

**THE HISTORY AND FUTURE OF AFRICAN CENSUS ANALYSIS
PROJECT (ACAP)**

Tukufu Zuberi and Martin W Bangha
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104

DRAFT

Paper prepared for presentation at the session “Censuses: New trends, new formulas.” The “Chaire Quetelet 2006” on *Information Systems in Demography and in Social Sciences, New Questions, New Tools!* Louvain-la-Neuve (Belgium), November 29-December 1, 2006.

THE HISTORY AND FUTURE OF AFRICAN CENSUS ANALYSIS PROJECT (ACAP)

Tukufu Zuberi and Martin Bangha

Abstract

The history of census taking in most of Africa is characterized by gross under-utilization of census data, the primary preoccupation having been the execution of the processes of enumeration, data processing, the production and publication of census tables and in some cases the preparation, production and dissemination of population projections. Very little attention has been paid to systematic evaluation of the quality of census data, conducting in-depth or further analyses of data, including spatial analysis, and disseminating data in format and media that meet the diverse needs of users. Consequently, census results have not adequately informed policy formulation, program implementation, and socio-economic development in general. Not only have census data been under-utilized, but until recently they were poorly stored, resulting in considerable loss in a number of countries. The few countries that have managed to preserve their data tend to have poor meta-data on census codes and definitions, which has limited the utility of their census data.

It was in recognition of the need to preserve Africