is independent of pattern. However, if the implied level is sought, its effectiveness reduces drastically. Generally, Coale-Hoover technique will imply a level of mortality which is slightly heavier than the actual figure because l(10) is usually on the higher side because of inflation of the 5-9 age group by the transfere across age 5 of those 3 and 4 years of age children while T(10) is on the lower side because of the transfere of children ten years and over across age ten.

93. The method thus involves, locating a level of mortality with an expectation of life at age ten equal to the computed figure given above. Note, once again that, a knowledge of the pattern of mortality is very essential for the technique. The technique is also sensitive to age mistatement and the estimated rate of growth. One advantage of the technique though is its avoidance of the age group 0-4 which is usually under enumerated. The same technique could also be used to estimate the level of mortality at other ages such as age 15,20,.... etc.

Stolnitz Method

94. This method is a variant of the Stolnitz method (Ramachandran, 1977). The equation used for estimting the level of mortality relates the stationary population L(0,4), and T(0) through the Lotka's equation.

$$\text{Let } L(0,4) = K \qquad (34)$$

From equation (24)

$$\frac{L(5,9)}{L(0,4)} = \frac{C(5,9)}{b}\, e^{7.5r} \cdot \frac{b}{C(0,4)e^{2.5r}} = \frac{C(5,9)}{C(0,4)}\, e^{5r} \qquad (35)$$

From equation (30) and (31)

$$L(5,9) = \frac{C(5,9)}{C(0,4)}\, e^{5r} \cdot K \qquad (36)$$

$$\frac{L(10,14)}{L(5,9)} = \frac{C(10,14)}{b}\, e^{12.5r} \cdot \frac{b}{C(5,9)e^{7.5r}} = \frac{C(10,14)}{C(5,9)}\, e^{5r}$$

$$L(10,14) = \frac{C(10,14)}{C(5,9)} . e^{5r} . L(5,9) \qquad (37)$$

From equations (32) and (33)

$$\therefore L(10,14) = \frac{C(10,14)}{C(5,9)} . e^{5r} . \frac{C(5,9)}{C(0,4)} . e^{5r} \quad K$$

$$= \frac{C(10,14)}{C(0,4)} . (e^{5r})^2 . K \qquad (38)$$

Similarly, for any age x, $0 \leq x \leq 60$

$$L(x,x+4) = \frac{C(x,x+4)}{C(0,4)} . (\exp 5r)^v K \quad \text{where } v = x/5 \qquad (39)$$

$$T(0) = \sum_{x=0}^{60} L(x,x+4) = \sum_{x=0}^{60} \frac{C(x,x+4)}{C(0,4)} [\text{Exp}(5r)]^v . k$$

$$+ \frac{C(65+)}{C(0,4)} [\text{EXP}(5r)]^{14} . K \qquad (40)$$

$$\frac{T(0)}{K} = \frac{T(0)}{L(0,4)} = \sum_{x=0}^{60} \frac{C(x,x+4)}{C(0,4)} \left[\text{Exp}(5r)\right]^v + \frac{C(65+)}{C(0,4)} \left[\text{Exp}(5r)\right]^{14} \qquad (41)$$

Here an assumption is made that A = 65 and that, the mid point for those who are 65 years and over is 72.5 years as assumed with the Coale-Hoover method.

94. The life table that has T(0)/L(0,4) equal to the right hand side of equation (41) is selected from among the life tables in a one parameter model system. It is important to note that one should have an idea of the pattern of mortality of the population under study.

95. An application of the technique to Kenyan Female 1979, is given in Table 3.12. Columns 6 and 7 are the ratios C(x,x+4)/C(0,4) and $[C(x,x+4)/C(0,4)](\text{Exp } 5r)^V$ respectively.

96. The summation of all figures under column 7 gives T(0). Hence for Kenya 1979 females where the rate of growth (r) is assumed to be 0.0346, the estimated T(0)/L(0,4) = 12.9181.

97. A search for a level of mortality with a ratio of T(0) to L(0,4) of 12.9181 will have to be made in one of the one-parameter model life table systems of Coale and Demeny(1966) to estimate the mortality level by this method. For Kenya, a ratio of T(0) to L(0,4) of 12.9181 is bracketed by levels 12 and 13 in the North family of Coale and Demeny. Interpolating for a T(0)/L(0,4) of 12.9181 gives a level of mortality of 12.4.

98. Under ideal conditions, i.e. where data are not subject to errors, all the four methods will give the same results except for the problems of approximation at the open-ended age group which can not be discarded as in the Coale and Demeny or Arriaga methods. Because of the heavy weight of the exponential growth term, the estimated rates will be correct to the nearest unit. The four techniques are however sensitive to the selected pattern. Among the four, the Coale-Hoover is the least robust method while the Stolnitz appears to be less sensitive to changes in the patterns of mortality. The sensitivity of the Coale-Hoover method lies in that models having similar patterns of adult mortality can be subject to varying levels of infant and child mortality.

Methods Based on Age Distribution from Two Successive Censuses

99. Lack of adequate data on vital statistics for estimating mortality has already been noted. In this section indirect techniques based on two age distributions collected from two successive censuses will be discussed. The techniques that will be discussed are i) the Carrier-Hobcraft Method, (ii) the Coale-Demeny Projection and Cumulation Method and (iii) the Preston-Bennett Methods. Other methods such as the Brass variant of Carrier-Hobcraft which use the logit smoothing method is not discussed here because they require relatively good age distribution.

Carrier-Hobcraft Method

100. The method is based on the following assumptions:

i) The population is closed to migration during the intercensal period

ii) The magnitude of age misstatement in the various reported age groups is not severe.

iii) The age pattern of mortality of the population conforms to one of the known pattern of mortality.

Based on these assumptions, the technique can be used to estimate the level of mortality in the intercensal periods.

101. The mechanics of the method starts by computing the ten-year cohort survival ratios;

If we assume that a population is sectionally stationary

$$\frac{L(x+10,x+14)}{L(x,x+4)} = \frac{N2(x+10,x+14)}{N1(x,x+4)} \qquad (42)$$

where L(z,W) is a stationery population in a life table age z to W years and N1(z,W) and N2(z,W) are respectively populations in a first and second censuses in age groups z to W.

Let $S(i) = N2(x+10,x+14)/N1(x,x+4)$ (43)
where S(i) is the ten years survival ratio for those aged x to x+4 years in the first census. From equations (42) and (43), it can be seen that

$$L(x+10,x+14) = S(i).L(x,x+4) \quad (44)$$

using equations (44),

$$L(10,14) = S(1).L(0,4)$$
$$L(15,19) = S(2).L(5,9)$$
$$L(20,24) = S(3).L(10,14) = S(3).S(1).L(0,4)$$
$$L(25,29) = S(4).L(15,19) = S(4).S(2).L(5,9)$$

etc.

Similarly for any age u, $5 < u \leq A-5$, where A is the beginning age of the open age interval, L(u,u+4) can be expressed as a product of the alternating ten years survivorship ratios depending on whether u is an odd or even multiple of five. Thus, L(u,u+4) can be represented as

$$L(u,u+4) = S(1).S(3).S(5). \ldots, S(2n+1).L(0,4) \quad (45)$$

for u an even multiple of 5

$$L(u,u+4) = S(2).S(4).S(6). \ldots, S(2n).L(5,9) \quad (46)$$

for u an odd multiple of 5

For an open ended interval, the total number of person years lived, T(A) , can be expressed as a function of L(A-10,A-6) and L(A-5,A-1) which will be estimated by equations (45) and (46).

Let $T(A)/T(A-10) = Z$

$$T(A) = Z.T(A-10) = Z(L(A-10,A-6)+L(A-5,A-1)+T(A)) \text{ , ie.,}$$
$$T(A)-Z(T(A)) = Z(L(A-10,A-6) + L(A-5,A-1)) \text{ and hence}$$
$$T(A) = (Z/1-Z) (L(A-10, A-6) + L(A-5, A-1))$$

The total person-years lived (or total number of persons in a stationary population) will be given as

$$T(0) = L(0,4)+L(5,9)+\ldots+L(A-10,A-6)+L(A-5,A-1)+T(A) \quad (47)$$

Substituting the equivalents of the person years lived in equation (47) and re-arranging the terms gives

$$\begin{aligned}T(0) = {} & L(0,4)[1+S(1)+S(1).S(3)+..+S(1).S(3)\ S(5)..S(2n+1) \\ & +(Z/(1-Z).S(1).S(3).S(5) \ldots S(2n+1)]+ \\ & + L(5,9)[1+S(2)+S(2).S(4)+ .. + S(2).S(4)\ S(6) .. S(2n) \\ & +(Z/(1-Z)).S(2).S(4).S(6) \ldots S(2n)]\end{aligned} \quad (48)$$

Let C and D be the constants in parenthesis which form the coefficients of L(0,4) and L(5,9) respectively.

i.e. $T(0) = C.L(0,4) + D.L(5,9)$

where

$$C = 1 + S(1)+S(1).S(3)+\ldots\ldots.+ S(1).S(3).S(5) ..S(2n+1) + (z/1-z).S(1).S(3)\ldots.S(2n+1)$$

and

$D = 1 + S(2) + S(2)S(4) + S(2)S(4)S(6) +..+ S(2)S(4)S(6).. S(2n) + (z/1-z)\ S(2).S(4).......S(2n)$

for $l(0) = 1$ the expectation of life at age zero is given by

$$T(0)/l(0) = e(0) = C.L(0,4)+D.L(5,9).$$

103. The procedure therefore reduces to locating a life table that satisfies this equation. In other words locating a life table whose expectation of life at birth is equal to the sum of the products of its stationary populations at ages 0-4 and 5-9 and the constants C and D respectively. This can easily be located by trial and error once the pattern of mortality is known.

104. The technique is inflexible in that, it is difficult to apply if intercensal interval other than five or ten years are involved as it requires tedious calculations. In addition, it is very sensitive to differential completeness in the census enumerations. Greater completeness in the second census will give an impression of lighter mortality than the actual prevailing level in the intercensal period. In most cases, the most recent census will be an improvement over its predecessor though deterioration in quality are also likely in some cases.

105. Another problem with the technique is the age of truncation. Experience with African data indicate that truncation below age 45 years gives unrealistic mortality estimates. Beyond age 45, implied levels alternate depending on whether the open ended age started with digit five or digit zero perhaps in response to the net pattern of age heaping. The other problem with the technique is its dependence on model life tables. But if L(0,4) and L(5,9) can be estimated from data on surviving children and children ever-born, its dependence on models could be reduced.

<u>An Illustrative Example of the Carrier-Hobcraft Method to Kenya 1969-1979 Data</u>

<u>Step 1</u>: Calculate the intercensal Cohort Survival ratios from the 1969 and 1979 population distributions. If N1(x,x+4) and N2(x,x+4) are the 1969 and 1979 populations by age and sex, respectively a set of S(i) values can be calculated using equation (43). For example,

N2(10,14) = 1025678 and N1(0,4) = 1046380.

Therefore S(1) = 1025678/1046380 = 0.9802.

The 10 year survival ratios computed using equation (43) are given in column 4 Table 3.13. Note that the last survival ratio is for the open age group 45+

<u>Step 2</u>: Divide the survival ratios into two groups, ie, those with ages starting in digits 0 and 5 respectively. Then with L(0,4) and L(5,9) taken seperately as equal to 1, the survival ratios in the two groups are multiplied succesively and are presented in columns 5 and 6. For instance the fourth value in

column 5 is obtained by multiplying the survival ratio at age 20-24 in column 4 by the third value in column 5 and so on. Note that the last two values in columns 5 and 6 are obtained by multiplying the last but one values in those columns by the value of (Z/(1-Z)) where Z is the survival ratio of the last open age group.

Table 3.13 An Application of the Carrier Hobcraft Method to the 1969-79 Female Population of Kenya

Age Group x-x+4	Female Population 1969	1979	10 Year Survival Ratio	Multipliers for:* L(0,4)	L(5,9)
1	2	3	4	5	6
0-4	1046380	1423936	0.9802	1.000	
5-9	893359	1246983	0.9955		1.00
10-14	663808	1025678	1.0353	0.9802	
15-19	554847	889316	0.9773		0.9955
20-24	450096	687234	0.9185	1.0148	
25-29	411245	542233	0.7926		0.9729
30-34	299241	413432	0.9163	0.9321	
35-39	264819	325951	0.8397		0.7711
40-44	201936	274193	0.9475	0.8540	
45-49	684593	222363	0.6971		0.6475
50-54		191365		0.8092	
55+		477264		1.8624[a]	1.4902[a]

* The values are obtained by using equation 44.
a) The value is obtained by multiplying the value above by (Z/1-Z)

$Z = 0.6971, \qquad Z/1-Z = 2.3014.$

$C = 1+0.9802+1.0148+0.9321+0.8541+0.8092+1.8624 = 7.4528$

$D = 1.00+0.9955+0.9729+0.7711+0.6475+1.4902 = 5.8772$

$T(0) = 7.4528\ L(0,4) + 5.8772\ L(5,9)$

for $l(o) = 1$,

$$e(0) = \frac{T(0)}{l(0)} = 7.4528\ L(0,4) + 5.8772\ L(5,9) \qquad (49)$$

I N P U T			OUTPUT	Difference
e(0)	L(0,4)	L(5,9)	e(0)	
1	2	3	5	5=Col.1-Col.2
57.50	4.5309	4.3362	59.253	-1.753
60.0	4.5933	4.4309	60.272	-0.272
62.5	4.6528	4.5204	61.243	1.257

Expectation of life at birth = 60.44

A life table that satisfies equation (49) has expectation of life at birth of 60.44 years in the North family of the Coale-Demeny model life table.

The Coale-Demeny Projections and Cumulation Method

106. The Coale and Demeny method makes the same assumptions as the Carrier and Hobcraft method. The difference is in the extent to which assumption three is utilized to arrive at the level of mortality. In theory, Carrier and Hobcraft method can do without it. In the case of Coale and Demeny method, it is impossible to apply the technique where a one parameter system of model life table is not avialble.

107. The application of the Coale and Demeny method starts by first selecting one of the many "families" of one-parameter system as the pattern of mortality prevailing in the area. Once this assumption is accepted, there can only be one life table with a survival ratio that corresponds to the observed survival ratio for a given age group. In practice, the ten year cohort survival ratios are unrealistically too high in some ages and too low in others. To reduce the effect of age misstatement errors, Coale and Demeny suggest a search for a level of mortality among one of the four regions of model life tables that will satisfy the equation.

$$\sum_{x=y}^{w} N1(x,x+4).S(x,x+4) = \sum_{x=10+y}^{w} N2(x,x+4) \qquad (50)$$

where N1(x,x+4) and N2(x,x+4) are observed populations in the first and second censuses in age groups x to x+4; S(i) is a ten year survival ratio from age (x,x+4) to (x+10,x+14) as defined earlier. In other words a level of mortality is selected so that the survived population aged y and above in the first census matches the observed populations aged y+10 and above in the second census. For each age y in the first census, an S(y,y+4) can be found corresponding to a level of mortality that will satisfy equation (50). In the absence of errors in the data or same pattern of errors in both censuses, the set of implied levels by each age y are supposed to converge to the

same level, but in practice it is not the case because of errors among other things. A choice is therefore made as a mean or a median among the estimated values in a range in which age reporting is considered to be less severe.

108. The same technique can also be used by comparing the survived population of the second census aged x+10 years and over with the population aged x years and over in the first censes i.e. If S(x,x+4) is the ten year survival ratio from age x,x+4 to x+10,x+14 in an intercensal period,

$$\sum_{x=y}^{w} N1(x,x+4) = \sum_{x=y}^{w} N2(x+10,x+14)/S(x,x+4)$$

Once again the median level is selected among the first set of seven or nine values. The first procedure is called the forward method and the latter the backward method. In theory both are expected to give the same levels of mortality, but in practice they do not. The reasons for giving different results are likely to be one or both of the following,

i) differences in completeness in the two censuses
ii) differences in the pattern of age misreporting in both censuses

One advantage of the Projection and cumulation technique is its process of cumulation which drastically reduces the effect of age misstatement except for substantial transfers across the initial age used in the cumulation.

<u>Illustration of Coale-Demeny Projection and Cumulation Method</u>

<u>Step 1</u>: Select the pattern of mortality to be used. For Kenya, North Model is selected as the most appropriate model. Once the pattern is selected, the range of plausible estimates are made. In the case of Kenya, the level is not expected to be less than level 10 and not more than level 18.

<u>Step 2:</u> The first population is projected successively by levels, 10,12,14,16,18 after ten year survival ratios for these levels have been calculated. Where computer facilities and the required softwares are available, the projections can be made for single levels of mortality. The ten year survival ratio can be estimated by multiplying consective five year survival ratios. For example, ten year survival ratio for 0-4 can be calculated as:

$$\frac{L(10,14)}{L(0,4)} = \frac{L(5,9)}{L(0,4)} \cdot \frac{L(10,14)}{L(5,9)} = S(1).S(2).$$

Similarly, ten year survival ratios can be estimated for each of these levels and the first census projected forward by these survival ratios for each level.

<u>Step 3</u>: The projected populations are cumulated from the oldest to the younger age groups such that for each level, NP(10+), NP(15+), NP(20+), ..., NP(45+) are computed. NP stands for projected population.

<u>Step 4</u>: The reported population in the second census is also cumulated in the same pattern as in step 3. For each age 10 years and over, the projected cumulated population is bracketed by two numbers for two levels. By interpolation, the level/corresponding to the cumulated population can be determined. Repeat the process for age ranges 10+, 15+, 20+, 25+, ..., 45+ and select the median. For Kenya the illustration is given in Table 3.14.

109. The problems associated with this technique are of two kinds. First, it is difficult to apply the method when the interval between censuses is not a multiple of 5 years. Infact for multiples of five years other than the period of 10 years it will be very difficult or perhaps time consuming to generate the survival ratios before the method can be applied. With the advent of computers and development of softwares though the technique can be handled very easily.

110. The other problem is its dependence on reference tables. In situations where existing tables do not resemble the adult mortality pattern of the population under investigation the technique can not be used. Fortunately, with the addition of the New United Nations (1983) and the OECD model life table systems (Clairin, et al, 1980) to the Coale-Demeny systems, the lack of the flexibility of the technique is reduced.

111. As with the Carrier and Hobcraft method, the method is sensitive to differential coverage between the censuses and net migration. Adjustments will have to be made before the application of the technique for better esimates.

<u>The Preston - Bennette Method</u>

112. Preston and Coale (1982) have shown that for any population,

$$C(x) = b.EXP[-\int_0^x r(a)da].l(x) \qquad (51)$$

where C(x) is the proportion at age x; b is the intrinsic birth rate, r(x) the growth rate at age x and p(x) is the survivors to exact age x. This general equation can be used to estimate mortaliy levels p(x), of any country with two censuses. One big advantage of the technique is that it can be applied to any intercensal duration unlike the previous techniques which can only be used for multiples of five year intervals. A second

Table 3.14 An Illustration of the Coale-Demeny Forward Projection and Cumulation Method Kenya 1969 and 1979

Age Group	1969 Female Population	1979 Cumulated Population	PROJECTED AND CUMULATED POPULATION BY VARIOUS LEVELS OF MORTALITY FROM NORTH FAMILY					Mortality Levels
			10	12	14	16	18	
0-4	1046380							
5-8	893359							
10-14	663808	5049027	4758927	4860549	4950945	5034614	5110722	16.19
15-19	544847	4023350	3859239	3931541	3992691	4051195	4106490	16.19
20-24	450096	3134034	3027745	3088121	3137961	3186433	3233133	14.52
25-29	411245	2446800	2404330	2451490	2500951	2543343	2584612	12.92
30-34	299241	1904567	1897012	1943754	1981712	2018715	2055064	10.80
35-39	264819	1491135	1482735	1523387	1556142	1588081	1619763	10.16
40-44	201936	1165184	1109172	1143355	1170592	1197209	1223977	10.21
45-49	163852	890991	840935	869891	892720	915094	937903	12.80
50-54	139072	668628	606364	630476	64933	667867	687011	12.92
55-59	102235	477263	430353	450616	466407	481971	687011	16.04
60-64	94508	342488	292157	308876	321881	334768	348334	
65-69	63307	232773	181922	194876	204941	215024	225750	
70+	12169	149403	184915	118528	125892	133380	141435	

advantage of the technique is that it does not require the use of model life tables and like the others it avoids the assumption of stability. However, the population is assumed to be closed to migration or adjustment will have to be made before its application if the relevant data are available.
From the above equation,

$$p(x) = (N(x)/N(0))). \text{EXP}(\int_0^x r(a)\, da)$$

$$\text{i.e.} \quad e(0) = \int_0^w p(x)dx = \int_0^w (N(x)/N(0). \text{EXP}(\int_0^x r(a)da)dx \quad (52)$$

The above equation requires an estimate of N(0), the number of births in the intercensal period, in order to estimate the expectation of life at birth. Since the data will not usually be available, another age u can be chosen such that

$$e(u) = \int_u^w p(x)dx = \int_u^w (N(x)/N(0).\text{EXP}(\int_u^x r(a)da)dx$$

For discrete data,

$$e(u) = \frac{\sum_{i=u}^{w} N((i,i+4) \text{ EXP}[5 \sum_{x=u}^{i-5} r(x,x+4)+2.5(r(i,i+4))]}{[N(u-5,\ u-1) + N(u,u+4)]/10} \quad (53)$$

112. The application of the technique to Kenyan female age distribution given in 1969 and 1979 is given in Table 3.16. The steps to be followed are given below:

Illustration of Preston-Bennett Method

Step 1: The intercensal growth rates for age group x to x+4, r(x,x+4), are calculated using the exponential growth law ie.,

$$r(x,x+4) = [\ln\ (N2(x,x+4)/N1(x,x+4)]/n \quad (54)$$

where N2(x,x+4)and N1(x,x+4) are respectively the number of females in age group x to x+4 in 1979 and 1969 censuses respectivelly. n is the intercensal interval in years and ln stands for natural logarithm. In Table 3.16, N1(x,x+4) and N2(x,x+4) are given in columns 2 and 3.

For example,

$$r(5,9) = \ln[(1246983/893359)]/10=[\ln(1.39584)]/10=0.33350/10$$
$$= 0.03335$$

The results of applying the exponential growth law formula are shown in column 4.

Step 2: The mid-period population for which the estimated level of mortality refers is given by:

$$N(x,x+4) = 0.5[N2(x,x+4)+N1(x,x+4)] \qquad (55)$$

where N1(x,x+4) and N2(x,x+4) are as defined in step 1.

For x = 10,

$$N(10,14) = 0.5\ (1025677+663808) = 844742$$

Similarly, the average mid-period population for the rest of the age groups is estimated using the same formula and the results are displayed in column 5.

Step 3: The intercensal rates of growth are cumulated from age 5 to the mid-point of each successive higher age group using the equation

$$R(x+2.5) = \sum_{i=5}^{x-5} r(i,i+4) + 2.5\ r\ (x,x+4) \qquad (56)$$

for x=5,

$$R(7.5) = 2.5 \times 0.03335 = 0.0834$$

for x = 10,

$$R(12.5) = 5 \times r(5,9) + 2.5r\ (10,14)$$
$$= 5 \times 0.03335 + 2.5 \times 0.04351 = 0.2755$$

For x = 25,

$$R(27.5) = 5 \sum_{i=5}^{20} r(i,i+4) + 2.5r(25,29)$$
$$=5[r(5,9) + r(10,14) + r(15,19) + r(20,24)]+2.5r(25,29)$$
$$=5(0.03335+0.04351+0.04899+0.04232)+2.5 \times 0.02765$$
$$= 0.9000$$

Similarly, the cumulated intercensal rates of growth are computed and entered in column 6.

For the last open ended age interval,

$$R(A) = \rho(A) + 5 \sum_{i=5}^{A-5} r(i,i+4) \qquad (57)$$

where $\rho(w)$ is calculated by using coefficients given in Table 3.15. The equation used to estimate $\rho(w)$ is:

$$\rho(A) = a(A)+b(A).r(10+) + c(A)\ \ln(N(45+)/N(10+)) \qquad (58)$$

Table 3.15: Coefficients for Estimation of the Equivalent Growth Rate over Age A, P(A), From the Growth Rate over Age 10 and the Ratio of the Population Over Age 45 to the Population Over Age 10

Age A	a(A)	b(A)	c(A)
45	0.229	20.43	0.258
50	0.205	18.28	0.235
55	0.179	16.02	0.207
60	0.150	13.66	0.176
65	0.119	11.22	0.141
70	0.086	8.77	0.102
75	0.053	6.40	0.063
80	0.025	4.30	0.006
85	0.006	2.68	0.006

Source: United Nations. (Manual X) Indirect Techniques for Demographic Estimation. Population Studies of the United Nations No. 81. ST/ESA/SER.A/81. P. 219.

where a(A), b(A) and c(A) are given in Table 3.15 for A = 45, 50,..,80, and 85. r(10+) is the intercensal rate of growth of those 10 years and over in the first census. N(10+) and N(45+) are the mid-period populations 10 years and over and 45 years and over respectively. They are estimated by cumulating the mid-period population 10 years and over and 45 years and over respectively, ie.,.

$$N(10+) = N(10,14) + N(15,19) + \cdots + N(70,74) + N(75+)$$
$$=844742+717084+\cdots +54319+81192 =4284802$$
$$N(45+) = N(45,49) + N(50,54) + \cdots + N(70,74) + N(75+)$$
$$=193107 + 165218 + \cdots + 54319+81192=787790$$

Step 4: The average mid-period population is then converted into a stationary population using:

L(x,x+4) = N(x,x+4) Exp (R(x)) for five year age groups and N(A+) = Exp(R(A+) for the open age interval (59)

for x = 25,

$$L(25,29) = N(25,29)\ Exp(R(25))$$
$$= 476739\ Exp\ (0.9101) = 1184373$$

For the open ended age interval where w=75+ and N(75+) = 81192, equation (58) gives,

$$R(75+)= a(75+)+b(75+).r(10+)+c(75+).\ln(N(45+)/N(10+)+ 5\sum_{x=5}^{70} r(i,i+4).$$

From Table 3.15, a(75) = 0.053, b(75) = 6.4 and c(75) = 0.063. Application of equation (54) to the population 10 years and over gives an intercensal rate of growth of r(10+) = 0.03606

Hence,

$$R(75+)=0.053+(6.4x0.03606)+0.063x\ln(787790/4284802)+5\sum_{i=5}^{70} r(i,i+4)$$

$$= 0.17707 + 5\sum_{i=5}^{70} r(i,i+4)$$

$$=0.17707+2.21511 = 2.39218$$

.·. L(75+) = T(75) = 81192 xExp (2.39218) = 888018

Column 8 shows the stationary population estimated using equation (59)

Step 5: The numbers of survisors to exact age x, l(x), are estimated by the relation:

$$l(x) = (L(x-5,x-1) + L(x,x+4)/10 \qquad (60)$$

For example, for x = 15
L(10,14) = 1112756 (from col. 8)
L(15,19) = 1190406 (" " ")

.·. l(15) = (1112756+1190406)/10 = 230316

The l(x) values for other exact ages are given in column 7.

Step 6: If T(x) denotes the person-years lived above age x, its value can be calculated as:

$$T(x) = \sum_{i=x}^{A-5} L(i,i+4) + T(A)$$ where L(x,x+4) and T(A)

are as defined in step 4.

For x = 65,
T(65) = L(65,69) +L(70,74) + T(75)
= 537414 + 460681 + 888081
= 1886176

The T(x) values for x > 5 are given in column 9.

Step 7: The expectation of life at age x, for $10 \leq x \leq 70$ are given in column 10 and were computed using the formula

$$e(x) = T(x)/l(x) = \text{col. 9/col. 7}$$

Note that, the computation should not necessarily begin with age 10, it can start from age five or any other higher age. Age ten is prefered for this exercise, because the 0-4 age group are considered to sufer from omission as well as shifting of those aged 3 and 4 across age 5.

Step 8: The implied levels of mortality, expectations of life at birth and age 5 are also given in columns 11, 12 and 13 respectively for Coale and Demeny (1966) North Model life tables.

113. From column 11, it can be seen that the implied mortality level decreases from a high level of 14.68 for age 10 to 10.13 for age 25 and then increases gradually with increasing age up to almost level 22 for age 65. The pattern of implied levels depicted by the Kenyan data is not unique. Differential completeness and age mistatement error are likely to create such a pattern. It is suggested that an average of the first three levels should be selected as an estimate of the true level of mortality. In the case of Kenya, the estimated level of mortality will be (14.68+13.53+11.21)/3 = 13.1. A level of mortality of 13.1 for Kenyan females for the period 1969 to 1979 appears to be on the lower side.

Mortality Estimation Based on Orphanhood

114. The possibility of using data on survival of parents to estimate female and male adult survivorship was suggested by Henry (1960). Later Brass and Hill (1973) developed a convenient procedure for obtaining estimates of survivorship from some fixed age from proportions of respondents with surviving parents. Since it is found that data on maternal orphanhood is more reliable than those on paternal orphanhood, it is only the former which will be considered here. The main idea is the following.

115. Consider a group of children a years old at the time of the survey. The mothers of these children must have been alive a years before the census or survey i.e. at the time of birth of these children. If assumptions are made that the maternal mortality of the living and dead children are the same and that the surviving children are typical of all the births a years earlier in terms of maternal mortality, the proportion of the respondents with a surviving mother will represent the probability of those mothers surviving for a years.

116. The mathematical relationship between proportion of children with surviving mothers and mortality of parents is:

Therefore

$$P(a) = \frac{\int_{15}^{49} b(x) \cdot \frac{l(a+x)}{l(x)} \, dx}{\int_{15}^{49} b(x)\,dx} \qquad (61)$$

Where b(x) = the density distribution of mothers by age x
l(x)= the survivor function of a life table from birth to age x
P(a) = the proportion of children of exact age a whose mothers were alive at the time of the census/survey.

Table 3.16 An Illustration of the Preston and Bennette Method of Estimating Adult Mortality for Kenya 1969-1979

	Enumerated Female Population		Intercensor Growth Rate	Average Population	Cumulated Growth Rate	Stationary Population			Expectation of Life at Age x	Corresponding		
						At exact age x	In five year age group	Over Age x		Level	Expectation of Life at Age Zero e(0)	Expectation of Life at Age Five e(5)
	N1(x,x+4)	N2(x,x+4)	r(x,x+4)	N(x)	R(x+2.5)	l(x)	L(x,x+4)	T(x)	e(x)			
1	2	3	4	5	6	7	8	9	10	11	12	13
0-4	1046380	1423936	0.03081	1235158	0.0000				0.00	0.00	0.00	0.00
5-9	893359	1246983	0.03335	1070171	.0834		1163233		0.00	0.00	0.00	0.00
10-14	663808	1025677	0.04351	844742	.2756	227598	1112756	12516476	54.99	14.68	54.20	58.37
15-19	544847	889316	0.04800	717081	.5069	230316	1190406	11403720	49.51	13.53	51.33	56.62
20-24	450096	687234	0.04232	568665	.7352	237656	1186156	10213313	42.98	11.21	45.54	53.06
25-29	411245	542233	0.02765	476739	.9101	237069	1184535	9027156	38.08	10.13	42.83	51.32
30-34	299241	413432	0.03232	356336	1.0601	221315	1028616	7842621	35.44	11.00	45.00	52.72
35-39	264819	325951	0.02072	295385	1.1928	200234	973724	6814004	34.03	13.55	51.37	56.65
40-44	201936	274193	0.03059	238064	1.3213	186602	892301	5840280	31.30	15.05	55.11	58.92
45-49	163852	222363	0.03053	193107	1.4741	173561	843312	4947979	28.51	16.71	59.28	61.30
50-54	139072	191365	0.03192	165218	1.6302	168677	843462	4104667	24.33	16.47	58.67	61.12
55-59	102235	134775	0.02763	118505	1.7791	154558	702118	3261204	21.10	17.64	61.61	62.95
60-64	94508	109715	0.01492	102111	1.8855	137502	672909	2559086	18.61	20.26	68.14	67.07
65-69	63307	83370	0.02753	73338	1.9917	121032	537414	1886176	15.58	21.86	72.15	69.62
70-74	45987	62651	0.03092	54319	2.1378	99809	460681	1348762	13.51	99.00	99.00	99.00
75+	75632	86752	0.01372	81192	2.3922	0	888081	888081	0.00	0.00	0.00	0.00

117. The right hand side of equation (61) is a ratio of two quantities. The denominator is the number of mothers involved and the numerator a product of the mothers and their conditional probability of survival from age y at the time of birth of the child to age a+x, the age of mothers at the time of the census/or survey. The integral covers the reproductive age range 15-49. In general, it could have been stated from u to v where u and v are the start and end of childbearing in the society respectively.

118. Equation (61) above can be expressed further in the form of l(A+B)/l(B) where A relates to the age of the respondent and B to the average age of mothers of children born in a particular time period. In fact, using models, Brass has developed Weighting factors such that the proportion of those with surviving mothers can be converted into conditional probabilities starting from age 25 to age 25+N where N is the age of children. The converting equation is

$$\frac{l(25+N)}{l(25)} = W(N)\ P(N-5,N-1) + [(1-W(N)]P(N,N+4)) \qquad (62)$$

where l(25+N)/l(25) is the conditional probability, W(N) is a weighting factor dependent on age of child and mean age of mothers at birth of children, $\bar{M}$, and P(N,N+4) is the proportion of children with surviving mothers in age group N to N+4. N is the central age of two consecutive age groups. For example, for respondents 5-9 and 10-14 age group, the central age N is equal to 10 years. The weighting factors ,W(N), for converting proportion of surviving mothers into life table survival probabilities from age 25 years are given in Table 3.17.

119. The mean age difference between mothers and children at the time of birth of the child is calculated from the distribution of births in a year by age of mother. If the number of children born in the last 12 months is given as B(i) for age group i, the mean age difference between mothers and children at the time of birth is given as:

$$\bar{M} = \sum_{i=1}^{7} B(i).X(i) / \sum_{i=1}^{7} B(i) \qquad (63)$$

where X(i) is the mid-point of the age interval i. The index i takes values 1, 2,.., 7 for age groups 15-19, 20-24, .., 45-49. For M estimated from data on births in the last 12 months, half of a year is subtracted because the mothers are on the average half of a year younger at the time of birth of the child than at the time of the census or survey.

Table 3.17 Weighting Factors, W(N), For Conversion of Proportions of Respondents With Mother Alive into Survivorship Probabilities for Females

Age	Mean age, M, for mothers at Maternity								
N	22	23	24	25	26	27	28	29	30
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
10....	0.420	0.470	0.517	0.557	0.596	0.634	0.674	0.717	0.758
15....	0.418	0.489	0.556	0.618	0.678	0.738	0.800	0.863	0.924
20....	0.404	0.500	0.590	0.673	0.756	0.838	0.921	1.004	1.085
25....	0.366	0.485	0.598	0.704	0.809	0.913	1.016	1.118	1.218
30....	0.303	0.445	0.580	0.708	0.834	0.957	1.080	1.203	1.323
35....	0.241	0.401	0.554	0.701	0.844	0.986	1.128	1.270	1.412
40....	0.125	0.299	0.467	0.630	0.791	0.950	1.111	1.274	1.442
45....	0.007	0.186	0.361	0.535	0.708	0.884	1.063	1.250	1.447
50....	-0.190	0.017	0.158	0.334	0.514	0.699	0.890	1.095	1.318
55....	-0.368	0.220	0.059	0.101	0.270	0.456	0.645	0.856	1.083
60....	0.466	0.352	0.217	0.084	0.053	0.220	0.378	0.579	0.800

Estimation equation:

l(25+N)/l(25)=W(N).P(N-5,N-1)+(1-W(N).P(N,N+4)

Source: United Nations. (Manual X). Indirect Techniques for Demographic Estimation. Population Studies of the United Nations, No.81. ST/ESA/SER.A/81. New York. p.103.

120. One advantage of equation (62), is its smoothing characteristic. Once the conditional probabilities are estimated, an abridged life table begining at age 25 can be estimated. Otherwise, using models, the conditional probabilities can be converted into levels of mortality or expectations of life at birth or any other age. In addition, a complete life table can be generated by splicing together results from childhood mortality from children dead among ever born and adult mortality from survival of mother.

121. Another method of converting proportions of persons with parent alive into standard indices of mortality has been devised by the United Nations (1983). But since experience suggests that the Brass-Hill method gives better results for ages of children less than 30, the application of the technique will be demonstrated for the former and not the latter.

122. The orphanhood method suffers from some problems. The first problem relates to the young ages. Invariably, the orphanhood data for children under 15 implies unrealistically very low mortality levels which are usually explained by what is called the "adoption effect" ie. the failure to differentiate between foster parents and true parents. In addition there is another problem with the data. While mothers with large number of surviving children are reported as many times as there are surviving children, the childless and those with few surviving children are not well represented. Another common problem is that male and female children reporting about their parents give different proportions of survivng parents. The proportion orphan are higher for females than for males of the same age. Such a differential could be explained if the ages of males are exaggerated relative to the ages of females, or if the sex ratio at birth declines with age of mother. Usually it is prefered to use data for both sexes combined when estimating the mortality level.

123. To reduce the effect of multiple reporting due to many surviving children by some women, suggestions were made to address the questions to the first-born or the eldest surviving child. Unfortunately the results were not successful. The plausible reason is the confusion as to who is the first-born or the eldest surviving child especially in the African context where once again polygyny is substantial and child birth not restricted within marriage.

124. The problem of timing was tackled by Brass and Bamgboye (1981) . Their procedure helps to estimate the time period (t) to which each estimate of l(B+N)/l(B) would refer prior to the censuses or survey on the assumption that mortality has been changing in a regular fashion. In addition to the linear decline in mortality change, an assumption is also made that the pattern of change conforms to a one parameter logit life-table system. In practise, it is evident that the ts are sensitive to changes in the age pattern of mortality but that they are much less affected if the other assumptions are violated. To appreciate the trends in mortality, it will be

necessary to relate each estimate of l(B+N)/l(B) to an equivalent standard index of mortality such as the implied level of mortality, the expectation of life at age five, the chance of surviving to age sixty ie., l((60)/l(B), etc.

An Illustrative Example of Orphanhood Method to Kenya 1979 Census Data

125. The data requirements are:

(1) the total population by age (2) Those with mother alive by age and (3) number of women by age group in the reproductive ages. In addition the total number of births in the 12 months preceeding the survey by age of women is also needed. Let N(x,x+4) and MS(x,x+4) be the total number of respondents and those with surviving mothers respectively in the same age group.

Step 1: The proportion with surviving mothers, P(x,x+4), is calculated using the formula

$$P(x,x+4) = MS(x,x+4)/N(x,x+4)$$

In the case of Kenya, the proportion with surviving mothers is given in column 2 of Table 3.18.

Step 2: For the 12 months preceeding the survey, the number of women who had births by age was given as

Age Group x-x+4	15-19	20-24	25-29	30-34	35-39	40-44	45-49	Total
No. of Women W(x,x+4)	87343	181078	149275	93286	56204	24836	9234	601256

If we assume that each woman had only one birth in the 12 months preceeding the census, the W(x,x+4) will equal the B(x,x+4) defined above. The mean age difference between mothers and children at the time of birth is thus:

$$\bar{M}=[(17.5x87343+22.5x181078+.....47.5x9234)/601256]-0.5 = 26.7$$

Step 3: Using $\bar{M}$ = 26.7, weighting factors (W(N)), were computed from Table_3.17 developed for this purpose by_Brass. Since the computed $\bar{M}$ = 26.7 lies between table values $\bar{M}$ = 26 and $\bar{M}$ = 27, linear interpolation with weights 0.3 and 0.7 resepctively will give the required multiplier. For instance: For central age N = 20, W(20) = 0.3x0.756+0.7x0.838 = 0.813
Similarly, the weighting factors for other central ages were computed and the results entered in column 4.

Step 4: The conditional probabilities are computed using equation (62). For example, For N = 15,

$$l(25+15)/l(25) = l(40)/l(25) = 0.720 \times 0.974 + 0.280 \times 0.950 = \underline{0.967}$$

Column 6 shows the conditional probabilities for N=10 to N=50 which are similarly computed. The corresponding levels in the Coale-Demeny North Family are given in column 7.

126. It can be seen that the implied levels decline from a level of 21.3 for the 5-9 to 14.4 for the 45-49. The impression given is that adult mortality has been on a decline. Whether this implied trend is a fact or an artefact is difficult to say. What can be said for certain is that the level of mortality in Kenya as implied by children under age 25 appears to be too light to be believed. The lighter mortality of the younger age group children is likely to be due to misconceptions arrising from failure to distinguish between foster and true biological mother.

Table 3.18: An Application of the Orphanhood Method of Estimating Female Adult Mortality Using 1979 Census Data of Kenya

Age Group x-x+4	Proportion with Mother Alive P(x,x+4)	Central Age N	Weighting Factors W(N)	Weighting Factors [1-W(N)]	Conditional Probabilities l(N+25)/l(25)	Corresponding Levels on the North Model
1	2	3	4	5	6	7
5-9	0.985	10	0.623	0.377	0.981	21.3
10-14	0.974	15	0.720	0.280	0.967	20.9
15-19	0.950	20	0.813	0.187	0.941	20.0
20-24	0.904	25	0.882	0.118	0.898	18.6
25-29	0.852	30	0.920	0.080	0.845	17.6
30-34	0.759	35	0.943	0.057	0.754	15.7
35-39	0.678	40	0.902	0.098	0.666	15.3
40-44	0.555	45	0.831	0.169	0.536	14.5
45-49	0.445	50	0.644	0.356	0.395	14.4
50-54	0.305	55				

References

Arriaga, Eduardo E.
New Life Tables for Latin American Populations in the Nineteenth and Twentieth Centuries. Population Monograph series No. 3. Berkeley: Innstitute of International Studies, University of California, 1968.

El Badry, M.A.,
"Failure of Enumerators to Make Entry of Zero: Error in Recording Childless Cases in Population Censuses," Journal of the American Statistical Association, Vol.56, No.26, 1961.

Bennett Neil G. and Shiro Horiuchi.
"Estimating the Completeness of Death Registration in a Closed Population, Population Index 47(2):207-221, 1981.

Blacker, J. and Allan Hill.
"Age Patterns of Mortality in Africa: An Examination of Recent Evidence" Proceedings of the IUSSP Conference, Florence. 2: 1985.

Brass, W.
Uses of Census and Survey Data for Estimation of Vital Rates", African Seminar on Vital Statistics, United Nations Economic Commission for Africa, Addis Ababa, 1964.

Brass, W. and K.H. Hill
"Estimating Adult Mortality From Orphanhood," Proceedings of the International Population Conference, Liege, 1973 Vol.3, International Union for the Scientific Study of Population Liege, 1973.

Brass, W. and E.A.Bamgboye
"The Time Location of Reports of Survivorship: Estimates for Maternal and Paternal Orphanhood and the Ever-Widowed." Working Paper No. 81-1, Centre for Population Studies, London School of Hygiene and Tropical Medicine, London, 1981.

Brass, W.
Methods for Estimating Fertility and Mortality from Limited and Defective Data, Laboratories for Population Statistics, Chapel Hill, N.C., 1975.

Brass, W.
"The Assessment of the Validity of Fertility Trend Estimates from Maternity Histories", Proceedings of the IUSSP Conference, Mexico City, Vol.1. 1977.

Brass, W.
Methods for Estimating Fertility and Mortality from Limited and Defective Data Laboratories for Population Statistics, Chapel Hill, N.C., 1975.

Brass, W. and K.H. Hill.
"Estimating Adult Mortality from Orphanhood,' in Proceedings of the International Population Conference, Liege, 1973, Vol.3, IUSSP, Liege, 1973.

Carrier, N.H. and Hobcraft, J.
Demographic Estimation for Developing Societies, Population Investigation Committee, London, 1971.

Clairin R, Julien Conde, Michele Fleury-Brousse, Dominique Waltisperges and Guillaume Wunsch
Mortality in Developing Countries, TOME 111, Vol.V/Vol.VI, New Model Life Tables for Use in Developing Countries. Development Centre of the Organization for Economic Co-operation and Development, Paris, 1980.

Coale, A.J. and Hoover, E.M.
Population Growth and Economic Devevelopment in Low Income Countries, Princeton University Press, New Jersey, 1958.

Coale, A.J. and J. Trussell.
"Model Fertility Schedules: Variations in the Age Structure of Child bearing in human Populations." Population Index, Vol.40,No.2, 1974.

Coale, A.J. and E.M. Hoover.
Population Growth and Economic Development in Low-Income Countries (A case Study of India's Prospects). Princeton University Press. Princeton, U.S.A. 1958.

Coale, A.J. and T.J. Trussell,
"Estimating the Time of Which Brass Estimates Apply," Population Bulletin of the United Nations, No.10, (United Nations Publication, Sales No. E.78 XIII.6), 1978.

Coale, A.J. and P. Demeny
Regional Stable Populations, Princeton University Press, Princeton , New Jersey, 1966.

Courbage, Y. and P. Fargues
"A Method for Deriving Mortality Estimates from Incomplete Vital Statistics," Population Studies, Vol.33, No.1, 1979.

Ekanem, Ita I. and R.K. Som,
The Problem of Choosing Model Life Tables for African Countries. Genn 11(1-2), 1984.

Feeney, G.,
"Estimating Infant Mortality Trends from Child Survivorship Data," Population Studies, Vol.34, No.1, 1980.

Henry, Louis.
Mesure indirecte de la mortalite des adultes. Population, V.15, No.3, 1960.

Hill, K.
"An Evaluation of Indirect Methods for Estimating Mortality" in Methodologies for the Collection and Analysis of Mortality Data by Jacques Vallin, John H. Pollard and Larry Heligman (eds), Ordina Editions, Liege (Belgium), 1984

Kraly, E.P. and D.A. Norris,
"An Evaluation of Brass Mortality Estimation Under Conditions of Declining Mortality." Demography. 15(4) 1978.

Lotka, A.J. and F.R. Sharpe,
"A Problem in Age Distribution" Philosophical Magazine, 1911.

Mabuza, S.E.
"Some Aspects of Mortality in Swaziland" Unpublished M.A. Thesis, held by the University of Ghana Legon. 1981

Palloni, A., and L. Hleigman,
"Re-Estimation of Structural Parameters to Obtain Estimate of Mortality in Developing Countries," Population Bulletin of the United Nations No.18, 1985. ST/ESA/Ser.N/18, 1986.

Palloni, A.
"A Review of Infant Mortality in the Third World: Some New Estimates." Population Research Centre. University of Texas. U.S.A. (Mimeographed), 1979.

Palloni, A.
"Estimating Infant and Child Mortalaity under Conditions of Changing Mortality". Population Studies. 34(1), 1980.

Potter, J.E.
"Problems in Using Birth History Analysis to Estimate Trends in Fertility," Population Studies 34(2), 1977.

Preston, S.H. and Neil Bennett.
"A Census Based Method for Estimating Adult Mortality," Population Studies 37(1). 1982

Preston, S.H. et al;
"Estimating the Completeness of Reporting of Adult Deaths in Populations that are Approximately Stable," Population Index, Vol.56 No.2. (Summer 1980).

Preston, S.H. and A.J. Coale.
"Age Structure, Growth, Attrition and Accession: A new synthesis." Population Index. 48(2), 1982

Preston, S.H. and K. Hill,
"Estimating the Completeness of Death Registration," Population Studies, Vol.34, No.2, 1980.

Preston, S.H. and A. Palloni,
Five-tuning Brass-type Mortality Estimates with Data on Ages of Surviving Children." Population Bulletin of the United Nations, No.10, 1977 (United Nations Publication, Sales No.E.78 XIII.6), 1978.

Preston, S.H.
"Use of Direct and Indirect Techniques for Estimating the Completeness of Death Registration Systems" in Data Bases for Mortality Measurement, United Nations Population Studies No.84, ST/ESA/SER.A/84. New York, 1984.

Ramachandran, K.V.
"Evaluation of the 1973 Census of Libya". Workshop on Methods of Demographic Data Evaluation, Adjustment and Analysis, Regional Institute for Population Studies (Legon, Ghana) 1977.

Smith, D.P.
"Life Table Analysis, World Fertility Survey" Technical Bulletins, No.6/Tech. 1365, 1980.

Stolnitz, G.J.
Life Tables from Limited Data: A Demographic Approach, Princeton. University Press, Princeton, U.S.A.

Sullivan, J.M.
"Models for the Estimation of the Probability of Dying Between Birth and Exact Ages of Early Childhood", Population Studies, Vo.26, No.1, 1972.

Sullivan,J. and G. Udofia
"Interpreting Childhood Survivorship Statistics under Conditions of Changing Mortality. Population Studies, Vol. 33, No. 2, 1979.

Tesfay Teklu,
"Some Problems in the use of Model Life Tables to African Data" in Fertility and Mortality Estimation in Africa, RIPS Monograph No.4. (Forthcoming).

Tesfay Teklu and K. Venkatacharya,
"Estimation of Infant and Child Mortality in Africa: Recent Experiences," (Mimeograph). A Paper Presented at the World Fertility Survey Regional Review Seminar, held in Addis Ababa 17-21 March 1986.

Trussell, T.J.
"A Re-Estimation of the Multiplying Factors for The Brass Techniques for Determining Childhood Survivorship Rates," Population Studies, Vol. 29, No.1, 1975.

United Nations
Methods of Estimating Basic Demographic Measures from Incomplete Data, 1967

United Nations
(MANUAL IV) Methods of Estimating Basic Demographic Measures from Incomplete Data (United Nations Publication, Sales No. E.67. XIII.2), 1967.

United Nations.
Model Life Tables for De veloping Countries Sales No. E.81. XIII.7., New York, 1982.

Venkatacharya, K.
Notes on Effect of Past Changes in Fertility on Brass Estimates of Mortality Using Number of Children Ever Born and Surviving (Mimeographed), 1979.

Venkatacharya, K.
"Indirect Estimation of Infant and Child Mortality using Child Survivorship Data: Current Methods and Research Issues" in Fertility and Mortality Estimation in Africa, RIPS Monograph No.4 (Forthcoming).

Zltonik, H. and K.H. Hill.
"The Use of Hypothetical Cohorts in Estimating Demographic Parameters Under Conditions of Changing Fertility and Mortality", Demography, Vol.18, No.1, 1981.

CHAPTER 4

ESTIMATION OF FERTILITY

4.1 Introduction

1. Estimation of current level and future trends in fertility forms an important element in demographic studies. Better estimation of levels, differentials and trends in fertility in sub-Saharan Africa can be made with the information available now than ten or fifteen years ago.

2. In sub-Saharan Africa, registration of vital events is mostly non-existent and in most cases where it exists the results are not useful. For this reason methods of estimating fertility using registered births is not dealt with here.

3. This leaves censuses and single round surveys (and a few multiround surveys especially in Francophone countries) as the main source of fertility data. Many countries in Africa have at least two censuses and more countries will have at least two censuses in the future. Thus the data situation in quantitative terms is much better now than what it was some twenty years ago, but in qualitative terms there are still a number of problems, and probably many of them will be inherited into the nineties.

4. Since the sixties with the publication of 'The Demography of Tropical Africa' (Brass et al, 1967) and the United Nations Manual IV (United Nations, 1967), the indirect techniques of fertility and mortality estimation have become popular and these have filled a certain gap in the methodology of estimation of vital rates in the less developed countries. The recent United Nations Manual X (United Nations, 1983) has updated and augmented many indirect estimation techniques. The original methods have been modified or improved upon and new methods have also been developed but the type of the data collected more or less remained the same. Consequently,it appears that any further improvements in the estimation of fertility and mortality can be achieved by improvements in the quality and quantity of data than by sharpening the existing indirect techniques. Further, as many aspects of population change with time including vital rates, these methods become less effective. Pleas for better data collection and improved methods have been voiced in almost all major population conferences.

5. While direct estimates of crude birth rate and age specific fertility rates are possible from census or survey data, most often these estimates turn out to be gross underestimates to be of any use for serious work. Therefore, efforts are made to adjust these rates in a number of ways. Some use the post-enumeration survey and others use an estimate of the completeness of vital registration to deduce the correction factor. The most commonly used approach to obtain

fertility rates in the context of developing countries in general and African countries in particular is the indirect measurement of fertility.

6. The indirect techniques of estimating fertility can be broadly divided into two categories. In the first category all methods discussed use the retrospective data on the mean number of children ever born (MNCEB) and births during the last year (BLY) by age of the surveyed women. These data are used to yield adjusted estimates of current age specific fertility rates (ASFR) and the total fertility rate. The Brass P/F ratio method, which is well known, has been found to be very useful not only as a consistency check but also ,in some cases, as a method of adjusting current fertility level. Recent work of Brass using 'relational Gompertz model' (RGM) to estimate total fertility rate using the same type of data has also been tried in some African countries. These methods are discussed under group A methods.

7. In situations where data are available for two points of time seperated by approximately five or ten years the method of hypothetical cohorts developed by Zlotnik and Hill (1981) can be used to derive the parity and current fertility data for the period between the two points of time. The methods described in group A can be directly applied to the data obtained for hypothetical cohorts and estimates of recent fertility can be derived. These methods are discussed under group B.

8. In the second category, sex-age distributions of a population ,along with information on the rate of growth or a life table available at a point of time are used to estimate the crude birth rate or gross reproduction rate. Assumptions on stability are made in many cases. But methods are also available for adjusting stable estimates of fertility to take care of declines in mortality (quasi-stability), and also change in fertility. These are discussed under group C methods.

9. In cases where two or more census sex-age distributions are available, recent work of Preston (1983) using generalized stable population equations can be used. These techniques are dealt with under group D methods. Table 4.1 lists the methods under each of these groups.

10. All the methods listed in groups A to D are applied to many sub-Saharan African countries and inferences on fertility levels and patterns are made. In the recent United Nations Manual X (United Nations, 1983) the indirect techniques of fertility estimation are exhaustively and clearly described. Therefore only outlines of the selected methods are given in section 4.2 . Also a brief review of the problems encountered in applying the methods is presented. A discussion of the estimates of fertility derived using various methods for some African countries is also given.

Table 4.1 Summary of Methods of Estimating the Level of Fertility

Main Type		Method of Estimation	Data Required	Fertility Measures
A	A.1	Brass P/F Ratio Method	MNCEB and BLY at a point of time	Adjusted ASFR and TFR
	A.2	Brass Relational Gompertz Model with Pi/Pi+1	MNCEB and BLY at a point of time	Adjusted ASFR and TFR
	A.3	Brass Relational Gompertz Model with Fi/Fi+1	MNCEB and BLY at a point of time	Adjusted ASFR and TFR
B	B.1	Computation of Hypothetical Cohorts	MNCEB at two points of time	Hypothetical parity Distribution
	B.2	Zlotnik-Hill Method for Hypothetical Cohorts	MNCEB at two points of time	ASFR for inter-survey period
	B.3	Brass Relational Gompertz Model for Hypothetical Cohorts using Pi/Pi+1	MNCEB at two points of time	Adjusted inter-survey ASFR and TFR
	B.4	Brass Relational Gompertz Model for Hypothetical Cohorts using Fi/Fi+1	MNCEB and BLY at two points of time	Adjusted inter-survey ASFR and TFR

C	C.1 Reverse Survival Method for five year age groups	Sex-age distribution and a life table or rate of growth.	Crude Birth Rate
	C.2 Reverse Survival Method for cumulated population by age.	Sex-age distribution and a life table or rate of growth.	Crude Birth Rate
	C.3 U.N Manual IV Method	Sex-age distribution and a life table or rate of growth.	Crude Birth Rate
	C.4 Coale Method Using C(15) and l_5	Proportion under age 15 for both sexes, an estimate of l_5 and intercensal rate of growth.	Crude Birth Rate
D	D.1 Method of using C(15), l_5, and intercensal age specific rates of growth by Preston-Coale equation.	Sex-age distributions for two recent Censuses and estimate of l_5.	Crude Birth Rate
	D.2 Preston intergrated Method	Sex-age distributions for two recent censuses and estimate of l_5.	Crude Birth Rate and an estimate of adult mortality

4.2 METHODS OF ESTIMATING FERTILITY USING DATA ON MNCEB, BLY AND SEX- AGE DISTRIBUTIONS

4.2.1 METHODS OF ESTIMATING FERTILITY USING DATA ON MNCEB AND BLY:

11. Estimation of TFR has been achieved by various methods. Coale and Demeny (United Nations, 1967) suggested on the basis of empirical data that period TFR may be approximated by $(P_3)^2/P_2$ where P_2 and P_3 are parities for age groups 20-24 and 25-29 respectively. Brass showed that in cases where Gompertz function is appropriate , that is whenever $P_2(P_4/P_3)^4$ is less than $(P_3)^2/P_2$,the smaller quantity will be a better estimate of TFR (Brass 1980). Coale and Trussell (1974) have also developed a three parameter model for the age specific fertility which is a combination of two seperate models of nuptiality and marital fertility rates. But the fitting of this model to data is involved and is not dealt with here.

12. Brass developed a very useful device of comparing lifetime fertility to cumulative current fertility and the ratio of the two, under certain assumptions, is used as a correction factor for the current fertility estimate. This method known as the P/F ratio method will be discussed first before proceeding to the fitting of the relational Gompertz model.

A.1 - Brass P/F Ratio Method and variants

13. This method is based on the following assumptions:
(a) fertility for the population under study remained constant for sometime in the past, (b) the reported number of children ever born for women in the early ages, say, 15-35 is more or less accurately reported, and (c) the reported number of births in the last year may suffer from errors resulting from inaccurate perception by the respondents of the reference period , but these errors are invariant with age.

14. The last assumption (c) implies that the reported age specific fertility rates based on data on BLY may underestimate or overestimate the level of fertility but their age structure is correctly reported. This assumption has not been verified empirically in any study but is accepted on intuitive grounds. The failure of the P/F ratio method in the case of some Anglophone countries may be attributed to the violation of this assumption (Blacker,1979).

15. The first assumption (a) in certain cases can be verified at least indirectly. Assumption (b) is believed to be reasonable for the early ages of women.

16. Under the assumption of constant fertility the cumulated current fertility upto a certain age x, (F_x), which is deduced from data on BLY, should be equal to the lifetime fertility (MNCEB) to the same age x, (P_x). Any difference between the two can be attributed to the errors that are normally present in the data on MNCEB and BLY. Thus any deviation of P/F ratio from unity for various age groups can be taken as an indication of the presence of errors in the data.

17. What the P/F ratio method does is to use the age structure of the fertility depicted by current fertility data on BLY, which by assumption (c) is reliable, and compare it or adjust it, with the data on MNCEB by age of young women. In otherwords, the most reliable pieces of information from the current and lifetime fertility data are combined to yield an adjusted current fertility estimate.

18. As most data on MNCEB and BLY are not provided by single year ages but by five year age groups, comparison of lifetime and current fertility rates is not straight forward. For instance, MNCEB for women in the age group 20-24 refers to the mid-point 22.5 years, whereas cumulated current fertility for the age group 20-24 refers to the end of the age-group, namely 24.5 years. Further, as current fertility data are based on questions on BLY, the age of the women at the time of the occurrence of birth is on the average six months less. A set of multipliers (Table 4.2) to derive comparable parity equivalents from current fertility rates were first derived by Brass and later by Coale and Trussell (Table 4.3) (United Nations ,1983). For ages above 20 the various multipliers give almost the same result. With the help of these multipliers it is possible to obtain parity equivalents of current fertility rates (F_i) which can be compared with the lifetime fertility data (P_i). Tables 4.4 and 4.5 illustrate the application of the methods to the data of Swaziland (1976).

Table 4.2 Multiplying factors for estimating the average value over five-year age groups of cumulated fertility (Fi) according to the formula $Fi = \phi_i + k_i f_i$ (when f_i is for ages 14.5 to 19.5, 19.5 to 24.5, etc)

Exact limits of age interval								
15-20	1.120	1.310	1.615	1.950	2.305	2.640	2.925	3.170
20-25	2.555	2.690	2.780	2.840	2.890	2.925	2.960	2.985
25-30	2.925	2.960	2.985	3.010	3.035	3.055	3.075	3.095
30-35	3.055	3.075	3.095	3.120	3.140	3.165	3.190	3.215
35-40	3.165	3.190	3.215	3.245	3.285	3.325	3.375	3.435
40-45	3.325	3.375	3.435	3.510	3.610	3.740	3.915	3.150
45-50	3.640	3.895	4.150	4.395	4.630	4.840	4.985	5.000
f_1/f_2	.036	.113	.213	.330	.460	.605	.764	.939
$\bar{m}$ (years)	31.7	30.7	29.7	28.7	27.7	26.7	25.7	24.7

Source: Brass et al 1967, p. 94

Table 4.3 Coefficients for interpolation between cumulated fertility rates to estimate parity equivalents

Age group (1)	Index i (2)	Coefficients a(i) (3)	b(i) (4)	c(i) (5)
		(a) Fertility rates calculated from births in a 12-month period by age of mother at end of period		
15-19-40-44	1-6	3.392	-0.392	-
45-49......	7	0.392	2.608	-
15-19......	1	2.531	-0.188	0.0024
20-24......	2	3.321	-0.754	0.0161
25-29......	3	3.265	-0.627	0.0145
30-34......	4	3.442	-0.563	0.0029
35-39......	5	3.518	-0.763	0.0006
40-44......	6	3.862	-2.481	-0.0001
45-49......	7	3.828	0.016[a]	-0.0002
		(b) Fertility rates calculated from births by age of mother at delivery		
15-19-40-44	1-6	2.917	-0.417	-
45-49......	7	0.417	2.083	-
15-19......	1	2.147	-0.244	0.0034
20-24......	2	2.838	-0.758	0.0162
25-29......	3	2.760	-0.594	0.0133
30-34......	4	2.949	-0.566	0.0025
35-39......	5	3.029	-0.823	0.0006
40-44......	6	3.419	-2.966	-0.0001
45-49......	7	3.535	0.007[a]	-0.0002

a This coefficient should be applied to f(i-1), not f(i+1), that is, to f(6) instead of f(8).

Source: United Nations, 1983, p. 34

Table 4.4 P/F ratio method of adjusting current fertility estimates, Swaziland 1976

Brass Method

Age group	i	Reported parity (P_i) (MNCEB)	ASFR f_i	ϕ_i	K_i	F_i	P_i/F_i
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
15-19	1	.325	.117	0	2.307	.270	1.20
20-24	2	1.652	.254	.585	2.890	1.319	1.25
25-29	3	3.211	.246	1.855	3.035	2.602	1.23
30-34	4	4.667	.209	3.085	3.091	3.731	1.25
35-39	5	5.609	.159	4.130	3.210	4.640	1.20
40-44	6	6.124	.092	4.925	3.423	5.240	1.17
45-49	7	6.258	.055	5.435	4.099	5.610	1.12

a) Col. (4) is compiled from (3) by the relation $\phi_i = 5 \sum f_s$ for s = 1 to i-1.

b) Col. (5) Is interpolated from Brass Multipliers (Table 4.2) using $f_1/f_2 = 0.461$ and m = 29.4 as arguments.

c) Col. (6) is obtained from col. (5) and col. (4) by the relation $F_i = \phi_i + K_i f_i$.

Table 4.5 Coale and Trussel Method, Swaziland 1976
(UN Manual X method)

Age group	i	Reported parity (P_i) (MNCEB)	ASFR f_i	ϕ_i	F_i	P_i/F_i
	(1)	(2)	(3)	(4)	(5)	(6)
15-19	1	.325	.117	.585	.262	1.24
20-24	2	1.652	.254	1.855	1.334	1.24
25-29	3	3.211	.246	3.085	2.609	1.23
30-34	4	4.667	.209	4.130	3.731	1.25
35-39	5	5.609	.159	4.925	4.623	1.21
40-44	6	6.124	.092	5.385	5.143	1.19
45-49	7	6.258	.055	5.660	5.597	1.12

Notes: a) Col. (4) is obtained from col. (3) by the relation
$\phi_i = 5 \sum f_s$ for s = 1 to i.

b) Col. (5) is obtained by the relation
$$F_i = \phi_{i-1} + a(i)f(i) + b(i)f(i+1) + c(i)\ \phi(7)$$

where the coefficients a(i), b(i) and c(i) are obtained from Table 4.3

A.2 and A.3 - Brass Relational Gompertz Model
(Parity and cumulated fertility situations)

19. The Gompertz function has been used by many authors to represent the cumulative fertility of a population. The function is

$F(x)/F = A^{B^x}$, where F(x) is the cumulative fertility

by age x and F is the total fertility rate by the end of reproductive life. A and B are constants and lie between zero and unity. The F values can be taken as cumulative fertility derived from age specific rates or parity data. The above function can be reduced to a linear function of age by taking logarithms twice as shown below

$$Y(x) = -\ln\,[-\ln(F(x)/F)] = a+bx \qquad \ldots(1)$$

20. As it stands the model represents the fertility pattern sufficiently well but more efficient methods are deviced especially to improve the fit at the extreme ages of reproductive life. This is done by transforming the age scale. In the transformed system Y(x) values of various populations are related to one another and hence any population could be chosen as a standard against which others could be compared. As we are not sure that the Gompertz model represents the fertility pattern exactly, a model pattern is chosen such that the deviations of the real populations from linearity are minimized. Such a standard was developed by Booth(1984) based on Coale and Trussell models of fertility. If $Y_s(x)$ are standard values then we have the relational Gompertz model

$$Y(x) = a+bY_s(x) \qquad \ldots(2)$$

where a and b are constants that reflect the pattern of fertility of the particular population. The actual fitting of the above equation can only be done when the value of F is known-- which in fact is to be estimated.

21. Brass suggested an alternative fitting procedure using Z(x), instead of Y(x), which circumvents the knowledge of F. Two procedures are indicated . In one, the Gompertz parameters are estimated using parity data, whereas in the second, both lifetime and current fertility data are used. The Z values are defined as:

$$Z(i) = -\ln\,[\,-\ln(P_i/P_{i+1})] \qquad \ldots(3)$$

$$Z(x) = -\ln\,[\,-\ln(F_x/F_{x+1})] \qquad \ldots(4)$$

where i (=1,2,....,7) refers to the five year age groups and x (=20,25,...50) refers to age.

22. The estimates of a and b are obtained by fitting the approximately equivalent relations:

$$Z(i) - e_i = a' + b' g_i \qquad \ldots(5)$$

$$Z(x) - e_x = a' + b' g_x \qquad \ldots(6)$$

where Z(i) and Z(x) are as defined earlier. The values of e_i, e_x, g_i and g_x are based on the standard fertility schedule chosen and are given in Table 4.6. The application of the method for the data from Swaziland is illustrated in Tables 4.7 and 4.8.

B.1 - Intersurvey Fertility Estimation Using Hypothetical Cohorts

23. Use of one point data on lifetime fertility and cumulative current fertility gives an idea of internal consistency and in some cases enables us to adjust the current fertility level. But this can be viewed as a static situation, because the method assumes a constant fertility or a steady state. If we have 'comparable' data on MNCEB and BLY for two points seperated by five or ten years then we can obtain a hypothetical parity distribution based on the fertility experience of the women between the two dates. These two rates can then be compared as in the conventional P/F ratio method. One of the advantages of this method is that fertility between the two points of observation is used to derive the hypothetical parities and this is compared with the cumulated fertility which is also derived on the fertility experience of the same period. Hence even if fertility has changed moderately the method is expected to perform well unlike methods A.1 to A.3.

24. The method ,however, requires us to assume that the fertility experience of those women who die or migrate between the two points of time is not different from that of the survivors. The other important assumption, which is very crucial for the method is that the completeness of reporting births and accuracy of age reporting are of the same magnitude at both the points of time. In Africa it is this assumption that may not hold and results may be unsatisfactory on account of this.

25. It is not necessary that data for both the points come from surveys or censuses.For this reason the word 'survey' is used in this method in the broader sense to include census as well.

26. If we denote by P(i,1),P(i,2) and P(i,h) the parities of women in the ith age group for the first survey, second survey and hypothetical cohort respectively then we have for surveys seperated by five years:

Table 4.6 Standard values for Brass Relational Gompertz Model
Table A

Age	e(x)	g(x)
15	.9866	-2.3138
20	1.3539	-1.3753
25	1.4127	-.6748
30	1.2750	.0393
35	.9157	.9450
40	.3966	2.3489
45	-	4.8097

Table B
Standard values for cumulated fertilities and their ratios with half year shifts

Age X	$F_s(x)/F$	$Y_s(x)$	e(x)	g(x)
14 1/2	.0018	-1.8444	.9760	-2.4020
19 1/2	.1151	-.7712	1.3364	-1.4501
24 1/2	.3528	-.0410	1.4184	-.7430
29 1/2	.5869	.6294	1.2978	-.0382
34 1/2	.7795	1.3897	.9670	.8356
39 1/2	.9192	2.4736	.4509	2.1649
44 1/2	.9889	4.4984	.0462	4.4564
49 1/2	.9999	9.3416	-	-

Table C
Standard values for mean parities and their ratios

Age	i	P_i/F	$Y_s(i)$	e(i)	g(i)
10-14	0	.0004	-2.0545	1.0632	-2.6447
15-19	1	.0528	-1.0787	1.2897	-1.7438
20-24	2	.2551	-.3119	1.4252	-1.0157
25-29	3	.4956	.3538	1.3725	-.3353
30-34	4	.7064	1.0569	1.1421	.4391
35-39	5	.8678	1.9534	.7061	1.5117
40-44	6	.9676	3.4130	.2763	3.2105
45-49	7	.9977	6.0557	-	-

Source: Brass 1981, p. 360

Table 4.7 Fitting Brass Relational Gompertz model to retrospective fertility data, Swaziland 1976

Age	i	MNCEB (P_i)	$\frac{P_i}{P_{i+1}}$	Z(i)	e(i)	Z(i)-e(i)	g(i)	$Y_s(i)$	Fitted Y(i)	$\frac{P_i}{F_i}$	F_i
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
15-19	1	.325	.1967	-.4862	1.2897	-1.7759	-1.7438	-1.0787	-1.0950	.0503	6.5
20-24	2	1.652	.5145	.4086	1.4252	-1.0166	-1.0157	-0.3119	-0.2876	.2641	6.3
25-29	3	3.211	.6880	.9836	1.3725	-0.3889	-0.3353	0.3538	.4157	.5169	6.2
30-34	4	4.667	.8321	1.6939	1.1421	0.5518	0.4391	1.0569	1.1145	.7203	6.5
35-39	5	5.609	.9159	2.4321	0.7061	1.7260	1.5117	1.9534			
40-44	6	6.124	.9786	3.8336	.2763.	3.5573	3.2105	3.4130			
45-49	7	6.258	-	-	-	-	-	-			

Notes: (a) Col. (4) is computed by the relation $Z(i) = -\ln [-\ln P_i/P_i+1]$

(b) Cols. (5) and (7) are copied from table 4.6

(c) Col. (8) and Col. (9): the parameters a and b are estimated by the least squares method, using the first four values in Col. (6) as dependent variable values and those in Col. (7) as independent variable values. For the example presented they turn out to be $a = 0.0426$ and $b = 1.0545$, so that $Y(i) = 0.0426 + 1.0545\ Y_s\ (i)$. In Col. (8) Y_s (i) values are taken from Brass (1981). The fitted Y(i) are computed accordingly and placed in Col. (9).

(d) Col. (10) is computed by the relation $P_i/F = \text{Exp}\ [-\text{Exp}\ [-Y(i)]]$.

(e) Col. (11) = Col. (2)/Col. (10).

Table 4.8 Fitting Brass Relational Gompertz Model to current fertility data, Swaziland 1976

Age (x)	i	*Cumulative Fertility F_x	$\frac{F_x}{F_{x+5}}$	Z(x)	e(x)	Z(x)-e(x)	g(x)	$Y_s(i)$	Fitted Y(i)	$\frac{P_i}{F_i}$	F_i
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
20	1	.585	.3154	-.1432	1.3364	-1.4796	-1.4501	-1.0787	-1.0953	.0503	6.5
25	2	1.855	.6013	.6760	1.4184	-0.7424	-.7430	-.3119	-.3431	.2443	6.8
30	3	3.085	.7470	1.2319	1.2978	-0.0659	-.0382	0.3538	.3099	.4802	6.7
35	4	4.130	.8386	1.7371	.9670	0.7701	.8356	1.0569	.9997	.6921	7.7
40	5	4.925	.9146	2.4161	.4509	1.9652	2.1649	1.9534			
45	6	5.385	.9514	2.9993	.0462	2.9531	4.4564	3.4130			
50	7	5.660									

Notes: (a) Col. (2) is obtained by multiplying by 5 the cumulated observed ASFR data of Swaziland 1976, based on births in the last year.

(b) Col. (4) computed by using the relation $Z(x) = -\ln(-\ln F_x/F_{x+5})$

(c) Col. (5) and (7) copied from table 4.6

(d) Col. (8) and Col. (9): the parameters a and b are estimated by the least squares method, using the first four values in Col. (6) as dependent variable values and those in Col. (7) as independent variable values. For the example presented they turn out to be a =-0.0371 and b = 0.981, so that Y(i) = 0.981 Ys(i) - 0.0371. In Col. (8) Ys(i) values are taken from Table C of Brass (1981). Using above relation, the fitted Y(i) are calculated and placed in Col.(9). The rest of the procedure is the same as in table 4.7.

$$P(1,h) = P(1,2) \qquad ...(7)$$

$$P(2,h) = P(1,h) + [P(2,2) - P(1,1)] \qquad ...(8)$$

$$P(i,h) = P(i-1,h) + [P(i,2) - P(i-1,1] \qquad ...(9)$$
where $i = 3,4,..\ 7.$

If the surveys are seperated by ten years then we have:

$$P(1,h) = P(1,2) \qquad ...(10)$$

$$P(2,h) = P(2,2) \qquad ...(11)$$

and
$$P(i,h) = P(i-2,h) + [P(i,2) - P(i-2,1)] \qquad ...(12)$$
where $i = 3,4,...\ 7.$

27. It is not always practicable to have surveys seperated by five or ten years. However, if the fertility has not changed drastically during the intersurvey period , and the interval is one or two years different from five or ten years, no correction is required (Zlotnik and Hill,1981). In Table 4.9 an illustration on the data from Swaziland is given.

B.2- Estimation of ASFR From Parities of Hypothetical Cohorts

28. In method B.1 we obtain the hypothetical parity distribution based on the intersurvey parity increments. It is possible to derive from these hypothetical parities the equivalent cumulative fertility rates ,F', by suitable interpolation. This is the inverse process of obtaining equivalent cumulative fertility to match parities (method A.1). One method of interpolation used in Zlotnik and Hill (1981) and United Nations (1983) is given here :

$$F'(i,h) = .9283\ P(i,h) + .4547\ P(i+1,h) - .0585\ P(i+2,h) - .3245\ F'(i-1,h) \qquad ...(13)$$
for $i = 1,\ 2,..5$ and $F'(0,h)=0$

$$F'(6,h) = .0209\ P(4,h) - .5574\ P(5,h) + 1.0478\ P(6,h) + .2869\ P(7,h) + .2018\ F'(4,h) \qquad ...(14)$$

$$F'(7,h) = 1.007\ P(7,h) \qquad ...(15)$$

29. Once the cumulative age specific fertility rates ,F' are obtained, the ASFR can be deduced by successive subtraction. Table 4.10 gives an illustration using the data from Swaziland for 1976.

30. The above procedure can be used even when we have one point data and obtain ASFR from parity data. These derived ASFR can be compared with those obtained from the data on BLY. This is an alternative way to the earlier P/F ratio of checking consistency of the lifetime and current fertility data.

Table 4.9. Computation of hypothetical cohorts for Swaziland, 1966-1976

Age group	i	Parities P(i,1)	Parities P(i,2)	Inter-censal increments	Parities of hypothetical cohorts
	(1)	(2)	(3)	(4)	(5)
15-19	1	.3168	.3251	-	.3251
20-24	2	1.6589	1.6520	-	1.6520
25-29	3	3.0126	3.2114	2.8946	3.2197
30-34	4	4.1361	4.6665	3.0076	4.6596
35-39	5	4.8499	5.6090	2.5964	5.8161
40-44	6	5.4689	6.1240	1.9879	6.6475
45-49	7	5.7451	6.2583	1.4084	7.2245

Notes: a) Col. (4) obtined as P(i, 2)-P(i-2,1) for i = 3 .....,7

b) Col. (5): for i=1 and 2 the values are copied from column (3)-see equations (10) and (11) of the text; for i=3, ...7, see equation (12) of the text.

Table 4.10. Computation of ASFR from the parity data of hypothetical cohorts for Swaziland, 1966-1976

Age group	i	Parity data of hypothetical cohorts	Cumulated fertility by the end of age group F'(i,h)	ASFR (i) (per woman)
	(1)	(2)	(3)	(4)
15-19	1	.3251	.8645	.173
20-24	2	1.6520	2.4444	.316
25-29	3	3.2197	3.9741	.306
30-34	4	4.6596	5.2916	.263
35-39	5	5.8161	6.2819	.198
40-44	6	6.6475	6.9613	.136
45-49	7	7.2245	7.2751	.063

Notes: a) Col.(2) obtained from Col.(5) of table 4.9.

b) Col.(3) obtained by interpolation using equations (13) (14) and (15) of the text.

c) Col.(4) obtained from Col.(3) as the difference between F'(i,h) and F'(i-1,h), divided by 5.

B.3 and B.4 - Brass Relational Gompertz Method for Hypothetical Cohorts

31. These methods are identical to those described in A.2 and A.3 except that the observed lifetime fertility will be replaced by intersurvey hypothetical parity estimates obtained in B.1 and the cumulative current fertility is based on the average ASFR for the intersurvey period.

4.2.2 METHODS OF ESTIMATING FERTILITY USING DATA ON SEX-AGE DISTRIBUTIONS:

C.1 - Reverse Survival Method

32. In any population the persons aged 0-4 ,in year t, are survivors of births in the five years preceding the census. Similarly persons aged 5-9 are the survivors of births during (t-10) and (t-5) years.

33. In the reverse survival method we project persons in given age groups backwards in time. For example, persons in the age group 5-9 at time t when projected five years and ten years backwards results in persons aged 0-4 at time (t-5), and births during the five years preceding (t-5). Once survival ratios are available for specific periods, reverse surviving becomes easy. We are particularly interested in reverse projecting a population five or ten years backwards in time.

34. It can easily be seen that

$$B(t-5,t).\ S_b(t-5,t) = P_{0-4}(t) \qquad ...(16)$$

where $B(t-5,t)$ denotes births in the five year period $(t-5,t)$, $S_b(t-5,t)$ the survival ratio of births to age group 0-4 at time t and P_{0-4} the population aged 0-4 at time t. From this it follows :

$$B(t-5,t) = P_{0-4}(t) \ / \ S_b(t-5,t) \qquad ...(17)$$

and the right hand side of the equation can be calculated if an appropriate life table is available. In this case

$$S_b = L_{0-4} \ / \ 5\ l_o \qquad ...(18)$$

We get annual births by taking a fifth of the value obtained in equation(17), assuming that births are uniformly distributed.

35. In order to obtain birth rate for the period (t-5,t) we need the mid-year population which forms the denominator. There are two ways of obtaining this. One is to reverse project all the age groups from time t to time t-5, and take the mean of the total populations at time t-5 and t. The second method is to use an estimated rate of growth of

population during (t-5,t) and use this rate of growth to obtain population at the year t-2.5. The birth rates by these two approaches can be written as follows:

$$b_{0-4}=(1/5)[P_{0-4}(t)(5\ l_o/{}_5L_o)]/(1/2)[T(t)+T(t-5)] \quad ...(19)$$

where T(t): Total population at time t.

36. If we use rate of growth method we have to replace the value 1/2[T(t)+T(t-5)] by T(t) Exp(-2.5r) in the denominator of the above equation to obtain the birth rate.

37. In case we are interested in obtaining crude birth rate based on persons aged 5-9 in the year t, we need to reverse survive them first to age group 0-4 at time (t-5), which can be done as follows:

$$P_{0-4}(t-5) = P_{5-9}(t)/\ S_{0-4}(t-5,t) \quad ...(20)$$

The quantity ${}_5L_5$ / ${}_5L_o$ taken from an appropriate life table gives the survival ratio from age 0-4 to 5-9 in equation (20). From the population aged 0-4 in the year (t-5) we can obtain births (t-10) and (t-5) by using equation (16), leading to :

$$B(t-10,t-5)= P_{5-9}\ /\ ({}_5L_5/{}_5L_o)({}_5L_o/5\ l_o) \quad ...(21)$$

If the life tables for the periods (t-10,t-5) and (t-5,t) are equal then ${}_5L_o$ on the right hand side of the above equation will cancel out.

38. The mid-year population during (t-10,t-5) is obtained on the same lines as in the case for 0-4 age group discussed above, with the difference that in this case reverse projection is done two times.

39. If the rates of growth of the population during the periods t-10,t-5 and t-5,t are denoted by r_1 and r_o then the mid-year population for the period t-10,t-5 can be written as :

$$T(t-7.5) = T(t)\ Exp(-2.5r_1)\ Exp(-5r_o) \quad ...(22)$$

From equations (21) and (22) the birth rate based on population aged 5-9 can be obtained. However, if we have two censuses separated by 10 years, then the intercensal rate of growth r will replace r_0 and r_1.

40. The application of the above methods require certain conditions to be satisfied. The most important assumption is that the reported proportion of populations aged 0-4 and 5-9 in particular are correct. Quite often this assumption is violated. The second assumption is the availability and

reliability of mortality data in the sub-Saharan African situation. In most of the countries of this region the main source of mortality is the Brass type indirect estimates. Using these indirect estimates of l_2, l_3 and l_5 a model life table can be selected on the assumption that the age pattern of mortality is known. However, it has been noticed that the estimates of birth rate are more influenced by the age pattern of mortality than over-all level.

C.2 Reverse Survival of Cumulated Populations

41. The methods described above , namely,reverse survival of populations aged 0-4 and 5-9 are well known and are extensively used. However, it is possible to extend the logic to cumulated populations under age 10 or 15. Recent work of Coale(1981) and (Venkatacharya and Tesfay Teklu, 1987a) indicated that reverse survival of populations of both sexes under age 15 , using the life table corresponding to the Brass type of indirect estimate l_5 gave very robust estimates of birth rate. This can be written as:*

$$b_R = C'(15)\ \mathrm{Exp}[\ 7.5\ r\ ]\ /\ {}_{15}L_0 \quad \ldots(23)$$

where C'(15) is proportion of both sexes under age 15 at the second census, and r the intercensal rate of growth. The birth rate thus obtained refers to a point of time 7.5 years prior to the second census. r and L_0 can also be used for both sexes. Computations can be made using either both sexes or one sex i.e females. Once C'(15) value for both sexes is taken as the proportion under 15, say for females, r and L_0 can also be used for females only.

42. Similarly,the birth rate based on proportion under age 10, and a life table corresponding to the Brass type indirect estimate of l_2 can be obtained as :

$$b_R = C'(10)\ \mathrm{Exp}[\ 5r\ ]\ /\ {}_{10}L_0 \quad \ldots(24)$$

The above birth rate refers to the mid-census period. We will later come to the discussion of these methods. Table 4.11 illustrates the application of the method for Swaziland.

---

* Here C'(15) and C'(10) indicate cumulated values upto ages 15 and 10 respectively for the recent census.

Table 4.11 Application of Reverse Survival Methods to Census data of Swaziland (1976).

(1) 1966 Female Pop.	(2) Survival ratios 1966-71	(3) 1971 Female Pop.	(4) Survival ratios 1971-76	(5) 1976 Female Pop.	(6) Age Group
(49018)	-	(51196)	-	-	Births
36910	.85933	42123	.87214	44650	0-4
29635	.92500	34142	.93393	39340	5-9
23942	.96982	28741	.97322	33228	10-14
20177	.97775	23409	.97985	28162	15-19
14482	.97475	19668	.97685	22867	20-24
13810	.97056	14056	.97290	19135	25-29
10141	.96603	13341	.96878	13617	30-34
9539	.96104	9746	.96423	12864	35-39
6765	.95588	9118	.95923	9349	40-44
5813	.95095	6433	.95427	8701	45-49
4868	.94196	5476	.94541	6080	50-54
4220	.92452	4501	.92872	5086	55-59
3434	.89458	3775	.90010	4051	60-64
4155	.84505	2902	.85219	3217	65-69
9165	.77065	3202	.77970	2263	70-74
	.52494	4811	.68085		75-79
	-		.42177	2029	80+
197056		225444		256821	

* The Survival ratios in columns (2) and (4) are taken from Coale-Demeny North model life tables corresponding to estimated levels of 12.65 and 13.65 respectively.

43. From the reverse projection of table 4.11 we get,

(a) Female births per year during 1971-76 = 1/5 (51196) = 10239
Mid-year population for the period 1971-76 = 1/2 (225444 + 256821) = 241133
CBR based on 0-4 population (10239/241133) * 1000 = <u>42.5</u>.

(b) Female births per year during 1966-71 = 1/5 (49018) = 9804
Mid-year population for the period 1966-71 = 1/2 (197056+225444) = 211250.
C.B.R based on 5-9 population = (9804/211250) * 1000 = <u>46.4</u>

(c) Intercensal rate of growth 1966-1976 = .0301 (computed from census figures).
Mid-year population for the period 1971-76 = 256821* Exp(-2.5x.0301) = 238204.
C.B.R based on 0-4 population = (10239/238204) * 1000 =<u>43.0</u>

(d) Mid-year population for the period 1966-71 = 256821. Exp(-7.5 X.0.301).
C.B.R based on 5-9 population = (9804/204922) * 1000 = <u>47.8</u>

4.3 METHODS BASED ON STABLE AND GENERALIZED STABLE POPULATION EQUATIONS

44. It is well known to demographers that a population which is closed to migration and is experiencing constant age specific fertility and mortality rates eventually attains a constant age distribution given by :

$$c(a) = b \ \text{Exp}[-a\ r\ ]\ p(a) \qquad ...(25)$$

where the various letters used in the equation carry the conventional meaning. Such a population is known as a Stable population. When mortality gradually declines, without any change in fertility the population looses its stability and becomes what is known as a 'quasi-stable' population. It is observed that the age distribution of a quasi-stable population is close to the age distribution of that stable population which has the same level of fertility and the current mortality level.

45. Several methods of estimating fertility and mortality based on stable population theory are available. For a detailed description of these methods the reader should refer to (United Nations , 1967; United Nations, 1983). In this chapter we briefly mention a couple of them.

C.3- United Nations Manual IV method.

46. In the United Nations (1967) Manual IV it is suggested that given an age distribution of a population by sex and an estimate of its rate of natural increase, the fertility and mortality levels of the population can be estimated by matching the observed age distribution with that of an appropriate stable population.

47. In sub-Saharan Africa it is very difficult to assume that the countries have been closed to migration, an assumption required for the application of stable population methods. In fact, most of the countries under study such as Ghana, Swaziland, Lesotho, Kenya, Malawi and Tanzania have in the last twenty or so years been either receiving or sending countries. In this study to reduce the effect of migration in most cases the female age distribution of those born in the country or de jure population is used. It is quite safe to assume that fertility in most sub-Saharan African countries has not substantially changed to vitiate stable assumptions.

48. In the case of mortality , many countries have experienced declines in the past but further reductions in the most recent years appear to be sluggish or non-existent. However, a number of authors have suggested corrections to birth rates derived on the stable assumptions to take account of the declines in mortality. (United Nations, 1967; Zachariah, 1970: and Abou-Gamrah, 1976). Some of these methods require the time when the mortality started to decline and the

magnitude of the decline. However, recent methods of estimating birth rates under nonstable conditions renders these corrections unnecessary. These methods will be discussed in the subsequent sections.

C.4 - Coale Method based on C(15) for both sexes and l_5

49. Coale has observed that the birth rates of stable populations having the same proportion of population under age 15 and the same level of mortality did not vary much irrespective of the family of model life table used. Based on this property Coale suggested a robust birth rate estimate:

$$b_C = b_s \text{ Exp}[\ 7.5\ (r - r_s)] \qquad \ldots (26)$$

where b_s and r_s are birth rate and rate of growth of the stable population, which is obtained by matching the observed proportion under age 15 (for both sexes) and observed l_5.

50. The logic behind the above adjustment is explained by Coale by treating the estimation of stable birth rate from C(15) and l_5 as a form of reverse survival method that gives an estimate of the average birth rate during the 15 years preceding the census. The persons under age 15 when reverse survived by life table survival ratios corresponding to estimated l_5 lead to the births during the 15 years preceding the census. To obtain the birth rate we need the denominator, namely, the person years lived during preceding 15 years. These person years lived are obtained by using the rate of increase r which differs for a stable and a non-stable or observed population. For example, the person years lived in the stable population are obtained as P_o Exp[$-7.5\ r_s$] and for the non-stable observed population as P_o Exp[$-7.5\ r_o$]. The value 7.5 is the number of years before the census where the midpoint occurs and P_o is the total current population. Thus the person years in the stable situation and hence the stable birth rate can be adjusted by the factor Exp[$7.5(r_o - r_s)$] to take care of the non-stable situation. For more details the reader should refer to Coale (1981) or United Nations (1983). Since b_C is the average value for fifteen years prior to the most recent census date, t, and l_5 refers to the time about 6-7 years prior to T, b_C should be treated to refer to the time (t-7.5).

51. Equation (26) is shown to be equal to the well known reverse survival method of obtaining birth rate from C'(15), with a life table corresponding to l_5, and r (Venkatacharya

and Tesfay Teklu, 1987a). Denoting by b_R ,the birth rate by reverse survival method, we have (== is used to denote "approximately equal to"):

$$b_R == b_C = C'(15) .Exp[7.5 r] / {}_{15}L_o \quad ...(27)$$

52. Equation(27) is simpler than equation(26) in that it does not contain stable parameters b_s and r_s. This results in a certain amount of computational ease. However, the denominator of equation (27) contains ${}_{15}L_o$, requiring the use of Model Life Tables , if a life table for the study population is not available. Further, simplification of the above equation is made by exploiting the very high correlation between l_5 and ${}_{15}L_o$ leading to :

$$b_V = C'(15) \ Exp[7.5 r] / [u + vl_5] \quad ...(28)$$

where u and v are constants obtained by fitting a straight line between ${}_{15}L_o$ and l_5 for various Model Life Table Families. Values of u and v are tabulated and presented in Venkatacharya and Tesfay Teklu (1987b, p11) and reproduced in table 4.12. If prior knowledge of the mortality pattern of the study population is available then the appropriate set of u and v can be used in equation (28). However,it has been observed that variations in the model life table families only result in minor changes in the estimated birth rate (Coale, 1981; Venkatacharya and Tesfay Teklu, 1987b). Therefore, in the context of developing countries where one does not have a firm knowledge of the pattern of mortality, the values of u and v that correspond to Coale and Demeny West model life tables can be used as they result in minimum percentage error in the estimated birth rate. Thus, the following birth rate estimate which uses only r, C'(15) and l_5 can be used irrespective of the true mortality pattern of the study population without any serious error in the estimated value (Venkatacharya and Tesfay Teklu, 1987 b):

$$b'_V = C'(15) \ Exp[7.5 r] / [.365+14.6 \ l_5] \quad ...(29)$$

<u>Sensitivity of birth rate</u> :

53. One of the advantages of equations (28) and(29) is that they permit easily, to make sensitivity analysis of the birth rate for errors or changes in the arguments, namely, C'(15),l_5 and r. From equation (28) one can deduce that a change of 100k% (i.e for 5% change k = .05) in C'(15) leads to 100k% change in the estimated birth rate. In other words the

Table 4.12. The values of u and v for various Model Life Table Families assuming a linear relationship between $_{15}L_0$ and l_5.

Model Life Table Family	u	v
Coale-Demeny		
West	.365	14.599
North	.145	14.809
East	.309	14.663
South	.553	14.415
United Nations (New)		
Latin American	.302	14.691
Chilean	.284	14.712
South Asian	.511	14.475
Far Eastern	.024	15.018
General	.147	14.851
O.E.C.D		
Region A	.871	14.061
Region B	.708	14.246
Region C	.749	14.197
Region D	.512	14.451
Region E	.618	14.340

Note: Except for Region A of O.E.C.D in all cases the R^2 is equal to 0.9999.

birth rate estimate is as good as our observed C'(15). A 100k% error in r leads to 100[Exp(7.5rk)-1]% change in the estimated birth rate. A 100k% change in l_5 results in

$$-100\left[1-\frac{u+vl_5}{u+vl_5+vl_5k}\right]\%$$

in the estimated birth rate. In the case of l_5 the change works in the opposite direction. Calculations made for some Asian and African countries indicate that a 5% change in each of C'(15),r and l_5 , one at a time , leads to about 5%, 1.2% and -4.5% change in the estimated birth rate respectively. Two very interesting points emerge from this. (i) The robust estimate of birth rate is more affected by changes in C'(15) and l_5, and to a lesser extent by changes in r. (ii) The errors in C'(15) and l_5 act in opposite directions thereby reducing the net effect on the birth rate in some situations.

D.1 <u>Robust Birth Rate Estimate by Preston -Coale Equations</u>

54. For a nonstable population , we have the folowing relation given by Preston and Coale (1982):

$$c(a,t) = b_n(t)\ \mathrm{Exp}\left[-\int_0^a r(x,t)\,dx\right] p(a,t) \qquad \ldots(30)$$

where $b_n(t)$ is the birth rate of the observed nonstable population, r(x,t) are the age specific rates of increase of the population aged x and p(a,t) is the life table survival probability to age a, <u>all referring to time t</u>. In the application of the above equation to data of two successive censuses, c(a,t), r(x,t) and p(a,t) are taken such that t refers to the midpoint of the intercensal period.

55. In the method suggested by Coale which was discussed in the previous section we have noticed that the robustness of the birth rate originates from the dependability of the three pieces of data C'(15), l_5 and r in that order of importance. Taking advantage of this knowledge, we can use the Preston - Coale equation to estimate the birth rate using C(15), l_5 and r(x) which should lead to a robust birth rate estimate. Here C(15) denotes the person years (both sexes) lived under age 15 centred at the midpoint of the two censuses. Dropping the subscript 't' we have from equation (30):

$$C(15) = \int_0^{15} c(a)\,da = b_n \int_0^{15} \mathrm{Exp}\left[-\int_0^a r(x)dx\right] p(a)\,da$$

$$== b_n\ \mathrm{Exp}\,[-\ 7.5\ r_w]\ \ {}_{15}L_0$$

leading to :

$$b_n = C(15) \, Exp[7.5 \, r_w] \; / \; {}_{15}L_o \qquad \ldots(31)$$

where

$r(x)$ refers to intercensal rate of growth, $p(a)$ to a life table corresponding to l_5, and $r_w = (1/7.5) \; [5 \, r_{0-4} + 2.5 \, r_{5-9}]$.
r_{0-4} and r_{5-9} refer to intercensal rates of growth for the age groups 0-4 and 5-9 respectively.

56. The equation(31) is very similar to equation(27) and has been found to give robust estimates for a number of sub-Saharan African countries. There are other variants of the above estimate, one of them which is likely to give better results is as follows:

$$b^*_n = C(15) \, Exp[\, 7.5 \, r'_w] \; / \; {}_{15}L_o \qquad \ldots (32)$$

where :

$$r'_w = (1/7.5) \; Ln \, [\, (Exp(-a_1) \; {}_5L_o/{}_{15}L_o) + Exp(-a_2) \; ({}_5L_5/{}_{15}L_o) + Exp(-a_3) \; ({}_5L_{10}/{}_{15}L_o)]$$

and $a_1 = 2.5 \, r_{0-4}$
$a_2 = 5 \, r_{0-4} + 2.5 \, r_{5-9}$
$a_3 = 5 \, r_{0-4} + 5 \, r_{5-9} + 2.5 \, r_{10-14}$

57. For a fuller discussion of these estimates in relation to the earlier Coale estimates reference can be made to Venkatacharya and Tesfay Teklu (1987a). It is very easy to see that equations parallel to (28) and (29) can be obtained in a straight forward manner from equations (31) or (32).

58. However, one important point should be mentioned here. The estimates of the birth rates given by b_n and b^*_n are applicable for the mid-point of the intercensal period or 5 years before the second census when censuses are ten years apart; whereas b_C, b_R, b_V and b'_V are applicable to a point of time 7.5 years before the second census. Unless the birth rate is nearly constant during the period under question these two sets of estimates will not be strictly comparable. Therefore, b_n and b^*_n have a very useful property, namely, that they give birth rates for more recent period than the estimates like b_C.

59. Thus using the same data as in Preston integrated method, and also using the same basic equation we have obtained a robust estimate of birth rate. The results are presented below:
C.B.R using equations (27), (28) and (29)
Brass type estimate (using Multipliers of U.N Manual X, p 77) of l_5 = .7927 for both sexes; and l_5 for Females is .8106.

Also $_{15}L_o$ for females (North Family Model Life tables) = 12.12604 where l_o = 1.
Proportion of females under age 15 in 1976 = .4564
Proportion of males and females under age 15 in 1976 = .4776
Using equation (27)

$$C.B.R = [.4776 \times Exp.(7.5 X .0301)] /12.12604] * 1000$$
$$= \underline{49.4}.$$

Using equation (28) for North family

$$C.B.R =[.4776 \times Exp.(7.5 X .0301)] / (.145+14.806 X .8106) X 1000$$
$$= \underline{49.3}$$

Using equation (29)

$$C.B.R =[.4776 X Exp.(7.5 X .0301)] /(.365 + 14.6 X .8106)]X1000$$
$$= \underline{49.1}$$

D.2 Preston Integrated Method

60. Preston(1983) approached the problem of estimating birth rate using equation (30) in a way different from the one indicated in the previous section. He suggested a method of estimating both birth rate and adult mortality simultaneously. The method can be briefly stated as follows :

61. We have from equation(30), dropping t :

$$1/p(a) = b \, Exp[-\int_0^a r(x) \, dx] \; / \; c(a) \qquad ...(33)$$

but p(a) can be written as:

= l_5 $_5p(a)$, where $_5p(a)$ is the conditional probability of survival from age 5 to age a and l_o is taken as unity throughout this chapter, so that l_5 gives the probability of survival to age 5. Therefore equation (33) can be written as :

$$1/_5p(a) = l_5 \, b \, Exp[-\int_0^a r(x) \, dx] \; / \; c(a) \qquad ...(34)$$

62. Further, for ages above 5 assuming that any life table in one Model Life Table family can be obtained by logit transformation of an arbitrarily chosen life table from the same family we have :

$$1/_5p(a) = e^A [\, _5q^s(a) / \, _5p^s(a)]^B + 1 \qquad ...(35)$$

where $_5q^s(a)$ and $_5p^s(a)$ are taken from standard life tables.

Table 4.13. Application of Preston intergrated method to Swaziland data, 1966-1976

Age group (x- x+5)	Female Population $_5N_x$ 1966	1976	Age specific growth rate(ASGR) $_5r_x$	Person years lived by each age group during the inter-censal period ($_5PY_x$)	Cumulated ASGR to beginning of next age group	Anuualized Proportion of inter-censal PY lived at each age		$\frac{q^{*s}(x)}{p^{*s}(x)}$	
(1)	(2)	(3)	(4)*	(5)*	(6)*	(7)*	(8)*	(9)*	(10)*
00-04	32678	44896	0.03176	38469.77	0.0				
05-09	30627	39739	0.02605	34978.89	0.03176	0.03209	21.083	0.0	47.1(1)
10-14	25136	33834	0.02972	29266.49	0.08753	0.02807	21.159	0.0450	48.8(2)
15-19	19473	29201	0.04052	24007.90	0.08753	0.02328	21.990	0.0700	49.0(3)
20-24	16433	23998	0.03787	19976.23	0.12805	0.01922	21.750	0.0976	49.7(4)
25-29	16319	20208	0.02137	19198.41	0.16592	0.01668	20.739	0.1307	51.4(5)
30-34	10851	14387	0.02821	12534.56	0.18729	0.01343	23.147	0.1705	51.9(6)
35-39	10291	13470	0.02692	11809.06	0.21550	0.01064	25.373	0.2183	
40-44	7120	9768	0.03162	8374.46	0.24242	0.00882	26.754	0.2755	
45-49	6619	8988	0.03059	7744.36	0.27404	0.00704	28.617	0.3426	
50-54	5501	6285	0.01332	5885.89	0.30463	0.00596	29.009	0.4205	
55-59	3415	5195	0.04195	4243.15	0.31795	0.00443	36.513	0.5252	
60-64	3361	4138	0.02080	3735.58	0.35990	0.00349	37.578	0.6794	
65+	9472	9814	0.00355	9633.80	0.38070				

*Notes:

Column (4) $_5r_x = \ln(_5N_x(1976)/_5N_x(1966)) / 10$, where $_5N_x$ is the population between ages x and (x+5).

Column (5) $_5PY_x = [_5N_x(1976) - _5N_x(1966)] / [10 * _5r_x]$

Column (6) $S_x = \sum {}_5r_{i-5}$; for x = 10 , 15, ...,and i = 5,10,.. (x-5)

Column (7) $C(x+5) = (_5PY_x + {}_5PY_{x+5}) /$[10 times the total of values in column (5)]

Column (8) $[l_5/c(x)].\text{Exp}[-5 * S_x]$

Column (9) $q^{*s}(x) = [1 - p^{*s}(x)]$. The values are obtained from Coale-Demeny North Model Life Tables of Females with expectation of life at birth equal to 30 years

Column(10) It is obtained by linear fit to data in columns (8) and (9). See text at equation (37)

For most practical purposes B can be taken as unity leading to:

$1/{}_5p(a) = K\,[{}_5q^s(a)/{}_5p^s(a)]+1$, where $e^{\lambda} = K$ ...(36)

From equations (34) and (36) we have

$$l_5\,\mathrm{Exp}[-\int_0^a r(x)dx\,]/c(a)$$

$$= (1/b)+(K/b)\,[{}_5q^s(a)/{}_5p^s(a)] \qquad ...(37)$$

63. The left hand side of the above equation involves known quantities. On the right hand side the quantity in the square brackets is also known, based on the standard model life table chosen. Thus equation(37) expresses a linear relation between two known quantities, and the intercept of this line gives the reciprocal of birth rate and the ratio of the slope and intercept gives a measure of the level of adult mortality.

64. On the left hand side of Equation(37) two quantities need some explanation as to how they could be computed. The integral in the exponent is evaluated by the method illustrated for the computation of r'_w in equation (32). The quantity c(a) is the ratio of a^{th} birth days to total person years lived between the ages a and a+5 during the intercensal period and is calcilated by the following method:
First Calculate:

$$PY(a,\ a+5) = \frac{[\text{Popn. }(a,a+5)\text{ in the second census}] - [\text{popn. }(a,a+5)\text{ in first census}]}{r(a,\ a+5).h} \qquad ...(38)$$

where r(a,a+5) is the age specific rate of growth for the age group (a,a+5) during the intercensal period and h is the interval in years between censuses. We can then obtain c(a) as:

$$c(a) = \frac{Py(a,a+5) + Py(a-5,a)}{10.\ (\text{Total of PY over all ages a})} \qquad ...(39)$$

65. Equation (37) can be fitted by various methods. Preston (1983) suggested the use of group mean method. The mean values for the ages in the range 10-30 and 35-55 can be used as two pairs of points to estimate the intercept and slope. In Table 4.13 an application of this method is illustrated.

66. Clear cut linear trend is not always perceivable in data sets that suffer from errors . The other factor that may contribute to the lack of a linear trend is the assumption that the mortality pattern of the study population belongs to a known model family and can be represented as a logit transformation of the standard from this family.

67. Preston(1983) illustrated the application of the method to data from Korea and India . The same method has been applied to some sub-Saharan African countries (Venkatacharya and Tesfay

Teklu,1985). The results are shown in table (4.15) presented later. The age range chosen to fit the trend are taken as 10-30 and 35-55. For the same data birth rates obtained by other methods described in the earlier section are presented elsewhere and may be compared.

68. One other important point observed in dealing with rough data is that the choice of the age groups in the estimation purpose by equation (37) affects the results.

69. Preston (1983) in his study thoroughly investigated the sensitivity of his method to errors in the data. We can summarize these briefly as follows:

(a) The choice of the points at higher ages in the fitting of the trend by equation(31) does not affect the birth rate seriously as it is mostly affected by the cluster of points at the younger ages, whose mean value is almost invariant with respect to errors.

(b) Differential completeness of the two censuses, if invariant with age, does not affect birth rate seriously, but can affect estimated adult mortality seriously.

(c) The most important element that can affect the estimation is the net migration. The effect of errors in the net migration figures used in the estimation of birth rate and adult mortality depends on the magnitude of the error and the ages at which this occurs --i.e young age or old age . If the errors occur in the age range 15-30 then the estimates are less affected.

(d) Interestingly, the estimation of birth rate is insensitive to omission of young children , say, aged 0-4. However, systematic overstatement of ages, especially at higher ages, affects adult mortality estimates.

(e) Errors in the pattern of mortality assumed will affect adult mortality estimation more than the birth rate.

(f) The errors in the estimated value of l_5 are very important for the estimation of the birth rate. but not for the estimation of adult mortality. Fortunately, in most developing countries l_5 is found to be reliable.

4.4 PROBLEMS IN THE APPLICATION OF THE METHODS OF ESTIMATION

4.4.1 Problems in data

70. The application of P/F ratio method has been reported to be successful in the case of Francophone African countries (Brass et al,1967;Page,1975) and some Latin American Countries (Samoza,1981) but in the case of Anglophone African countries the results are different. The reasons for the method being not very successful in these countries could be many, but only some of the important ones are enumerated here. The failure of the method to estimate successfully the true level of fertility can be due to the violation of the basic assumptions and/or deficiencies in the data collected.

71. All the methods described under group A assume constant fertility preceding the survey date so that both lifetime and current fertility are comparable. But in many developing countries fertility is declining though in much of sub-Saharan Africa fertility is still likely to be constant.

72. In the P/F ratio analysis the recommended P_2/F_2 adjustment is based on certain assumptions. It is assumed that women in the age group 20-24 remember the number of children borne by them sufficiently accurately, because such women will have only a small number of births that could have occurred in recent years reducing recall error. Further, as such women must have entered into reproductive life in recent years , they do not form a 'selective' group. Even if fertility had declined in the years preceding the survey date its effect on this group will be minimal. Similarly the cumulative fertility F_2 which is based on women in the age range 15-25 is assumed to have only such errors that will have no effect on the age structure of current ASFR. In some situations these assumptions may not hold and it is difficult to foresee this or verify it in a direct way.

73. In order to average out the errors in the use of P_2/F_2 some authors have used means of various combinations of ratios such as $1/2(P_2/F_2+P_3/F_3)$ or $1/3(P_2/F_2+P_3/F_3+P_4/F_4)$ or weighted average of P/F values with number of women as weights in the respective ages.

74. Before proceeding to further discussion on the methods of estimation we need to look at the deficiencies in the data collected.

75. Estimated MNCEB can be affected by errors in the reported number of births (numerator) and women of specific age groups (denominator). The most important error in the reported number of births is the omission of births by older women, especially those births that ended in early deaths. Women in the older age groups also tend to forget grown up children, children born to another husband and children not present at home for various reasons. There are also factors that may tend to inflate the number of births by the inclusion of step or adopted children, grand children etc. Another error in the reported number of children is the inclusion of still births. The net effect of these errors is a tendency for mean number of children steadily decreasing as age of the women increases.

76. Among the serious errors that result primarily from the enumerator's inefficiency is the parity 'not stated' group. Sometimes a blank or a dash is used to indicate zero parity which then gets included in the 'not stated' group. If the relation between the reported proportion with parity not stated and the reported proportion childless is linear then El-Badry's method could be applied to filter out the childless women from the 'not stated' group. Otherwise it is recommended that these women be used in the denominator (United Nations ,1983).

77. One of the major problems in Africa is the measurement of age. Many types of errors are identified and adjustments suggested (Ramachandran,1983; Hobcraft et al, 1982). Certain types of age misstatement like random errors do not affect P/F ratio method, especially when the errors are such that they have the same distribution at all ages (Brass et al, 1967). However, there are other types of errors in the reproductive ages that are serious. A systematic shift of ages upwards till some boundary age or age group is reached and a shift downwards thereafter to the end of reproductive age affects the mean parities. The mean parities below the boundary age will be understated, the age group containing the boundary age will be correctly reported and those above the boundary will be overstated (United Nations, 1983).

78. A very interesting study that brought out a number of limitations as well as the evaluative capability of P/F ratio method based on real data was that, using the maternity histories of the WFS project. As the maternity history data contains two dimensions , age and cohort, many indices of fertility can be computed which can be used to test internal consistency more effectively than the situation where we have only data on MNCEB and BLY. For instance, the P/F ratios can be computed for various periods prior to the survey, for first birth orders and so on. On the basis of WFS data for nineteen countries for the most recent period preceding the survey it was observed that P/F ratios computed by duration of motherhood depicted fertility patterns better than the ratios computed by duration of marriage. The usual P/F ratios for age groups were found to be least effective. The P/F ratios by age for Jordan, Pakistan and Nepal showed steady decline by age groups, but when the ratios are indexed by duration of motherhood they showed a steady trend indicating that reporting errors are small and fertility remained constant in the past years. In the case of Malaysia, Fiji, Sri Lanka, Peru, Mexico, the Dominican Republic, Jamiaca and Indonesia the P/F ratios by age were almost constant but when computed by duration of marriage they showed a rising trend indicating recent declines in fertility. " A blind application of the P/F ratio technique, as originally propounded , would take these rather constant values by age as being indicative of a measure of dating errors for the five years before the survey and adjust the current fertility levels accordingly" (Hobcraft et al,1982). This clearly illustrates the pitfalls in the primary dependence on the P/F

adjustment. It appears that the P/F ratio should be used more as a diagnostic test of internal consistency than an adjustment factor. Even in this case the analyst will have to be cautious in interpreting the ratios.

79. Much of what is said about P/F ratios applies to the relational Gompertz model as they both use the same data. However, Brass states that the relational Gompertz model (using F_i/F_{i+1}) has certain adavantages. The RGM is fitted using the data on BLY for the younger age groups only thereby requiring the assumption of errors being constant for these ages alone and not for all ages as is assumed in the origianl P/F method. " Since the fitting 'averages' the current rates, the estimated F's are less vulnerable to chance and erratic errors in the measures at under 25 than with the traditional P_2/F_2 correction." (Brass, 1981).

80. Women who are included in the survey are the survivors of the cohort of women that entered the reproductive life some years ago. The parity and other fertility data of the surveyed women will give representative fertility estimates if the women who died are no different from those who survived with respect to fertility performance. If fertility and mortality are associated then the estimates based on the current survivors will give biased estimates. For Instance, if we assume that women are heterogeneous with respect to mortality risks, and are also heterogeneous with respect to the risk of bearing a child then those women who die early will on the average tend to be of higher fertility than the survivors. This effect of selectvity will be more pronounced for women at higher age groups than younger age groups. It is hoped that heterogeneity of women is small and that as the general mortality risks are low in the child bearing ages the effect of this on the mean parity, especially at lower ages, will be small.

81. The other important factor that leads to selectivity when dealing with sub-national populations or populations with heavy net migration is the movement of people in the reproductive age groups. If fertility of the women in places of destination and origin of migration are different, this selectivity becomes prominent and the estimated fertility will be affected. Analysis of fertility seperately by place of birth or duration of residence will reduce this error.

82. Retrospective data on the number of children ever born for women at higher ages suffers from various errors of omission. In the early questionnaires only one question on the number of live births ever born by a woman was asked. In the later questionnaires three pronged question on the number of children living in the household, elsewhere and dead was asked. It was reported that this lead to improvement in the reporting of births in Kenya (1969),Uganda(1969), Swaziland(1976) and the Gambia(1973). Still the data on parity was found to be deficient. Similarly instead of asking the occurrence of birth in the last

year, asking the date of last live birth is expected to give better results on the current fertility. Though these improvements are theoretically sound due to a number of practical problems mostly resulting from enumerator's negligence and fatigue, the data still show deficiencies.

4.4.2 Problems in the application of the Methods of estimating birth rate using two successive age distributions.

83. The use of reverse survival method for the estimation of fertility in Africa south of the Sahara is faced with a number of problems. The accuracy of the estimates derived by this method heavily depend on the quality of reporting of the age data especially for children under ten years. In most African censuses, save Sierra Leone 1963 and Liberia 1962, the proportion under age five always implied crude birth rates lower than age 10. Omission of infants from census enumeration for cultural and other reasons is alluded to as one of the reasons among a number of others that lead to a low proportion in 0-4 age group. Children who have not passed certain rituals are not considered as grown ups in some African countries and it is likely that these might not be reported in population censuses and demographic surveys. Another error that affects 0-4 population is 'heaping' at age five. Children in the age group 3-4 seem to be reported in the age group 5-9 either due to preference for digit five or due to the fact that age 6 is the mininum age of entry into school. In most cases, therefore, reverse projection of 0-4 age group is likely to yield a lower estimate of births in the five years preceding the census.

84. On the other hand the population 5-9 seems to be inflated by the movement of people under age five and over age ten into the 5-9 age group. An evaluation of most censuses in Africa clearly indicates that the crude birth rates based on the populaion under age ten is much higher than those based on population under five and fifteen. A relatively higher proportion of population aged 5-9 than those aged 10-14 is surprising because the terminal digit of the former group is an avoided digit. It therefore appears that the reason for people belonging to age group 10-14 to shift downwards might not be due to digit preference but to such factors as the minimum age for admission to schools, physical development and appearance etc.

85. In situations where the 0-4 age group is underreported but 5-9 is overreported, it might be advisable to use the proportion under ten to estimate the total number of births in the decade before the census. Such a solution is not without difficulties. As the age group widens it becomes difficult to locate the mortality level in time. Brass-type indirect estimates of l_5 will give the level of mortality as of 6-7 years preceding the census. Therefore in situations where mortality is declinig, use of this to reverse survive children leads to overestimation of births. However, the magnitude of this is likely to be small in view of the recent retardation in the decline of mortality.

86. The other serious error comes from the use of rate of growth in obtaining the mid-period total population which has been referred to earlier. In Africa, estimated rate of growth from two censuses is mostly in error because of unequal coverage in the two censuses. In many countries no serious evaluation of census coverage is done. Indirect methods of evaluation also suggest that in most cases the two censuses are not of equal coverage. That is, the intercensal rate of growth might be on the higher side if the coverage of the recent census is better than the earlier one, as we suspect in the cases of Liberia 62-74, Lesotho 66-76, and Malawi 66-77. On the other hand if the recent census is poorer in coverage than the earlier one the reverse is expected which is suspected in such cases as Swaziland 66-76 and Sierra Leone 63-74. A similar finding is reported for some countries in Asia (Luther, 1983)

87. A different problem arises when reverse surviving the entire set of age groups using a selected life table corresponding to Brass type estimate of l_i, (i=2,3 and 5). While Brass-type estimates may provide robust estimates of childhood mortality they may not be helpful in selecting adult mortality. Orphanhood and widowhood data could provide estimates of adult mortality for females above 25 years and males beginning from age 27.5 or 32.5. Experience has shown that orphanhood data did not give satisfactory results due to a number of difficulties in concepts in Africa such as fostering, age reporting. There are also some technical difficulties like more than one child reporting the survival of the same parents. The use of widowhood data has also problems due to the loose definition of marriage, prevalence of polygyny, remarriage, age reporting etc.

88. The age pattern of mortality of the life table used in the reverse surviving of the entire age range also affects the results. Demographers have preferred 'North' family life tables in the Coale-Demeny Model Life Tables for many sub-Saharan African countries. But, some studies based on small populations in Senegal and the Gambia indicate that the true pattern of mortality in Africa does not seem to indicate a decline in l_x with age as in the models we know so far (Cantrelle, 1974; Garenne, 1981). Although it is difficult to ascertain whether the mortality curve at early ages in sub-Saharan Africa forms a convex curve, indications are that the existing model life tables may not cover the patterns of mortality in the region.

89. The reverse projection of population of all the age groups, unlike the rate of growth method, tends to exaggerate the net census coverage error.

90. Some of the errors discussed above are not peculiar to reverse survival methods, but are also applicable to stable population methods.

91. No population is strictly stable. Fertility can vary due to factors as recent wars, epidemics, famines etc. Age and sex selective migration through its effect on age sex distribution and more particularly its impact on family formulation and fertility levels could affect the age-sex distributions. Though, less significantly, the declining mortality also affects stability.

92. Due to errors in the data, it is difficult to identify an appropriate model stable population to represent the study population, even though the study population is in fact stable. This is because, the characteristics or indices of the population we use to select a model stable population are affected by reporting errors, such as differential omission of people by age, the differential coverage of censuses etc.

93. As model stable populations are tabulated by various patterns of moratility by age and sex, knowledge about the patterns of moratlity is necessary, which is difficult to obtain in the context of developing countries.

94. Therefore, in order to obtain reliable birth rate estimates we need to choose those population charactersistics that are likely to be less affected by the various errors mentioned above. Secondly, it is also necessary that the charactersitics chosen are such that they serve to identify the appropriate stable population. In other words, the characteristics used are such that, the birth rate estimated is nearly invariant to the type of model life table system selected.

95. There are many methods used in the past in matching a model stable population to the population under study. But a method suggested by Coale satisfies the above two desirable properties. This is the method of selecting a model stable population using C(15) of both sexes and l_5. The birth rate of such a model stable population, surprisingly gave accurate estimates of birth rate even when the population in question is far from stable.

4.4.3 Direct Studies of Errors and Biases:

96. One of the major difficulties in the estimation of fertility is that to date not many studies are available that directly analysed the nature and magnitude of errors in data (Som,1973). This is partly because of the difficulty of measuring errors in a direct way. It is useful to briefly mention some of the recent studies which gave some encouraging results.

97. In the Gambia comparison of the 1973 census results on 573 women aged fifteen and above when matched with the efficiently conducted follow-up study by the Medical Research Council showed that the Census could only obtain correct parities in 60% of the cases (Gibril,1976). A further analysis of 99 tape recorded interviews comprising of 764 persons indicated that for about 30% of the time the enumerators did not bother to ask the three

questions on children at home, elsewhere and dead. Similarly in more than 50% of the cases no effort was made to see that the children are not adopted ones.

4.5 Estimates of Fertility for some countries South of Sahara

(a) Fertility estimates using data on MNCEB and BLY:

98. In Table 4.14 the P/F ratios along with the estimates of TFR derived by various methods described in section 4.2 for some countries in Africa are shown.

99. The P/F ratios for the age group 15-19 when compared to those of the middle age groups show erratic fluctuations which are due to errors in the data and partly due to small number of births in that age group. For countries such as Kenya(1978), Lesotho (1971), Libya(1973), Cameroon (1978), Ghana (1978) and Nigeria (1981) the P/F ratios in the middle age groups are very close to unity indicating consistency in the data and also the validity of underlying assumptions of the method.

100. For countries such as Ghana(1971), Kenya(1969), Tanzania (1973), Ethiopia(1970), Malawi(1972, 1977), Swaziland(1976) and Botswana(1971) the ratios for the age range 20-35, on the average, are between 1 and 1.2. Demographers consider this value as an indication of reasonably good data considering the magnitude of errors expected in African data.

101. The ratios for the countries such as the Sudan(1973), Swaziland(1966), Zambia(1974) and Uganda(1969) tend to be in the neighbourhood of 1.3 for the age range 20-35 and the data should be considered defective or the underlying assumptions are violated.

102. Taking a P/F ratio of 1.2 or less as an indication of reasonably good data we note that 12 cases out of 20 studied show that the data are useful.

103. The total fertility rate estimated by the two formulae $(P_3)^2/P_2$ and $P_2(P_4/P_3)^4$ are close in the cases of Tanzania (1973), Uganda(1969) and Ghana(1971), but in most cases they give different results. For a number of countries the two values fall on either side of the reported parity for the age group 45-49.

Table 4.14
The Values of P/F Ratios and Estimates of TFR by Various methods

	SOUTHERN AFRICA									*	*
	Malawi			Swaziland			Lesotho			Zam	Bot
	72	77	HYP	66	76	HYP	71	79	HYP	74	71
Age group	P/F Ratios										
15-19	.88	1.62	.88	1.61	1.20	1.21	1.10	3.77	3.77	1.83	1.08
20-24	1.17	1.36	1.52	1.54	1.25	1.26	1.07	1.22	1.22	1.64	1.22
25-29	1.20	1.23	1.42	1.36	1.23	1.24	1.10	1.00	1.15	1.29	1.17
30-34	1.13	1.21	1.36	1.26	1.25	1.25	1.06	1.03	1.14	1.37	1.16
35-39	1.14	1.15	1.32	1.17	1.20	1.25	1.06	0.98	1.10	1.23	1.09
40-44	0.91	1.10	1.25	1.15	1.17	1.27	1.04	0.96	1.09	1.01	1.04
45-49	1.04	1.10	1.19	1.12	1.11	1.29	0.97	0.96	1.08	0.96	0.97
Estimates of Total Fertility Rate by Coale-Demeny and Brass											
CO-DE *	7.1	6.4	7.6	5.5	6.2	6.3	5.9	4.8	6.4	5.2	5.8
BRASS *	6.0	8.0	8.3	5.9	7.4	7.2	5.1	7.7	6.4	13.0	6.5
P_7*	6.9	6.9	7.8	5.7	6.3	7.2	5.1	5.4	6.1	6.0	5.6
Estimates of Total fertility Rate by P/F Ratio Method											
$(P/F)_1$*	7.8	8.9	10.1	8.0	7.1	7.1	5.6	6.9	6.9	10.5	7.1
$(P/F)_2$*	7.9	8.5	9.8	7.5	7.0	7.1	5.6	6.3	6.7	9.4	6.9
$(P/F)_3$*	7.8	8.3	9.6	7.2	7.1	7.1	5.6	6.1	6.6	9.2	6.9
$(P/F)_4$*	7.8	8.0	9.3	6.8	7.0	7.0	5.6	5.8	6.5	8.5	6.8
Estimates of Total Fertility Rate by Brass RGM											
RGM(P)*	6.1	7.3	8.1	5.4	6.7	6.7	5.0	8.0	8.1	8.6	5.6
RGM(F)*	7.1	7.6	9.1	6.7	6.6	6.6	5.4	5.5	6.0	8.3	5.8
UNP	7.0	7.0	-	6.5	6.5	-	5.4	5.4	-	6.9	6.5

* Notes: Zam-- Zambia   Bot-- Botswana
CO-DE -- Coale-Demeny Formula= $(P_3)^2/P_2$,
BRASS -- Brass Formula = $(P_2).(P_4/P_3)^4$,
P_7 -- Reported MNCEB for women 45-49
$(P/F)_1$ -- Using P_2/F_2,
$(P/F)_2$ -- Using $1/2(P_2/F_2+P_3/F_3)$
$(P/F)_3$ -- Using $1/3(P_2/F_2+P_3/F_3+P_4/F_4)$
$(P/F)_4$ -- Using $1/2(P_3/F_3+P_4/F_4)$,
RGM(P) -- Brass RGM using P_i/P_{i+1},
RGM(F) -- Brass RGM Using F_i/F_{i+1},
UNP -- United Nations Projections

Table 4.14 (continued)
The Values of P/F Ratios and Estimates of TFR by Various methods

	------EASTERN AFRICA----------						--NORTHERN AFRICA--			
	Kenya			Tan*	Uga*	Eth*	Lib*	Sudan		
	69	78	HYP*	73	69	70	73	73	78	HYP*
Age group	P/F Ratios									
15-19	1.64	.88	.88	1.30	1.44	1.23	1.00	1.57	4.00	1.36
20-24	1.35	1.01	1.01	1.18	1.42	1.15	1.08	1.44	1.79	2.04
25-29	1.22	1.03	1.03	1.10	1.27	1.10	1.04	1.28	1.21	1.54
30-34	1.17	1.06	1.05	1.07	1.17	1.10	1.00	1.18	1.17	1.47
35-39	1.10	1.04	1.05	1.01	1.08	1.07	1.00	1.11	1.11	1.50
40-44	1.01	1.01	1.07	.92	.96	1.03	.97	1.00	1.01	1.53
45-49	.93	.99	1.10	.89	.94	1.04	.96	.94	.97	1.64
Estimates of Total Fertility Rate by Coale-Demeny and Brass										
CO-DE *	6.6	7.7	7.7	5.7	5.9	5.1	8.3	6.0	5.3	6.8
BRASS *	7.3	8.7	8.5	5.8	5.8	5.6	7.8	5.6	10.3	10.1
P_7*	6.1	7.9	8.8	5.0	5.0	5.1	7.6	4.8	6.0	9.4
Estimates of Total Fertility Rate by P/F Ratio Method										
$(P/F)_1$*	8.9	8.1	8.1	6.6	7.6	5.7	8.7	7.5	11.1	11.7
$(P/F)_2$*	8.5	8.2	8.2	6.4	7.2	5.5	8.5	7.1	9.3	10.3
$(P/F)_3$*	8.2	8.3	8.3	6.3	6.9	5.5	8.3	6.8	8.6	9.6
$(P/F)_4$*	7.9	8.4	8.4	6.1	6.6	5.4	8.1	6.4	7.4	8.6
Estimates of Total Fertility Rate by Brass RGM										
RGM(P)*	6.8	8.3	8.1	5.6	5.5	5.3	7.4	5.4	8.1	9.3
RGM(F)*	7.7	7.8	7.8	5.8	6.7	5.2	8.2	6.5	8.8	9.3
Estimates of Total Fertility Rate By U.N.Projections										
UNP	7.6	8.1	-	6.5	6.1	6.7	7.4	6.6	6.6	-

* Notes: Tan-- Tanzania Uga-- Uganda
Eth-- Ethiopia Lib-- Libya
HYP-- Using Hypothetical cohorts

CO-DE -- Coale-Demeny Formula= $(P_3)^2/P_2$,
BRASS -- Brass Formula = $(P_2).(P_4/P_3)^4$
P_7 -- Reported MNCEB for women 45-49,
$(P/F)_1$ -- Using P_2/F_2
$(P/F)_2$ -- Using $1/2(P_2/F_2+P_3/F_3)$,
$(P/F)_3$ -- Using $1/3(P_2/F_2+P_3/F_3+P_4/F_4)$
$(P/F)_4$ -- Using $1/2(P_3/F_3+P_4/F_4)$,
RGM(P) -- Brass RGM using P_i/P_{i+1},
RGM(F) -- Brass RGM Using F_i/F_{i+1}
UNP -- United Nations Projections

Table 4.14 (continued)
The Values of P/F Ratios and Estimates of TFR by Various methods

	WEST AFRICA						
	Ghana			Gam*	Cam*	Nig*	
	71	78	HYP*	73	78	78	
Age group	P/F Ratios						
15-19	1.09	.76	.76	1.28	1.18	.50	
20-24	1.19	.98	.98	1.29	1.06	1.06	
25-29	1.16	.98	.97	1.22	.99	1.03	
30-34	1.19	1.01	.97	1.20	.99	.99	
35-39	1.15	1.07	1.00	1.20	.95	.96	
40-44	1.13	1.06	.96	1.09	.91	.87	
45-49	1.10	1.08	.93	1.07	.86	.93	
Estimates of Total Fertility Rate by Coale-Demeny and Brass							
CO-DE *	6.1	5.3	5.2	5.5	5.1	5.8	
BRASS *	7.9	7.0	6.1	6.2	6.6	5.9	
P_7*	6.4	6.7	5.8	5.4	5.1	5.8	
Estimates of Total Fertility Rate by P/F Ratio Method							
$(P/F)_1$*	7.0	6.2	6.2	6.4	6.3	6.7	
$(P/F)_2$*	6.9	6.2	6.1	6.3	6.1	6.6	
$(P/F)_3$*	6.9	6.2	6.1	6.2	6.0	6.5	
$(P/F)_4$*	6.9	6.3	6.1	6.1	5.9	6.4	
Estimates of Total Fertility Rate by Brass RGM							
RGM(P)*	6.9	6.0	5.5	5.9	6.1	5.6	
RGM(F)*	6.8	5.9	5.9	5.9	5.9	6.1	
UNP	6.7	6.7	-	6.4	-	6.9	-

* Notes: Gam-- Gambia, Cam -- Cameroon, Nig -- Nigeria
HYP-- Using Hypothetical cohorts
CO-DE -- Coale-Demeny Formula= $(P_3)^2/P_2$,
BRASS -- Brass Formula=$(P_2).(P_4/P_3)^4$,
P_7 -- Reported MNCEB for women 45-49,
$(P/F)_1$ -- Using P_2/F_2,
$(P/F)_2$ -- Using $1/2(P_2/F_2+P_3/F_3)$,
$(P/F)_3$ -- Using $1/3(P_2/F_2+P_3/F_3+P_4/F_4)$,
$(P/F)_4$ -- Using $1/2(P_3/F_3+P_4/F_4)$,
RGM(P) -- Brass RGM using P_i/P_{i+1}
RGM(F) -- Brass RGM Using F_i/F_{i+1},
UNP -- United Nations Projections

104. Four types of adjustment factors based on the P/F ratio for the middle age groups are also shown in the Table 4.14. They are

$P_2/F_2, 1/2(P_2/F_2+P_3/F_3), 1/3(P_2/F_2+P_3/F_3+P_4/F_4)$

and $1/2(P_3/F_3+P_4/F_4)$. In general TFR estimates show a declining trend as we proceed from the first type of adjustment to the fourth type. For Libya(1973), Ghana(1971), Ghana(1978), Kenya(1969), Gambia(1973), Cameroon(1978), Tanzania(1969), Ethiopia(1970), Malawi(1972), Swaziland(1976), Lesotho(1971) and Botswana(1971) the estimated values are close to each other.

105. Towards the bottom of the table, the estimated TFR based on the fitting of RGM are shown. In RGM(P) the model is fitted to parity data alone whereas in RGM(F) the model is fitted to current fertility data and then TFR is derived. There are various approaches to the fitting of the model depending upon the age groups used and the estimation procedure applied. The values of TFR shown in the table are those obtained by fitting the model to the age range 20-35 by the least squares method.The estimated TFR derived by the two types give different results in most of the cases. In the method RGM(F) the age pattern of current fertility data is used to derive the model parameters and the parity data are adjusted by this pattern, and therefore this is similar to the P/F method. In RGM(P) the age pattern is deduced from the parity data itself.

106. Let' us first consider the cases that showed better P/F ratios. In the case of Kenya(1969) the RGM gave an estimate of TFR which is on the low side whereas P/F adjustment tends to give a value of 8.2, which is believed to be near to the truth. For Kenya(1978) the TFR estimated by P/F method and RGM(P) give very close results. But the estimate based on RGM(F) is on the low side possibly due to the effect of the reference period error being variable by age. The most plausible estimate for Kenya around 1969 and 1978 is about 8. The United Nations Projections (UNP) used a value of 8.1 which is very close to those obtained by P/F ratio method. The estimates of TFR for Libya in 1973 look to be slightly overestimated by P/F ratio adjustment, and also by RGM(F). The true value may be slightly lower than 8, and about 7.5 which is closer to RGM(P). The UNP estimate of 7.4 is close to this. For Ghana(1971) the estimates of TFR obtained by using P/F and that of RGM are close to 7.0 which is slightly higher than the estimate of UNP of 6.7. A value of 7.0 looks most probable. For Ghana(1978) the estimated TFR by P/F ratio method and RGM are quite close and indicate an estimate of 6.0 which is lower than the estimate for 1971. If

we treat this estimate as correct then we have to accept a reduction of 1 birth in the TFR during the seven years 1971-78, which looks implausible. The estimate for 1978 looks to be an underestimate and it appears that both the lifetime data and current fertility data are affected by errors. This is an interesting example of how data may be internally consistent but the results suffer from errors. For Lesotho (1971) the estimated total fertility by the P/F ratio method gives a value of 5.6, the RGM(F) gives 5.4 and UNP give a value of 5.4 and all values are quite close to each other.The estimate by RGM(P) of 5 is on the low side. For Cameroon (1978), the estimate of TFR by P/F method and the RGM give values close to 6.

107. Among the countries with the mean P/F ratios between 1 to 1.2 Ethiopia(1970) gives a TFR quite close to 5.5 by both P/F method and RGM, but the UNP gives a value of 6.7. It appears that both the lifetime and current fertility data may be suffering from omission of children as a result of which the estimated TFR is low. For Malawi(1972,1977), Swaziland(1976), Botswana(1971), Tanzania(1973), Kenya(1969) and Ghana (1978) it was found that the TFR estimated by RGM tend to give values lower than those estimated by P/F ratio method. In some cases the UNP estimates tend to lie between these two estimates but somewhat closer to RGM estimates.

108. By and large the estimates of total fertility obtained by P/F ratio method and RGM enclose the true estimates of TFR between them.

109. Table 4.14 also presents the results for hypothetical cohorts for six countries having data on MNCEB and BLY at two points of time. Let us look at the countries one by one. Kenya has requisite data for 1969 and 1978, the first one a census and the second one a survey. For the computations the intersurvey period is treated as ten years. The P/F ratios obtained for the hypothetical cohorts experiencing intersurvey cumulative and current fertility rates are found to be very close to that of the recent survey of 1978. Similarly the total fertility rate estimated for hypothetical cohorts are close to that of the recent survey. In the case of Swaziland and Lesotho also the estimates based on hypothetical cohorts are reasonable and give estimates close to the recent survey. This implies that errors in the two surveys are small or the errors in the surveys are such that they do not affect the intersurvey parity increments of fertility. In the case of Ghana the same pattern is observed, and the estimated TFR for hypothetical cohorts based on P/F ratio and RGM(F) are close to 6. It appears that the reported MNCEB and BLY in 1978 suffer from underreporting, a point mentioned earlier. In the case of the Sudan the data from the 1978 survey indicate errors, consequently the P/F ratios and the estimated TFR for hypothetical cohorts are found to give erratic and high estimates by various methods.

110. It is interesting to note that the three countries that gave reasonable values for TFR are based on surveys seperated by about ten years and the three countries that gave poor results have surveys seperated by 5 years. This could be accidental, it is also possible that, since hypothetical cohorts are obtained by using inter survey parity increments which in their turn are obtained by differencing parity data, a ten-year intersurvey period may produce a better balance of errors than a five year period. More study is needed of the influence of errors on the estimates based on hypothetical cohorts in Africa.

(b) Estimates of birth rates based on two age-sex distribution

111. Estimates of birth rates by methods that assume stable conditions are not shown here, as the recent methods which are described in the earlier sections are more general in that they do not require the stability assumptions and they give robust estimates of birth rates. In the earlier studies (Venkatacharya and Tesfay Teklu, 1986) it was observed that the birth rate estimates by Manual IV methods using l_2, or l_3 or l_5 gave values lower than the expected true values.

112. From table 4.15 it is clear that the birth rate estimated by reverse projecting 0-4 population is lower than the value obtained by reverse projecting 5-9 population. This could be explained by the omission of children under 5 and shifting of some children from the age group 0-4 to 5-9 due to age reporting errors. This is true of all the nine countries under study. The reverse projection of the population aged 10-14 gives birth rates that are lower than those obtained by reverse surviving the 5-9 age group for all countries excepting Kenya, Lesotho and Swaziland. This is due to some shifting of persons of the age group 10-14 to the adjacent age groups due to reporting errors. The mean value of the three birth rates obtained by reverse surviving 0-4, 5-9, 10-14 give values (column 5) which appear close to the expected true birth rates.

113. As errors in age reporting are reduced by cumulation, the reverse survival of population under 10 and 15 are expected to yield better results. The birth rates obtained by reverse surviving the population under 10 and 15 are shown in columns 6 and 7. For most of the countries the birth rate obtained by these two methods are close to each other. For Liberia, Kenya, Tanzania, Lesotho, Swaziland and Libya the birth rate based on population under 15 is slighlty higher than that obtained on population under 10. But the differences are small.

114. The birth rate obtained by the reverse survival of the population under 15 seems to give very robust estimates agreeing with the findings of earlier studies (United Nations, 1983, Coale, 1981). However, the comparison of birth rate obtained by this method, with other estimates in columns (2) to (6) should take into account the fact that the estimated birth

Table 4.15: The estimates of birth rates obtained by various methods for nine African countries having two censuses.

Country	Birth rate by reverse survival of age group						Birth rates by Coale derived methods			Birth rates on generalized stable equation		
	0-4	5-9	10-14	Mean	0-9	0-14 b_R	b_C	b_V	b'_V		b_n	b''_n
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Ghana	47.0	50.6	45.9	47.8	48.9	40.3	49.7	50.4	50.0	47.2	48.1	55.1
Liberia	39.3	49.7	44.6	44.5	44.3	45.2	44.8	45.2	44.9	41.7	42.3	63.0
Sierra Leone	45.9	58.5	42.5	49.0	52.0	50.6	50.2	50.5	50.1	47.6	48.1	65.8
Kenya	45.4	50.3	51.6	49.1	48.1	49.9	49.2	49.9	49.7	48.0	48.4	55.6
Malawi	53.4	54.6	46.7	51.6	54.1	53.8	53.2	53.7	53.3	53.3	52.4	56.7
Tanzania	46.5	53.0	49.0	49.5	49.9	50.6	49.9	50.1	50.0	48.7	49.0	56.9
Lesotho	33.7	35.8	41.2	36.9	34.9	37.1	36.9	37.6	37.9	36.9	36.5	42.0
Swaziland	43.2	48.8	50.1	47.4	46.6	48.1	47.4	48.1	47.7	47.1	47.0	53.2
Libya	52.1	58.1	52.1	54.1	55.4	55.6	54.5	55.4	55.3	58.3	57.8	64.1

rates in all these columns do not refer to the same time. If t denotes the census year then columns (2), (3) and (4) give birth rates that refer to (t-2.5), (t-7.5) and (t-12.5) years respectively. On the other hand the birth rates in column (6) refer to year (t-5). The birth rates in columns (7) to (13) all refer to year (t-7.5). If the birth rates are assumed to be constant during the fifteen years prior to t, for all the countries under study then all these birth rates will be comparable. If fertility fluctuated during this period the comparison of birth rates between the various columns should take into account the time referencing.

115. Comparison of birth rates by reverse surviving persons under age 15 (column 7) with those obtained by Coale method (column 8) shows that the two birth rates are almost equal for all the countries in the table. This is not surprising as it has been shown earlier methematically that these two are almost identical.

116. The birth rate obtained by the short cut method, or simplified robust estimate shown in column (9) and (10) also show consistency with each other. In column (9) the coefficients for u and v in equation (28) are taken from North Family in Table 4.12 whereas in column (10) the values of u and v are for the West Model. The specification of Model Life Table family did not result in any great change in the estimated birth rate.

117. It is also interesting to note that birth rates in column (10) and (7) are remarkably close. Further, supporting the theoretical explanation given earlier, the birth rates in columns (7) to (10) are all nearly equal for all the countries under study. Thus the birth rate obtained by reverse surviving the population under 15 or the birth rate obtained through the use of Model Stable and Life Table Populations (Coale method) give similar results.

118. The birth rates obtained in columns (11), (12) and (13) are based on the recent equations developed by Preston and Coale which are applicable to non-stable and stable populations. The data used in this case are identical to the data used in reverse survival and Coale methods (columns 7-10). The birth rates in columns (11) and (12) are obtained by using the population under age 15, the intercensal rate of growth for age groups 0-4, 5-9 and 10-14; and a life table that has an l_5 value estimated by Brass type indirect estimate. The birth rates in column (13) are obtained by using the age distribution of the two censuses and Brass type indirect estimate of l_5, Thus the difference in the results shown in columns (13) compared to columns (11) and (12) lies in the fact that the same equations fitted to the same basic data in two

different ways. Comparison of birth rates in columns (11) and (12) indicate very close agreement. Comparison of these values with those in column (8)- (10) indicates that the former birth rates are generally higher than the latter values. We notice that the two methods use the same c(15), and l_5. But in columns (7) to (10) the over all population rate of growth is used, whereas in columns (11) to (12) the rate of growth for the age groups 0-4, 5-9 and 10-14 are used. Therefore we can attribute this slight underestimation of the birth rate by the use of generalised stable equations to the errors in the estimated age specific rates of growth. The birth rates in column (13) which are based on Preston integrated method, are on the high side compared to all other estimates in the table.

119. However, there are situations where this tendency to overestimate is likely to compensate for the general understatement of birth rate, usually found in the developing countries. But, it is difficult to know this in advance. The reported age -sex data for Sierra Leone and Liberia have larger errors than other countries and it is especially in these cases the birth rate by Preston integrated method goes out of range. In the case of Libya in addition to minor errors in the data, the omission of net migration in the estimation process leads to higher birth rates.

120. A study of the figures in table 4.15 indicates that the birth rates by all the methods except the Preston integrated method give by and large values closer to each other and very close to the true values. In some cases, due probably to under reporting of population under 15 , the birth rates tended to be slightly on the lower side (e.g. Kenya). This clearly indicates that the use of C(15) instead of the entire age distribution gives birth rate estimates which are more stable in poor data situations whether the equations used are based on stable or generalized stable populations.

121. From all the various estimates presented in the table, the values obtained in columns (7) to (10) look close to the true values. Thus the estimation of birth rate by reverse surviving the population under age 15 is a dependable method of estimating birth rate. The simplified estimation procedure given in equations (28) and (29) is very useful in this regard, as it does not involve the use of Model Stable or Life Table Populations.

122. However, the estimation by this procedure, is likely to be an underestimate of birth rate, in some instances due to the omission of children aged 0-4, like in the case of Kenya. In such situations, any information on the magnitude of this error can be helpful in improving the estimates by making end corrections.

REFERENCES

Abou Gamrah
Methods of Adjusting the Stable Estimates of Fertility for Effects of Mortality Decline. Population Studies. 30(1):165-177, 1976.

Blacker,J.G.C
The Estimation of Vital rates from Census Data in Kenya and Uganda. In: (Ed) Cantrelle et al. Population in African Development. Ordina Editions,Belgium, 1979.

Booth,H
Transforming Gompertz Function for Fertility Analysis: The Development of a Standard for the Relational Gompertz Function. Population Studies. 38(3):495-506, 1984.

Brass,W
Screening procedures for Detecting Errors in Maternity History Data. WFS Occasional Papers. 22:30-49, 1980.

Brass, W. et al
The Demography of Tropical Africa. Princeton University Press. Princeton.U.S.A, 1967.

Brass, W
The Use of Relational Gompertz Model to estimate Fertility. Proceedings of the IUSSP Conference. Manila. 3:345-362, 1981.

Caldwell,J.C
The Study of Fertility and Fertility Change in Tropical Africa. WFS.Occasional Papers. 7, 1974.

Cantrelle,P
Is There a Standard Pattern of Tropical Mortality? In: P.Cantelle et al (Ed). Population in African Development. Ordina Publications.Belgium, 1974.

Chidambaram,V.C and T.Pullum
Estimating Fertility Trends from Retrospective Birth Histories: Sensitivity to Imputation of Missing Dates. Population Studies. 35(2):307-320, 1981.

Coale, A.J and Paul Demeny
Regional Model Life Tables and Stable Populations. Princeton University Press. Princeton .U.S.A, 1966.

Coale,A.J and T.Trussell
Model Fertility Schedules : Variations in the Structure of Childbearing in Human Populations. Population Index. 40(2):185-258, 1974.

Coale,A.J
Robust Estimation of Fertility by The Use of Model Stable Populations. Asian and Pacific Census Forum.8(2):5-7, 1981.

Garenne,M.
The Age Pattern of Infant and Child Mortality in Ngayokheme (Rural West Africa). African Demography Working Paper No.9. Population studies Center. University of Pennsylvania 1981.

Gibril,M.A
Some Reporting Errors in the 1973 Gambian census (Unpublished). M.sc.Dissertation. London School of Hygiene and Tropical Medicine, 1976.

Hobcraft, J.N. et al
Advances in the P/F ratio Method for the Analysis of Birth Histories. Population Studies. 36(2):291-316.1982.

Luther,N.Y
Measuring changes in Census Coverage in Asia. Asian and Pacific Census Forum 9(3):1-11, 1983.

O'Muircheartaigh, C.A. and A.M. Marckwardt
1980. An Assessment of the Reliability of WFS Data. WFS Conference, 1980. London.315-380.

O'Muircheartaigh, C.A
Methodology of the Response Errors Project. WFS Scientific Reports.28.1982.

Page,H
Fertility Levels: patterns and Trends. In (Ed). J.C.Caldwell . Population Gowth and Socio-Economic Change in West Africa. The Population Council. New York.29-57, 1975.

Potter,J.E
The Validity of Measuring Change in Fertility by Analysing Birth Histories Obtained in Surveys. Ph.D.Thesis. Microfilms. Ann Arbor, 1975.

Preston,S.H and A.J Coale
Age Structure,Growth,Attrition and Accession: a New Synthesis. Population Index 48:217-259, 1982.

Preston,S.H
An Integrated System for Demographic Estimation from Two Age Distributions. Demography. 20(2):213-226, 1983.

Ramachandran,K.V
Errors and Deficiencies in Basic Demographic Data: Overview of Methods of Detection, Evaluation and Adjustment. Proceedings of the Workshop on the Estimation of Fertility and Mortality. (Mimeographed), 1983.

Samoza, J.L
An Evaluation of the Analysis of Defective Data. Proceedings of the IUSSP Coference. Manila.3:375-394, 1981.

Som, R.K
Recall Lapse in Demographic Enquiries. Asia Publishing House.Bombay, 1973.

United Nations
Methods of Estimating Basic Demographic Measures from Incomplete Data.(Manual IV).Sales No.E.67.XIII.2, 1967.

United Nations
The Concept of a Stable Population: etc.
ST/SOA/ Series A/39. New York, 1968.

United Nations
Model Life Tables for Developing Countries.
Sales No. ST/ESA/SER.A/77. New York, 1982.

United Nations
Indirect Techniques for Demographic Estimation (Manual X). Sales no. e.83.XIII.2.New York, 1983.

Venkatacharya, K and Tesfay Teklu
Estimation of Some Indices of Mortality and Fertility from Successive Censuses and Surveys With reference to Sub-Saharan Africa. Proceedings of IUSSP Conference, Florence.(4):541-55, 1985.

Venkatacharya, K and Tesfay Teklu
Problems and Solutions in The Estimation of Fertility in Africa. In: Report of the Regional Workshop on Demographic Estimates and Projections in Africa. ECA/RIPS 1:91-128, 1986.

Venkatacharya, K and Tesfay Teklu
On Some Robust estimates of Birth Rate Under non-Stable Conditions. Demography. 24(4):639-648, 1987a.

Venkatacharya, K and Tesfay Teklu
A Simplified Robust Estimate of the Birth Rate. Asian nad Pacific Population Forum. 1(4): 9-23, 1987b.

Venkatacharya, K and Tesfay Teklu
Reverse Survival Methods of Estimating Birth Rates Under Non-Stable Conditions. Genus,44(1-2):73-97, 1988.

Zachariah,K.C
An Alternative Approach to the Problem of Adjusting Birth Rates Estimated by the Stable Population Method for the Effects of Declines in Mortality. Demographic Measures and Population Growth in Arab Countries.Cairo Demographic Centre Research Monograph Series. No.1:327-34, 1970.

Zlotnik, H and K.Hill
The Use of Hypothetical Cohorts in Estimating Demographic Parameters Under Changing Fertility and Mortality. Demography.18(1):103-122, 1981.

CHAPTER 5

POPULATION PROJECTIONS

5.1 INTRODUCTION

1. In the earlier sections, attention has been focused on evaluation and adjustment of population data as well as on various methods of estimating vital rates from available data. These deal with information on the current situation as at the time of the census or survey. However, in practice, the task before the demographer goes beyond that of providing information about the current situation. Since planning is concerned with the future, policy makers and planners require, as of necessity, reasonable projections, which provide information about the future size, structure and distribution of population. Here, we will therefore deal with demographic projections.

5.1.1 The Role and Importance of Demographic Projections

2. The emergence of newly independent states in the African continent during the late 1950's and early 1960's stimulated among national governments the formulation of development plans designed to overcome poverty and under-development. The objective that featured in most of these plans was to satisfy the current and future needs of the population in the areas of health, education, housing, employment, etc, using the available human and material resources. As human resources and needs vary according to the size, composition, and distribution of the population, demographic data, not only for the current period but also estimates for the future, have become the essential tool and the fundamental background for development planning.

3. The tasks of defining goals and the selection and determination of available means to achieve them are made quite simple by translating them into quantitative terms. Where, for example, one of the principal objectives of the plan is to increase per capita income, there will be the need to know the future size of the population during the plan period so as to determine the magnitude of national output that can ensure an increase in per capita income in the future. Population projections also provide information about the size of the potential labour force that will be available for achieving the desired objectives. Similarly, calculating prospective food needs based on nutritional requirements of different population groups cannot be undertaken without a projection of the size and age-sex composition of the total population.

4. The increased awareness among national governments of the need to integrate population variables in planning and the subsequent shift of emphasis in planning strategy to satisfaction of the basic needs of the population has added new dimensions to the importance and role of demograpic

projections. Basic needs planning require projections of the releveant target population groups, who in the planning process constitute consumers on one hand and producers of goods and services on the other.

5. Efficient planning of future investments by both government and enterprises requires not only urban and rural population projections but also considerable detailed projections at sub-national levels. Planning in the educational sector involves development of adequate teacher training programmes, construction of schools and colleges, and provision of classrooms, books, etc. The quality of decision in these areas is extremely dependent on the quality of the available demographic projections of the target population, i.e, the school-age population at various levels. In the estimation of future housing needs, household or family projections give an indication of the new dwelling units that has to be provided during the plan period. Projection of the socio-economic sub-groups of the population thus enables planners to pursue a greater number of objectives, consider a larger number of policy options and render planning responsive to the needs of the population (Pyatt and Thorbecke, 1976).

6. In view of the fact that population censuses and surveys are costly affairs, they cannot be undertaken at every time and for every detail required by development planners. Indeed, the majority of African countries find it extremely difficult to keep to the schedule of one census every ten years. At the same time, birth and death registration in these countries are still too incomplete in coverage and defective in quality to provide any basis for regular estimation of population parameters. In consequence, population data required for the assessment of the situation in the base or inter-census years of the plan period are obtained, in effect, by means of projections.

7. Quite apart from their use in planning, demographic projections shed light on demographic processes and present the net results of interactions between the components of population change, namely - the interrelations between the levels of mortality, fertility and migration over time and their net effect on growth rate and age-sex composition of the population. This constitutes a very important aspect of demographic analysis. Benchmark demographic projections can be used to illustrate what could happen if observed levels and trends in one of the components of population change were to persist.

8. The current high rate of population growth against a background of a depressed economic condition has made it difficult for developing countries to meet the needs of their growing population for more employment, education, housing and other facilities. The problem of both international and rural-urban migration is also a matter for serious concern in these countries. Consequently, population policy-making is gradually being institutionalized as an integral part of development planning. In the formulation and implementation of

population policies, demographic projections play distinctive roles. They provide best possible estimates of what the future population will be. They can also be used to sound a warning of where present trends could lead, unless something is done to avert the situation. In this connection, the projected figures are usually considered in conjunction with other information on urbanization, on population density, on avialable land and on agricultural and industrial development. Demographic projections also serve the purpose of highlighting the sensitivity of vital rates to an on-going or a proposed population policy-oriented programme (Keyfitz, 1984).

5.2 TYPES OF DEMOGRAPHIC PROJECTIONS

9. Demographic projections can be undertaken at different levels and for different segments of the population, for example, at a global or regional level, at national and sub-national levels, or for sectors and population sub-groups within a country. The increased awareness of the pressure of a rapidly growing population on limited agricultural and industrial resources stimulated the demand for projections of future population trends at a global level. Such world and regional population projections are usually undertaken by international organisations. The United Nations (U.N.), for example, has since 1950 published global projection figures under the title "World Population Prospects". The latest of these projections, the 1984 assessment, gives projections of population by sex and age for the world, the more developed and the less developed regions, and for continents and countries (U.N, 1987).

10. Originally these global estimates served the purpose of illustrating the implications of the unprecedented growth in world population which was brought about by a decline in mortality accompanied by little change in the high level of fertility. However, because of the lack of expertise, deficiency of existing population data and the inability to conduct population censuses at regular intervals, a number of developing countries of Africa now use the results of UN global projections for planning purposes.

11. National population projections deal with the future size and structure of the population of residents or usual residents of a country. The components of population change taken into account in national population projections include fertility, mortality and international migration. Most national development plans and population policies are formulated on the basis of the results of national projections. Population projections for political units of the country such as districts, provinces or counties, for urban and rural areas as well as for individual cities, groups of cities and localities are regarded as sub-national population projections. For such sub-groups, changes in the size and structure of the population

are affected not only by fertility, mortality and international migration, but also by internal migration, reclassification of urban and rural areas and groups of localities as well as by boundary changes.

12. In planning for basic needs of the population, it is also essential to obtain projected population by sectors or socio-economic groups. Such specialized projections include those of the labour force, agricultural population, school enrolment and educational attainment and household and families. In addition to demographic components of population change, departure from and entry into the socio-economic group, through losing or acquiring the group's definitional charactersitics, affect their sizes and structures. Specialized sectoral projections therefore require additional information regarding labour force participation rates, school enrolment rates, grade progression rates, household headship rates, etc.

5.3 TECHNIQUES OF POPULATION PROJECTION

13. A number of methods have been developed for assessing future population size and structure. These methods differ from one another in the different circumstances in which they are applied. The determination of the method to be used in a particular situation depends on the quality and details of available population data and the purpose which the estimates are expected to serve. Two main methods, the mathematical and the component method of demographic projections are of more general application. However, other hybrid methods, such as the ratio and economic methods used in special circumstances and for special purposes will be briefly discussed.

Mathematical Projection Method:

14. Early attempts at population projection were mathematical by nature. The method assumes that a mathematical expression can be fitted to the observed counts of population in the past, and from such a fitted expression, future size of the population can be extrapolated. In cases where at least two censuses have been taken and there are no vital registers, the method can easily be used for projecting for future size of the population.

15. Mathematical projections are based on assumed laws of population growth. Where the size of population is postulated as a function of time, for example, a polynomial function can be used to derive the projected population at time t years after the base period, i.e.

$$P_t = a + bt + ct^2 + dt^3 + \dots \qquad \dots(1)$$

The values of the constants a,b,c,d, ... are usually estimated from past census results by an appropriate technique, such as

the least square method. As early as 1891 Pritchett fitted a polynomial of the third degree to the 1790 to 1880 United States census results and applied this method in projecting the population of United States beyond 1880 (Pritchett, 1891).

16. Among the most commonly used mathematical formulae for preparing demographic projections are those representing geometric curves

$$P_t = P_o (1+r)^t \qquad ...(2)$$

and

$$P_t = P_o e^{rt} \qquad ...(3)$$

where r is the observed rate of growth of population, t, the number of years after the base year and P_o, the size of the base year population. In this case, the increase of population is expressed as a function of the size of the population and the rate of growth of population is assumed to be constant over time. Over a long period of time, population projection by this method can produce impossibly large numbers. As no population with limited space can be expected to grow indefinitely at a constant rate, these formulae are recommended for use only in making short term projections.

17. The attempt to find a mathematical expression that adequately describes the laws of population growth over a long period of time led to the development of the logistic curve

$$P_t = \frac{K}{1 + \exp(a+bt)} \qquad ...(4)$$

where t is the number of years, a and b are constants. K is the upper asymptote of the curve. Fitting a logistic curve requires a large number of observations covering a long period in order to determine values of the constants. The property of having an upper limit ensures that the logistic curve, when used for projecting the size of the population, does not yield absurd results in the long run. Although the method was originally developed for projecting the total size of populations, the logistic curve is finding wider application in other areas of demographic projections, such as urban and rural and city population projections.

18. The merit of the mathematical projection method lies in its simplicity and the fact that it requires relatively fewer data. It can also be applied in the projection of the components of population change, for example, in projecting mortality instead of total numbers. However, it has its own limitations when used to project total population. The method does not show the resulting changes in population structure, such as distribution by sex and age. Furthermore,the size and

structure of the population at any point in time are influenced by social and economic forces. The mathematical projection method carries with it the implicit assumption that the social and economic factors which influenced population growth in the past will continue in exactly the same way into the future. It is therefore not recommended for populations subject to rapid changes in the components of population change. Generally speaking, the longer the span of the projection period, the more unreliable will be the results of projection by the mathematical method.

The Ratio Method

19. The ratio method of population projection is specific to the projection of the population of a sub-area of the national population. As its name implies, it merely provides projected sets of ratios which, when multiplied by already projected national figures, yield the projected population of the sub-area. Series of total population figure for a sub-area can be extrapolated into the future by a mathematical formula as a ratio of an already projected national population. The method is probably the most convenient for projecting the population of regions, divisions, states, counties, and districts, especially if the total population of the country has already been projected by other methods.

20. One way of projecting the sub-area to total population ratio is to hold the recently observed percentage distribution constant throughout the period of projection. The other is to modify the ratios in some way according to the past trends or even according to expected impact of regional development programmes. In populations where the tempo of urbanization and volume of internal migration are high, the holding constant of recently observed ratios may not be justified. As the percentage of sub-area to national population cannot decrease or increase indefinitely, it is necessary to ensure that in the modification of the observed trends for the future, the percentage distributions approach a stable state in the long run.

21. The general functional relationship between the sub-area and the national population can also be examined by regressing the sub-area population against the national populations especially if data exist for a large number of periods in the past. With the fitted regression equation, the sub-area projected population can easily be determined from the already projected national population. However, the problem of the application of this method to African populations lies with the extremely few number of censuses conducted in the past (less than three for most countries). Regression equation for ratios derived on the basis of such few number of observations can be very misleading in determining the future trends in the growth of sub-area populations.

Economic Method

22. Population growth cannot be entirely separated from changing economic conditions. A number of studies have shown that internal movement of population is influenced to a large extent by the changing economic advantages of one area over the other (Lowry, 1966). In populations where economic opportunities exert an over riding influence on population change through migration, population projection can be undertaken in such a way that it takes into account economic variables.

23. The major types of economic variables considered in such projections include, school enrolment, per capita income, relative changes in per capita income, and production. Provided adequate data exist on these variables over a long period, regression analysis can be used to relate net gain in migration to changes in one or more of the economic indicators. The Stanford Research Institute used this method in projecting net migration for various states in the United States, 1965-80 (Badestra and Rao, 1964).

24. In areas where new industries are sited, new mines operated or a new capital city is established, population projection can lend itself to the economic method. In this case the new establishments or projects are first related to expected labour force growth, or employment opportunities. Net migration resulting from the growth of labour force is calculated and projected population can be obtained after taking into account births and deaths.

25. The application of the economic method for population projections requires not only detailed data on the various economic variables but also accurate measurement of type of relationship that exists for example, between labour force or employment opportunity and net migration. There will also be the need to make some assumptions about what the future trend in the relationships will be during the projection period.

Component Method of Projection

26. The component method of population projection involves projecting separately the three components that are responsible for population change, namely, mortality, fertility and migration. The projected components are then applied to the base population in order to obtain projected population at future points in time. The method is also so called because it consists in the separate projections of numbers of males and females in each age group of the population. The projected total population is obtained as a sum of the separately projected segments of the population.

27. With the increasing availability of detailed data on age distribution and vital rates of the population obtained from censuses, surveys and vital registers, the component method has become the most widely applied technique of population projection in recent years. In contrast to the mathematical method, it attempts to take proper account of the components of population change and also of available information as to their trends in the past. On the basis of the observed trends, the soico-economic changes and the population policies expected to take place over the projection period, the assumption about the future course of the components can be made.

28. In the application of the component method, it is most convenient to project the population by time intervals that are equal to the width of the age intervals in which the population has been classified. In this way, at the end of each interval of projection, all surviving members of one age group will have moved into the subsequent age groups.

29. Population projection by the component method involves a sequence of computational steps. In the case where the base population is classified by sex and five-year age-groups, the population structure is modified over each five-year interval to allow for assumed mortality and migration conditions. In addition, the number of births reflecting the fertility assumptions is computed for each interval and further modified to allow for mortality and migration. The process yields the age-sex structure at the end of successive intervals.

30. The component projection method yields results that are indispensable to any planning exercise seeking to take future population change into account or designed to satisfy basic needs of certain segments of the population. These features make the component method particularly useful in the integration of population variables in development planning. It is important to note that the method merely provides an estimate of what the future size and structure of the population will be _if_ the components of population change take the course specified in the assumptions.

5.4 METHODOLOGY FOR NATIONAL POPULATION PROJECTION BY THE COMPONENT METHOD

5.4.1 Inputs Required for National Population Projection by the Component Method

31. One of the most important inputs required for the purposes of population projection by the component method is data on the age-sex structure of the population from the most recent population census. Experience has shown that in most developing countries- especially those of Africa- available population census data are defective due to under-enumeration, over-enumeration and/or age mistatement. They should therefore be first evaluated and where the nature and magnitude of the

defects are suspected to be serious, attempt should be made to adjust the data before any preliminary projection is undertaken. Where possible, the adjustments should remove errors of coverage and age reporting. To carry out the adjustments, it will be necessary to utilize the techniques already dealt with in the earlier sections.

32. The base year populaion age-sex structure used for component projection should pertain to the mid-point of the year. However, only in very rare cases are censuses taken in the middle of the year. Furthermore, in most instances, the most recent census data usually precedes the projection base year by a few if not several years. Consequently deriving the base year population age-sex structure as an input may require a preliminary projection to bring the most recent census data up to the mid-point of the base year. An alternative procedure would be to project the census year age-sex structure in five year intervals until immediately beyond the base year. The age-structure of population at the mid-point of the base year is then obtained by interpolating between the intervals just before and after the base year. This type of preliminary projection will ofcourse require appropriate assumptions regarding fertility, mortality and migration or the age structure and growth rate of total population.

33. A second type of input required for projecting population by the component method consists of detailed information on the past trends, current levels and patterns of vital rates. Various methods of estimating these rates have been dealt with in the earlier sections and will therefore not be discussed here. However, the type of indices of past and current levels of vital rates required as an input will be treated. For population projection by the component method, fertility input should preferably be specified in terms of gross reproduction rate or total fertility rate along with the age pattern of fertility rates. Similarly, past and current levels of mortality are most convenient for component projection if they are prepared in the form of expectations of life and age-sex specific mortality rates or survival rates. Where the nature of data permits, international and internal migration data presented in terms of annual or quinquennial net gains/losses should be preferred. The value of the sex ratio at birth is also needed to enable the splitting of total births into male and female births. The utilization of existing information on recent trends and current levels of mortality, fertility and migration makes the formulation of assumptions about their future trends a less hazardous exercise.

34. The third type of input required for population projection by the component method is a set of assumptions about what the future trends in the components of populaion change are expected to be. For the formulation of such assumptions, it would normally be necessary to take into account anticipated socio-economic changes and the impact of any population policies during the projection period.

In the case of fertility assumptions, for example, due consideration should be given to any expected changes in marriage patterns, breastfeeding habits, laws regarding abortion or sterilisation, availability of family planning methods and effectiveness of contraceptive, as studies have shown that changes in these intermediate variables are the most important fore-runners of fertility change (Chidambaram, 1984). Similarly, in making assumptions about the future course of mortality, it would be normally necessary to consider likely changes in factors affecting mortality and morbidity, such as environmental conditions, access to and quality of health services, household income, and female education. Formulating assumptions regarding future trends in migration is a very difficult and speculative exercise. Where international migration constitutes a significant part of population change, the assumptions should reflect government policies, particularly if international migration is subject to considerable government control. In countries where considerable international labour migration takes place, it may be necessary to take into account domestic labour requirements, domestic labour force and implied need for foreign workers. However, changes in the direction and volume of flow of international migration are highly unpredictable.

5.4.2 Assumptions and Techniques of Projecting Vital Rates

35. As mentioned ealier, population projection by the component method involves separate projection of the three components of population change (fertility, mortality and migration) and the application of the projected vital rates on the base year population age-sex structure to obtain the projected population. The projections, however, rest on the premise that orderly progress will be made and that catastrophes, such as wars, famines and epidemics will not occur during the projection period.

36. In projecting mortality as a component of population change, two indices of mortality may be employed, namely, age specific death rates and age specific survival rates. The method of projecting age-pattern of mortality for any specific

country depends on the type, extent and quality of data available at the time of projection. In general, such as age specific mortality rates or survival rates, can be projected by extrapolation using appropriate mathematical formulae; by reference to model life tables or to an optimal life table attainable under ideal conditions; and by cause of death.

37. Where sufficient data on mortality exist over a long period of time, life table functions can be projected by the use of a mathematical formula. One such formula is represented by the equation:

$$q_{x,n} = a_x + b_x r_x^n, \quad (0 < r < 1) \qquad ...(5)$$

where $q_{x,n}$ is the probability of dying at age x in the year n, b_x reflects the level of mortality at age x at a particular point in time, r_x takes account of the improvement in mortality at age x over time, while a_x represents the ultimate level of mortality at age x ($a_x, b_x > 0$).

38. The ultimate level of mortality, a_x, can be estimated with reference to its value in countries that have attained the lowest level of mortality such as Norway and Sweden, where as the values of b_x and r_x are obtained from trends in mortality in the country of interest. For example, the improvement factor, r_x can be estimated for various recent years by using the following between the years t and t+n:

$$r_x = (q_{x,t+n} / q_{x,t})^{1/n} \qquad ...(6)$$

After comparing the estimated r_x at neighbouring ages, a reasonable value of r_x is chosen for extrapolation purposes and the projected probability of death at age x, is then calculated using equation (5).

39. A graphical approach can also be used in projecting for each selected age, where the observed probability of dying at age x in the year m, $q_{x,m}$ is plotted agianst m, preferrably in a semi-log graph paper. A snooth curve is drawn representing the observed trend in $q_{x,m}$ over the years, the trend is then extrapolated graphically to yield the projected probability of dying at age x in the year n, $q_{x,n}$. With the graph of neighbouring ages drawn on the same sheet, trends at neighbouring ages can be taken into account.

40. Although the mathematical graphical method of projecting mortality may be simple in its application, a lot is left to the discretion of the demographer especially with regard to the final choice of the improvement factor, r_x. Besides, the method takes no direct account of other relevant information which may be available to the demeographer, such as the trends in mortality by cause and the suspected reasons for these trends.

41. A number of developed countries with detailed reliable mortality data covering several decades have been known to apply these extrapolation methods (e.g. United Kingdom, Canada, Australia, etc) However, their application in the African region is fraught with problems mainly because of the paucity of reliable mortality data. As far as is known these methods of mortality projection have not been applied in African countries.

42. A second method of mortality projection is by reference to a law of mortality. A few of such laws dealing with adult mortality include that of Gompertz:

$$\mu_x = Bc^x \qquad \ldots(7)$$

where the force of mortality. μ_x is defined as the negative differential of log $(1-q_x)$:
or one due to Makeham:

$$\mu_x = A + H_x + Bc^x \qquad \ldots(8)$$

and one due to Barnett of the form:

$$q_x/P_x = A - H_x + Bc^x \qquad \ldots(9)$$

In each of these laws $C > 1$ and is approximately 1.08. Recently Heligman and Pollard (1980) proposed a formula which takes into account child and adult mortality:

$$q_x/P_x = A^{(x+B)^C} + D \exp(-E(\ln(x/F))^2) + GH^x \qquad \ldots(10)$$

The first term reflects early childhood mortality, the second term accident mortality at the younger adult ages and the third a term for old age mortality.

43. A Campbell of the US Bureau of the Census gave a mortality rate projection method based on the formula:

$$m_{x,t} = (m_{x,0} - m_{x,\infty}) \exp(b_x t) + m_{x,\infty} ,$$

where $m_{x,t}$ is the projected value of mortality rate at age x and time t,

$m_{x,0}$ is the base mortality rate at age x,

$m_{x,\infty}$ is the expected value of mortality rate in the remote future and

b_x is a negative constant depending only on age x.

Campbell gave a set of $m_{x,\infty}$ values. The values of b_x are obtained from past trend of m_x values of the country. He also suggested that $m_{x,\infty}$ can be taken as one half of values already attained by lowest mortality countries like Norway of Sweden.

44. When the projection of age specific mortality is attempted with reference to any one of these laws, the curve representing the law is fitted to the observed age specific mortality rates at each of several different epochs (e.g. by least square or maximum likelihood). Trends in the parameters are then extrapolated to provide estimates of parameters at future periods of time. Projected age specific mortality rates are obtained by substituting the values of the paramaters at the various ages into the formula describing the law.

45. Examples of the application of the "law" of mortality of the projection of age specific mortality rates are few. Cramer and Wold applied Makeham law exensively to Swedish national mortality data (Cramer and Wold, 1935). Heligman and Pollard fitted their eight parameter curve to Australian mortality data at three epochs 1947, 1961, 1971 and found that the parameters seemed to exhibit an apparent trend over time and could be used for mortality projections.

46. Although mortality projection by reference to a 'law' of mortality has some theoretical appeal, it may be difficult to find a suitable law. Even when a suitable law can be found, it may be applicable only to the data at certain epochs. It is also possible that the variations in the parameters may not follow a discernable pattern over time. More pertinent in the case of developing countries of Africa is the fact that the method cannot be applied when mortality data are non-existent or available only for a single epoch.

47. Age speciic mortality rates can be projected by reference to model life tables such as Coale and Demeny (1966), Brass (1971), and United Nations (1982). These model life tables are particularly useful for estimating complete life tables or abridged life tables and for projecting mortality from limited data. The systems of model life tables differ from one another with respect to the number of parameters involved.

48. The projection of mortality by reference to model life tables involves first, the choice of the system of life table that suits the mortality data of the country of interest best. For example, in the case of Coale and Demeny system of model life tables, with a single parameter, (the life expectancy at

age X,(e_x^o), it will be necessary to select the appropriate model which will continue to represent the mortality pattern of the population under consideration (e.g. Coale and Demeny North Model tables). The value of the parameter e_x^o is measured or observed at each year or epoch. Any trend in the parameter is extrapolated by mathematical formula to provide estimates at future epochs. Projected age specific mortality rates are obtained by interpolating, where necessary, between model life tables.

49. In African countries where data may not exist over a long period for the determination of the trend in age pattern of mortality, the Coale and Demeny North model life tables, which have been found to yield consistent results with most sub-Saharan African population data can be used. Assumptions regarding future gains in life expectancy should be made taking into consideration the level of longetivity already attained and the possible effect of any mortality and morbidity reducing programmes underway.

50. The assumption in the United Nations 1978 projections that quinquennial gains in life expectancy at birth, e_o^o, tend to approximate to between 2 and 2.5 years when e_o^o is less than 55 years and to slow down thereafter, does not necessarily have to hold for all populations. For a number of countries, non-decreasing gains in life expectancy at birth, or gains exceeding 2.5 years may prevail beyond the 55 year cut off level.

51. A table indicating alternative models of assumptions about future gains in female life expectancy at birth as prepared by the United Nations Economic Commission for Africa (UNECA) in 1986 is presented in Table 5.1 below. It will be seen that the alternative models I, II and III differ in the assumed rate of improvement in mortality conditions: rapid, moderately rapid and slow improvements in mortality conditions. The assumed improvement in mortality conditions approach zero as life expectancy at birth rises beyond 80 years. Where data on past trends in mortality decline are lacking, current mortality conditions can be used as a guide in choosing one of the alternatives or the values for the purpose of mortality projections.

Table 5.1 Alternative Models of Mortality Decline in Terms of Annual Increment in Female Life Expectancy at Birth

Current $\overset{o}{e}_0$ in years	Average annual increase in $\overset{o}{e}_0$ in years		
	I	II	III
Less than 60	0.60	0.50	0.40
60-62.5	0.60	0.50	0.40
62.5 - 65.0	0.60	0.50	0.40
65.0 -67.5	0.52	0.50	0.40
67.5 - 70	0.44	0.44	0.40
70 - 72.5	0.34	0.34	0.34
72.5 - 75	0.25	0.25	0.25
75 - 77.5	0.20	0.20	0.20
77.5 - 80	0.14	0.14	0.14
80 - 82.5	0.04	0.04	0.04

Source: Economic Commission for Africa, Report of the Regional Training Workshop on Demographic Estimates and Projections in Africa. Vol. 1, Addis Ababa, 1986, Table 52(b).

52. Other methods of projecting mortality include projection by reference to an "optimal" life table attainable under ideal conditions and projections by cause of death. In the former technique, a suitable optimal life table attainable under ideal condition is selected from those developed by other researchers, taking into account optimal improvement for each cause of death. A decision is taken as to how the population will approach the optimal mortality-schedule and how quickly it wil do so by the use of a mathematical formula. The projected mortality at epoch n is then calculated.

53. The method of projection by cause of death involves calculaing age-specific mortality rates by cause of death at several recent epochs for selected ages. The age specific mortality rates by cause are projected separately using any one of the mathematical or graphical approaches described earlier. The projected age specific mortality rates by cause of death are combined to obtain the projected mortality rate at the selected ages.

54. The choice of method for the projection of mortaility depends to a large extent on the data available and their reliability, the resources available for the projection and on the purpose for which projection is undertaken. The cause of death approach is appropriate, if reliable and detailed mortality data are available over a long period. In some countries of Africa, resonably reliable estimate of current mortality rates are still lacking and even in those areas where they exist, the data may refer to one point or at most two points in time. Consequently estimates about past trends are based on conjectures. Statistics available on causes of death refer usually only to those admitted in hospitals. In the circumstances, most mortality projections in African countries are undertaken by reference to model life tables. Age pattern of probabilities of surviving corresponding to different levels of life expectancy at birth are chosen, and where necessary modified according to any existing reliable data.

55. Assumptions regarding the future trends in fertility require reasoned judgement. In preparing such assumptions, past and current fertility trends in each country are carefully studied, but generally, the determination of future trends are guided by the logic of demographic transition theory as well as by the expected success or failure of any family planning programmes in the country under consideration. The logic of demogarphic transition rests on the premise that the experiences of West European and North American countries with regard to decline in fertility will be duplicated in present high fertility areas of the world. Social and economic development will bring with it a lowering of the level of fertility. However, the problem with the application of this logic lies on the timing of the onset of decline and in the determination of the rate of decline (ECA, 1986). In countries where fertility rates are already low, the transition from high to low fertility took between 80 years in Western European and North American countries to less than 50 years in East Asia and some smaller countries of Latin America and Africa.

56. With regard to family planning programmes, most African countries have accepted the need for population policies designed to reduce fertility. A survey conducted by the United Nations in 1980 revealed that at least 12 African countries regarded their fertility rate as too high and had designed policies to reduce the rate. In the circumstances, projected trend of the level of fertility should be that of a decline at some future point in time, with the onset and the tempo of

decline determined by the effectiveness of family planning programmes and envisaged changes in the socio-cultural conditions in the country under consideration.

57. For component projection purposes, the projected trend in fertility should preferably be given in the form of gross reproduction rates and age pattern of fertility or age specific fertility rates. Unlike mortality which has an almost universal tendency to decline, fertility neither tends to its possible maximum or minimum level. In recognition of the fact that several possibilities for the future trends in fertility appear to be equally plausible, United Nations projections of fertility are carried out under three or more variants. The ranges among the high, medium and low variants are based on assumed length of time needed for the gross reproduction rate to reach 1.0 and the timing of decline. The variants correspond to the high, medium and low levels of population growth.

58. In Table 5.2, for example, the United Nations' projected gross reproduction rates for West Africa for the period 1975 to 2025 are presented. They indicate a constant gross reproduction rate (GRR) for the period 1975-1985 for each of the three variants, and thereafter there is a slow decline for the medium variant and a much faster decline for the low variant.

Table 5.2 Projected Gross Reproduction Rate for West Africa 1975-2025

Period	Gross Reproduction Rate		
	High	Medium	Low
1975-1980	3.33	3.33	3.33
1980-1985	3.33	3.33	3.33
1985-1990	3.33	3.24	2.77
1990-1995	3.26	3.09	2.18
1995-2000	3.16	2.89	1.86
2000-2005	3.02	2.64	1.63
2005-2010	2.85	2.37	1.48
2010-2015	2.65	2.10	1.35
2015-2020	2.44	1.83	1.26
2020-2025	2.22	1.56	1.29

Source: United Nations: Population Projections, Methodology of the United Nations, New York 1984. Table 2.

59. In projecting fertility for individual countries due regard should be given to any anticipated or possible changes in age pattern of fertility. Future improvement in fertility projections will be made if sufficient data are available for incorporating intermediate variables in fertility projections, and for predicting future birth expectations.

Techniques of Projecting Migration

60. For most countries, net migration is quite negligible and for projection purposes it is usually assumed to have an insignificant effect on population change. However, for a number of countries, especially the small ones, net migration may contribute significantly to population change. In such cases projection of migration should be attempted.

61. The greatest obstacle to the projection of migration is the fact that migration is affected to a large extent by social, economic and political circumstances in countries of origin and destination and predicting the future trends in these factors is a particularly precarious task. Furthermore, in most countries of the world, especially those of Africa, the current level of migration is unknown. Consequently predicting future migration on the basis of past and current levels constitutes a conjectural exercise.

62. One of the methods used to project migartion directly is the mathematical method. Migration can be related to time or to some socio-economic variable, such as labour force or employment opportunities, industrial growth etc. The mathematical method assumes that there exists a high correlation between proportional changes in net migration and changes in the socio-economic variables. In most African countries the detailed data required for studying these relationship are at present lacking.

63. Indirect methods can however be used to project migration in countries where data are available on population distribution by age and sex and place of birth for two consecutive censuses. The census cohort-change rates obtained from the results of the censuses can be used to project cohorts already alive at the later census, provided the two censuses are comparable in scope and coverage and assuming that such cohort change rates remain constant or can be suitably modified on the basis of current information. In this way both mortality and migration projection are taken care of at the same time. However, the method does not allow for calculation of cohort-change rates for those born during the intercensal period.

5.5 ILLUSTRATIVE EXAMPLE OF NATIONAL POPULATION PROJECTION BY THE COMPONENT METHOD

A Summary of Procedures:

64. The first step in national population projection by the cohort-component method is the compilation and evaluation of the available data on age-sex structure of the population and on the levels and trends of fertility, mortality and, where necessary, migration. Appropriate adjustments are made so that the measure of vital rates appear reasonable and age structure is internally consistent and compatible with estimated vital rates. Preliminary projection may have to be carried out in order to bring the most recent census data forward to the mid-year of the base period.

65. The next step is to derive (5 year) age specific survival rates from the mortality assumptions. In the case where use is made of model life tables, interpolation for survival ratios may be made for projected levels of mortality that fall within the printed levels.

66. Projected age and sex-specific five year survival ratios are successively applied to the age sex distribution of population at the beginning of the interval to obtain survivors aged 5 years and over at the end of the interval. For example, the numbers alive aged x to x+4 at the initial time t, $_5N_x^t$

are projected as urvivors aged x+5 to x+9 at time t+5, $_5N_{x+5}^{t+5}$ as follows:

$$_5N_x^t \; ({_5L_{x+5}}/{_5L_x}) = {_5N_{x+5}^{t+5}}, \quad x = 0,5,\ldots,w-5$$

where w is the highest given age taken as a multiple of 5, $_5L_x$ is the usual life table population whose value is chosen from the appropriate model life table representing the level of mortality of the country in the interval of time t to t+5. Similarly, the numbers aged w years and over at time t,

$_5N_{w+}^t$ are projected as survivors aged w+5 years and over at time t+5 ie.

$$_5N_{w+}^t (T_{(w+5)+}/T_{w+}) = {_5N_{(w+5)+}^{t+5}}$$

Using projected net international migration rates, if necessary, net changes among survivors aged 5 years and over due to migration are calculated. The number of survivors are modified by these changes.

67. Projected age specific fertility rates are then applied to the average number of females in each of reproductive age groups and the sum of the products is multiplied by 5 to obtain total number of births during the five year interval. In symbols,

$$B(t, t+5) = 5\sum_{i=15}^{49} 0.5\,[P(i, t) + P(i, t+5)]\,F(i) \quad \ldots\ldots(11)$$

where $B(t, t+5)$ = total births between time t and t+5
$F(i)$ = age specific fertility rates at age i
i = 15-19, 20-24, ..., 45-49.
$P(i,t)$ = female population aged i at time t
$P(i, t+5)$ = female population aged i at time t+5

68. The total number of births, B(t, t+5) in the 5 year interval is distributed according to an assumed sex ratio at birth. The numbers of births for each sex is then multiplied by the appropriate survival ratio of births in order to obtain the number of survivors at the end of the interval. The net balance of migration is added or subtracted from the surviving births of each age and sex to obtain population aged 0-4 at the end of the interval.

69. For projection to the end of subsequent five year periods, t+10, t+15 etc, the procedure outlined above is repeated using the already projected population at time t+5, t+10 etc respectively as the base and applying the assumed fertility, mortality and migration parameters as well as the sex ratio at birth for the respective periods.

Application of the Procedure in the Projection of the Population of The Republic of Sierra Leone, 1975-1995

70. The last published census of the Republic of Sierra Leone was conducted in 1974. An analysis of the data on age and sex distribution of the population indicated a joint score of 50.3 for the country. Similarly the data on age specific fertility rates were found defective. Brass' techniques of local fitting by the logit method was used in adjusting the age and sex data while the age specific fertility rates were equally adjusted using Brass' P/F ratio method. The adjusted age-sex distribution of the population and adjusted age specific female birth rates are presented in Table 5.3.

71. It is required to project the population for the period 1975-1990. As the most recent published census in this respect was taken in December 1974, there was the need to bring the census data forward by six months to mid-year 1975, the base year of the projection period. In order to do this, the official average annual growth rate of 2.1 percent was used to

bring the data forward to mid-year 1975, on the assumption that the age-sex structure of the population could not have changed much within six months. The resulting base year age distribution is presented in Table 5.3.

72. By comparing the reported age distribution of Sierra Leone, 1974 with those of Coale and Demeny Model life tables, it was found that North Model level 9 provided the closest mathematical fit to the age structure. On the basis of this finding, male and female life expectancy of birth was chosen as 38.9 and 41.4 respectively to represent the mortality level in the five year period before mid 1975. From then on life expectancy at birth was projected to increase by 0.5 year per annum as presented in Table 5.4. With regard to fertility projection, the observed age pattern was assumed to remain constant through out the projection period. A sex ratio, at birth of 1.03 was assumed. The projected age specific female birth rates under the three variants are as presented in Table 5.4.

73. During the 1974 census of Sierra Leone, the total number of foreign nationals enumerated came to 79,414 or 2.9% of the total population. As there was no evidence of any substantial emigration from Sierra Leone during the period, it was assumed that at the national level, international migration did not constitute a significant component of population growth. Consequently the effect of net international migration on future population size and age structure was assumed to be negligible. The enumerated population of nationals and non-nationals together was carried forward in the projection. No projections were therefore made for net international migration.

74. The assumptions of fertility and mortality made above were fed into the projection programme and the results for the medium variant are presented in Table 5.5.

5.6 ILLUSTRATIVE PROJECTIONS OF SUB-NATIONAL POPULATIONS

75. A number of methods described earlier can be used for projecting sub-national populations. These methods include mathematical method, the ratio method and some other miscellaneous methods, such as economic method and the holding capacity method. The sub-national population referred to in this context include those of regions, districts, provinces or geographical divisions of the country as well as urban/rural and city populations. The choice of method for sub-national projections depends to a large extent on the nature and quality of available data and the purpose which the projection is expected to serve.

Table 5.3 Adjusted Age-Sex Distribution and Adjusted Age-Specific Female Birth Rates, Sierra Leone 1974/1975

Age group	Adjusted Pop. Dec. 8, 1974 Male	Adjusted Pop. Dec. 8, 1974 Female	Age Specific Female Birth Rates	Base Year Population Mid Year 1975 Male	Base Year Population Mid Year 1975 Female
0-4	229,127	228,019		231,520	230,401
5-9	176,167	176,783		178,007	178,629
10-14	151,998	152,701		153,585	154,296
15-19	135,334	135,633	0.104	136,747	137,049
20-24	121,531	119,841	0.146	122,800	121,093
25-29	105,570	103,474	0.138	106,673	104,555
30-34	90,645	89,482	0.103	91,592	90,417
35-39	77,551	76,922	0.083	78,361	77,725
40-44	65,971	66,223	0.040	66,660	66,915
45-49	54,996	56,459	0.026	55,570	57,049
50-54	44,957	47,569		45,427	48,066
55-59	36,336	39,656		36,716	40,070
60-64	24,986	28,358		25,247	28,654
65-69	19,390	22,724		19,593	22,961
70-74	13,271	16,343		13,410	16,514
75-79	7,519	9,929		7,598	10,033
80+	3,972	5,722		4,013	5,782
TOTALS	1,359,321	1,375,838		1,373,519	1,390,209
	2,735,159			2,763,728	

Table 5.4 Projected Mortality and Fertility Schedule, Sierra Leone,1975-90

Year	Gross Reproduction Rate (GRR) under the three variants		Life Expectancy at birth (in years) Male	Female
	High	3.2		
1975-80	Medium	3.2	41.4	43.9
	Low	3.2		
	High	3.3		
1980-85	Medium	3.2	43.9	46.4
	Low	2.9		
	Hige	3.3		
1985-90	Medium	3.2	46.4	48.9
	Low	2.6		

Table 5.5 Projected Population of Sierra Leone, 1980, 1985 and 1990
Using the Medium Variant Assumption

Age	1980		1985		1990	
Group	Females	Males	Females	Males	Females	Males
00-04	287,809	289,891	334,780	338,077	381,008	385,853
05-09	208,036	208,674	262,919	264,441	309,364	312,180
10-14	171,715	171,239	200,842	201,574	254,894	256,481
15-19	150,020	149,329	167,426	166,932	196,333	196,970
20-24	132,826	131,873	145,823	144,451	163,167	161,905
25-29	116,769	117,527	128,509	126,684	141,499	139,207
30-34	100,230	101,823	112,379	112,634	124,112	121,821
35-39	86,113	87,012	95,898	9[illegible],[illegible]65	107,968	107,890
40-44	73,574	73,788	81,922	82,367	91,638	92,399
45-49	62,998	61,993	69,626	69,055	77,879	77,515
50-54	53,159	50,786	59,031	57,072	65,562	63,984
55-59	43,837	40,453	48,822	45,640	54,551	51,681
60-64	35,178	31,381	38,859	34,985	43,650	39,847
65-69	23,588	20,187	29,355	25,476	32,813	28,749
70-74	16,974	14,041	17,860	14,781	22,599	18,971
75+	24,265	15,500	20,840	14,975	20,549	15,614
T O T A L	1,587,091	1,565,497	1,814,891	1,796,309	2,087,586	2,071,067

Projection of Sub-divisions of a country:

76. Considering the nature and quality of data available for projection of sub-national populations in Africa, especially with regard to internal migration which constitutes a major component of population change in sub-national populations, the ratio method is probably one of the most satisfactory procedures for projecting the population of sub-divisions of a country, that is, assuming that the country's total population has already been projected by other methods. The ratio method is based on the relationship between the population of sub-areas and that of the nation as a whole. The ratios of sub-area population to that of the whole country are calculated for the dates for which census data are available. An attempt is made to project the trend in the ratios to future dates on the basis of certain assumptions

77. The inputs required include:

(i) Data for at least two censuses that are comparable in scope and coverage and show the distribution of the country's total population by sub-areas.

(ii) A projected population by sex and age of the whole country covering the dates for which the sub-area population projections are required.

78. In the case where data exist for two censuses on the distribution of the total country population by sub-areas, the projection of the sub-area population involves a number of steps. First, the ratio of the sub-area to total country population is calculated for each sub-area at the two census dates. The average annual rate of change in these ratios over the inter-censal period is calculated using, for example, the geometric growth formula. The derived average annual rate of change in the ratios is assigned to the midpoint of the inter-censal period.

79. The next step involves the formulation of assumptions regarding the future trends in the rates of change of observed sub-area ratios, ie. projecting the observed rates of change of sub-area ratios to future dates. Over the intercensal period, the observed ratios may have increased, decreased or remained constant. Where majority of the sub-area ratios have remained constant or shown very little change, the sub-area ratio may be held constant at the last observed level, especially if the projection is for a short period of time. In the cases where data for more than two censuses are available, the general methods are given in the ECA training workshop on demographic estimates and projections (ECA, 1986).

80. On the basis of the assumption made, the rates of change of the ratios for the different intervals of projection are obtained for each sub-area and by applying these rates of change to the base year ratios, the projected ratios for the sub-areas are obtained. In the cases where the projected ratios for all the sub-areas do not sum up to unity, the ratios are prorated to add up to 1.

81. It may be further required to distribute the projected sub-area total population by sex and age. To do this, the projected population by sex and age and regions is obtained by applying the difference elimination method to the two sexes separately (ECA, 1986).

Illustrative example with the data for Sierra Leone

82. The procedure for projecting sub-national population by the use of the ratio method is illustrated for the 13 districts of Sierra Leone (Tables 5.6 and 5.7). The ratios of the district population to that of the whole country in April 1963 and December 1974 are as shown in columns 2 and 3 of Table 5.6. The observed average annual percentage rate of change between 1963 and 1974 in the sub-area ratios, are calculated and presented in column 4. It is important to note that the observed rates of change are assigned to the mid-point of the intercensal period in this case approximately around February 1969. On the assumption that the observed intercensal percentage rates of change of the ratios will fall linearly to zero in 50 years time, the projected rates of change of the sub-area ratios for the periods 1974-5, 1975-80, 1980-85, and 1985-90 are calculated and shown in columns 5 to 8. On the basis of the results, the sub-area ratios observed in December 1974 were extrapolated to mid 1975, 1980, 1985 and 1990. For Bo district, for example, the ratio of Bo district to that of the whole country population was 0.0962 in April 1963 and 0.0796 in December 1974. The observed annual percentage rate of change of the ratio in the period 1963-74, r_{63-74} is derived from the equation:

$$0.0962(1+r_{63-74})^{11.67} = 0.0796$$

which yields an r_{63-74} value of -0.0161 or -1.61 percent. If the annual rate of change has to fall to zero in 50 years then annual fall is 1.61/50 = .0322. From December 1974 to mid 1975 ie. 6 months, the fall in the rate of change is 0.0161.
The Ratio of Bo District population to that of the whole country by mid 1975, R_{75} can be obtained as

$$R_{75} = R_{74}\ (1-0.0161)^{0.5}$$
$$= 0.0796(1-0.0161)^{0.5} = 0.0790$$

Similarly $R_{80} = 0.790(1-0.0124)^{5} = 0.0742$

These results are shown in the preliminary projected ratios in Table 5.7. It will be seen from the table that sum of the ratios of all the districts do not add up to one. A pro-rata adjustments is carried out to ensure that the ratios add up to one as shown in Table 5.8. The projected total population of Bo district is derived by multiplying the ratios shown in Table 5.8 by the projected total population of the whole country in Table 5.5.

Table 5.6 Projection of the Population of Sierra Leone By the Ratio Method from 1974 to 1990

District	Ratio of District Pop. to National Pop. April 1963	Dec. 1974	Observed Average Annual percentage rate of change 1963-1974	Projected avearge annual percentage change of the ratios 1974-75	1975-80	1980-85	1985-90
Bo	.0962	.0796	-1.61	-1.58	-1.24	-1.08	-0.92
B/Sherbro	.0368	.0320	-0.19	-1.04	-0.92	-0.80	-0.78
Moyamba	.0768	.0690	-0.91	-0.79	-0.70	-0.61	-0.52
Pujehun	.0389	.0376	0.29	-0.25	-0.22	-0.19	-0.17
Kailahun	.0689	.0659	-0.38	-0.33	-0.29	-0.26	-0.22
Kenema	.1043	.0975	-3.58	-0.51	-0.45	-0.39	-0.33
Kono	.0770	.1203	3.89	3.39	3.00	2.61	2.23
Bombali	.0912	.0854	-0.56	-0.49	-0.43	-0.38	-0.32
Kambia	.0632	.0568	-0.91	-0.79	-0.70	-0.61	-0.52
Koinadugu	.0592	.0581	-0.16	-0.14	-0.12	-0.11	-0.09
PortLoko	.1135	.1068	-0.52	-0.45	-0.40	-0.35	-0.30
Tonkolili	.0846	.0754	-0.98	-0.85	-0.76	-0.66	-0.56
W/Area	.0894	.1156	2.23	1.94	1.72	1.50	1.27
Total	1.0000	1.0000					

Table 5.7 Preliminary Projected Ratios of the District to National Population, Sierra Leone, 1975 to 1990

District	Preliminary Projected Ratios 1975	1980	1985	1990
Bo	0.0790	0.0742	0.0703	0.0671
Bonthe & Sherbro	0.0318	0.0304	0.0292	0.0281
Moyamba	0.0687	0.0664	0.0644	0.0627
Pujehun	0.0376	0.0371	0.0368	0.0365
Kailahun	0.0658	0.0647	0.0639	0.0632
Kenama	0.0973	0.0951	0.0932	0.0917
Kono	0.1223	0.1418	0.1613	0.1801
Bombali	0.0852	0.0834	0.0818	0.0805
Kambia	0.0566	0.0546	0.0530	0.0516
Koinadugu	0.0581	0.0577	0.0546	0.0543
Port Loko	0.1066	0.1044	0.1026	0.1011
Tonkolili	0.0751	0.0723	0.0699	0.0680
Western Area	0.1167	0.1271	0.1369	0.1458
Totals	1.0008	1.0092	1.0179	1.0307

Table 5.8 Projected Ratios of District to National Population, Sierra Leone, 1975 to 1990

District	Projected Ratios 1975	1980	1985	1990
Bo	0.0789	0.0735	0.0691	0.0651
Bonthe & Sherbro	0.0318	0.0301	0.0287	0.0273
Moyamba	0.0686	0.0658	0.0633	0.0608
Pujehun	0.0376	0.0368	0.0362	0.0354
Kailahun	0.0658	0.0641	0.0628	0.0613
Kenama	0.0972	0.0942	0.0916	0.0890
Kono	0.1222	0.1405	0.1584	0.1747
Bombali	0.0851	0.0826	0.0804	0.0781
Kambia	0.0566	0.0541	0.0521	0.0501
Koinadugu	0.0581	0.0572	0.0536	0.0527
Port Loko	0.1065	0.1035	0.1007	0.0981
Tonkolili	0.0750	0.0716	0.0687	0.0660
Western Area	0.1166	0.1260	0.1344	0.1414
Totals	1.0000	1.0000	1.0000	1.0000

Urban Rural Projection

83. Another type of sub-national population projection which is indispensable in development planning is the projection of urban and rural populations. It provides information about the levels and trends of urbanization and thus enables researchers and policy makers to anticipate some problems that are connected with urbanization and growth of large urban centres. The ratio method, the mathematical method and some miscellaneous methods can be used in projecting urban and rural populations. However, the nature of data available often dictates the type of method to be used. Mortality and fertility data in African countries are often not classified by urban and rural residence. Furthermore, the poor quality of migration data makes it extremely difficult to apply the component methods to urban rural projections.

84. The historical experience of industrial nations with regard to the process of urbanization has exhibited a fairly regular pattern. At the initial stage, when the economy is principally agrarian, the proportion of urban to total population is small and urban population grows at a slow rate. When industrialization takes off, the proportion of urban to total population increases rapidly for some time until it reaches a high level. Thereafter, the growth rate of the proportion slows down and eventually declines to zero. Because of the analogy of this growth pattern to that of the logistic curve, it was considered realistic to project urban rural percentâge by the use of the logistic curve.

85. The appropriateness of the logistic curve for projection of urban proportion becomes evident when one considers the index of urbanization, ie. the urban-rural ratio (URR). The rate of change in URR measures the tempo of urbanization. Symbolically, if we denote U(t) and R(t) as the urban, and rural population at time t respectively, then, using the exponential growth formula, urban-rural ratio at time t+n, URR(t+n) can be expressed as:

$$URR(t+n) = URR(t)\ \exp(rn) \qquad ...(12)$$

where r represents the rate of growth of URR or the tempo of urbanization.

From equation (12), it can be shown that the tempo of urbanization,

$$r = \frac{1}{n} \ln \frac{URR(t+n)}{URR(t)} \qquad ...(13)$$

Since Urban Rural ratio at time t, URR(t) = U(t)/R(t)

$$r = \frac{1}{n}\left[\ln \frac{U(t+n)}{R(t+n)} - \ln \frac{U(t)}{R(t)}\right] \qquad \ldots(14)$$

Re-arranging equation (14), yields

$$r = (1/n)[(\ln U(t+n)/U(t) - \ln (R(t+n)/R(t)] \qquad \ldots(15)$$

which is simply the difference between urban and rural growth rates for the n year period from time t years to t+n years. In other words, the tempo of urbanization is equal to the difference between urban and rural growth rates.

86. In view of the fact that empirical observations have shown that urban-rural growth difference (URGD) more or less remains constant at least over a short period of time, the urban-rural growth difference method has found wide application among demographers as a technique for projecting urban and rural populations. Assuming that the URGD remains constant over the period of projection from equation (12) as follows, the urban-rural ratio at time t+n is projected as follows:

$$URR(t+n) = URR(t)\exp(URGD\,.n) = URR(t)\,e^{rn} \ldots(16)$$

Once the urban-rural ratio at time (t+n) is derived, the projected proportion of total population urban,

$$PU(t+n) = \frac{URR(t+n)}{1 + URR(t+n)} = \frac{URR(t)e^{rn}}{1 + URR\,(t)e^{rn}} \qquad (17)$$

which is the equation of the logistic curve

87. The United Nations has however noted that assumption of the constancy of URGD may not be justified, as empirical evidence has indicated that the difference declines over time. In order to compensate for this tendency, a method of modifying the URGD as urbanization increases has been developed (United Nations, 1980). Weights were introduced to adjust the urban growth rate difference in such a way that as the initial level of urbanization increases, the difference between the urban and rural growth rate diminishes. To do this, a hypothetical urban-rural difference, URGDH is introduced with URGDH = 0.044177-0.028274 (initial proportion urban). Weights, w are then given to enable an adjusted urban-rural growth difference URGD* to be calculated for the different projection periods, as follows:

$$URGD^* = w.URGD + (1-w)URGDH \qquad \ldots(18)$$

The values of w start with 0.8 in the first projection period

and decline linearly by 0.2 for each subsequent five year period (UN, 1984). The respective values of URGD* for each period are then introduced into equation (16) so that

$$URR(t+n) = URR(t) \exp(URGD^{*}.n) \qquad ...(19)$$

88. In cases where the pattern of urbanization is not depicted by the logistic curve, other types of curve like Gompertz curve,

$U_t = a^{b^t}$, may be fitted to the data at

two points in time and the values of the constants a and b determined. With these values, the urban population at any point in time can be projected using the Gompertz formula.

Illustrative Example of Urban-Rural Projection Using Sierra Leone Data

89. The inputs required for the projection of urban and rural populations are the distribution of population by urban and rural residence at two census dates and the projected total population at future points in time. In the case of Sierra Leone, the inputs required were obtained from the census results of 1963 and 1974 and the projected total population for the future periods. For illustration, the medium variant projected population of Sierra Leone 1985 (Table 5.5) will be used. As shown by the April 1963 and December 1974 censuses, the urban-rural population distribution were as follows:

Date	Total Population	Urban	Rural
April 1963	2,180,355	412,254	1,768,101
Dec. 1974	2,735,159	752,026	1,983,133

90. As a first step, the urban-rural ratio and the exponential growth rates for urban and rural populations between 1963 and 1974 are calculated and the resulting initial URGD is determined for the period 1963-1974 as follows:

$$\text{Urban growth rate} = \frac{1}{11.67} \ln \frac{(752,026)}{(412,254)} = 0.0515$$

$$\text{Rural growth rate} = \frac{1}{11.67} \ln \frac{(1,983,133)}{(1,768,101)} = 0.0098$$

The urban-rural ratio in 1974 = 752,026/1982133 = 0.3792
Unadjusted urban-rural growth difference = 0.0417

Hypothetical urban-rural growth difference, URGDH, is then calculated as

$$0.044177 - 0.028274 \times 0.2749 = 0.0364$$

where the proportion urban in 1974 = .2749.
Using equation (18) above and the appropriate weights, the adjusted urban rural growth difference, URGD* are obtained for the projection periods as follows:

Projected Period	w	1-w	URGD*
1975 -80	0.8	0.2	0.0406
1980 -85	0.6	0.4	0.0396
1985 -90	0.4	0.6	0.0385

Finally the urban-rural ratio and the proportion of total population urban are derived using equations (19) and (17). For example, the projected URR for 1975,

$$URR_{75} = URR_{74} \exp[(0.0417)\ (0.5)] \text{ using equation (19)}$$

$$= 0.3872$$

and the projected proportion urban by mid 1975, PU_{75}

$$= URR_{75}/(1+URR_{75}), \text{ using equation (17)}$$

$$= 0.2791$$

The urban-rural ratios, projected proportion urban for the rest of the projection periods are presented in columns 2 and 3 of Table 5.9.

91. The projected urban and rural population in Table 5.9 refers to total populations. In most cases, it may be further required to distribute the projected total urban and total rural populations by sex and age. This can be done by the method of difference elimination, using the data on age and sex distribution of urban and rural populations at the base year as well as the projected total country population by sex and age and the projected urban and rural totals as inputs (United Nations, 1974). The procedure is illustrated with Sierra Leone data, where the 1974 population census data is the base year data and the projected total urban and total rural populations for 1985 in Table 5.9 are to be distribution by age and sex.

Table 5.9 Projected Total Urban and Total Rural Populations, Sierra Leone 1975-95

Year	Projected URR(t)	Projected Proportion Urban PU(t)	Projected* Total Population	Projected Urban Population	Projected Rural Population
1975	0.3872	0.2791	2,763,728	771,356	1,992,372
1980	0.4770	0.3230	3,152,588	1,018,286	2,134,302
1985	0.5875	0.3700	3,611,200	1,336,144	2,275,056
1990	0.7237	0.4198	4,158,653	1,745,803	2,412,850

* Obtained from Tables 5.3 and 5.5.

92. The format for the computation requires arranging the data in a matrix form as shown in Table 5.10. The base year population data in panel A are classified by age and sex as well as by urban and rural residence. Panel B contains the projected urban and rural totals in 1985 as well as the projected total population of the country by age groups in 1985 (see Table 5.9 and Table 5.10). The difference elimination method described here is an application of a general matrix solution for adjusting the information in the cells of the matrix in such a way that it conforms to an independent set of vertical and horizontal totals.

93. The first step in the computation involves prorating the figures in each row of the base data (Panel A) in the ratio of the marginal totals of panel B to that of panel A. The results are entered in panel C as a first approximation. Thus, in panel C, the number 213,055 entered for urban males aged 0-14 is obtained by multiplying 147,661 by the ratio of 804,096 to 557,292. It will be seen that the results in panel C, when added up vertically, will not agree with the vertical totals of panel B (ie. the projected urban and rural totals). As a next step, the figures in each column in panel C are further prorated in the ratio of the marginal totals of panel B to that of panel C and the results are entered in panel D. For instance, the rural population of 205,768 female, aged 45 years and over entered in panels D is obtained as 236,869 times 2275056/26[illegible]8918. Again, when added up horizontally, the results in panel D wil differ from the horizontal totals of panel B. The procedure is repeated several times by prorating alternatingly in the horizontal and vertical directions until

Table 5.10 Application of the Method of Difference Elimination to a Projection of Urban and Rural Population, and of Total Population by Sex and Broad Age Groups, Example of Sierra Leone 1985

Panel A (Base Year Data 1974)

Sex	Age	Urban	Rural	Total
Males				
	0-14	147661	409631	557292
	15-44	203878	392724	596602
	45+	40026	165401	205427
Females				
	0-14	164062	393441	557503
	15-44	158506	433069	591575
	45+	37893	188867	226760
	Totals	752026	1,983133	2,735159

Panel B (Projected Totals 1985)

Sex	Age	Urban	Rural	Total
Males				
	0-14	...	...	804096
	15-44	...	...	730233
	45+	...	...	261980
Females				
	0-14	...	...	798541
	15-44	...	...	731957
	45+	...	...	284393
Totals		133614	4 2275056	3611200

Panel C (First horizontal Prorating)

Sex	Age	Urban	Rural	Total
Males				
	0-14	213055	591041	804096
	15-44	249544	480689	730233
	45+	51045	210935	261980
Females				
	0-14	234995	563546	746170
	15-44	196120	535837	726546
	45+	47524	236869	276548
Totals		992282	2618918	3611199

Panel D (First Veritical Prorating)

Sex	Age	Urban	Rural	Total
Males				
	0-14	286886	513438	800324
	15-44	336020	417575	753595
	45+	68733	183239	251973
Females				
	0-14	316429	489553	805982
	15-44	264083	465482	729565
	45+	63993	205768	269761
Totals		1336144	2275056	3611200

the discrepancies become negligible. The final age and sex distribution of the projected urban and rural population of Sierra Leone obtained by using procedure is presented in Table 5.11.

94. The recently published preliminary report of the December 1985 population census of Sierra Leone indicates that the proportion of urban to total population was 0.3180 as against 0.3700 projected in Table 5.9 (Sierra Leone Government, 1986). Thus, judging from the difference between the enumerated and projected proportion of total population urban, it would seem that during the period 1974 to 1985, urbanization may have been sluggish. The intercensal period 1963 to 1974 which was used as a basis for estimating the urban -rural growth difference was also the period soon after the attainment of independence. It was a period that witnessed the establishment of more new industries and various new organs of government at the urban centres of the country to the neglect of the rural areas. The creation of such new employment opportunities at the urban areas in the period 1963-1974 may have triggered off a high rate of rural to urban movement of population that could not be repeated in the period 1975-1985, a period marked by economic depression and rising umemployment in the urban areas.

95. This phenomenon probably lends support to the view that urban-rural growth difference may not be constant over time. Unfortunately most African countries have conducted less than three censuses. The projection of future trends on the basis of few past observations may be very misleading. Until more data are available over a long period of time, results obtained by extrapolation of most recent trends should be interpreted with a great deal of caution.

Table 5.11 Distribution of Projected Total Urban and Rural Populations of Sierra Leone, 1985 by Sex and Age Using the Difference Elimination Method

Sex and Age Group	Projected Population 1985		
	Urban	Rural	Total
Males			
0-14	289336	514760	804096
15-44	326674	403559	730233
45 and over	71772	190208	261980
Females			
0-14	314639	483902	798541
15-44	265953	466004	731957
45 and over	67770	216623	284393
Totals	1,336144	2,275056	3,611200

R E F E R E N C E S

Badestra, P. and Rao, W.R.,
Basic Economic Projections, United States Population, 1965-80, Stanford Research Institute, California. pp. 37-44 (1964).

Brass, W.,
"On the Scale of Mortality" in Brass, W. (ed.) Biological Aspects of Demography, London, Taylor and Francis (1971).

Coale, J. and Demeny, P.
Regional Model Life Table and Stable Population, Princeton, Princeton University Press. (1966).

Chidambaram, V.C
"Use of intermediate factors in fertility projections" in United Nations, Population Projections, Methodology of the United Nations, Population Studies No. 83, U.N. New York (1984).

Crammer, H. and Wold, H.
Mortality Variation in Sweden Scandinavian Actuarial Journal Vol. II pp. 161-241 (1935).

Dominion Bureau of Statistics,
Memorandom on the Projection of Population Statistics (84-D-69), Ottawa (1950).

Dominion Bureau of Statistics,
Memorandum on the Projection of Statistics (84-d-69A) Ottawa.

Economic Commission for Africa
Report of the Regional Training Workshops on Demographic Estimates and Projections in Africa, Vol. 1., Addis Ababa (1986).

Golulapati, R., De Ravin, J.W, and Trickett
"Projections of Australia Mortality Rates, Occasional Paper No. 1983/2 Canberra, Australia Bureau of Statistic (1984).

Heligman, L. and Pollard, J.H.,
The Age Pattern of Mortality in Journal of Institute of Vol. 107 pp. 49-80 (1980).

Keyfitz, N,
"Population Projection as an Aid to the Formulation of Population Policies", in Population Projection, Methodology of the United Nations, UN, New York. BT/ESA/SER.A/83 (1984).

Lowry, I.S.,
Migration and Metropolitan Growth, Two Analysis Models. San Francisco, California, Chandler Publishing Company (1966),.

Pollard, J.H.,
Projections of Age Specific Mortality Rates, Population Bulletin of the United Nations Nos. 21/22-1987 New York (1988).

Pritchett
A formula for Projecting the Population of the United States, Publications of the American Statistical Association 2:278 (1891),.

Pyatt, G. and Thorbecke
Planning Techniques for a better future, ILO, Geneva (1976).

Ramachandran K.V., and Gyepi-Garbrah, B.
Population Projections in Planning Sub-national and Regional Aspects, RIPS Newsletter No. 14, 1974, RIPS, Legon (1974).

Regional Institute for Population Studies (RIPS),
Internal Migration and Regional Development in Africa, African Population Dynamics, RIPS Monograph, No. 2, 1987.

Shyrock, H.S. and Siegel, J.S.
The Methods and Materials of Demography Condensed Edition by Stockwell, E.G. Academic Press, London (1976).

United Kingdom, Government Actuary's Department
Projecting the Population of the United Kingdom, Economic Trends (May, 1965) London, H.M. Stationery Office (1965).

United Nations
Manual VIII, Methods for Prejections of Urban and Rural Population, Population Studies, No. 55. ST/ESA/SER.A/55, United Nations, New York (1974).

United Nations
Patterns of Urban and Rural population Growth Population Studies no. 68, ST/ESA/SER.A/68., United Nations, New York (1980),.

United Nations
Methodology of the United Nations, Population Studies No. 83, United Nations New York (1982).

United Nations
Model Life Tables for Developing Countries United Nations, New York, (Sales No. 81 XIII.7) (1982).

United Nations
Global Estimates and Projections of Population by sex and Age, The 1984 Assessments, U.N. New York, ST/ESA/SEr. R/70 (1987).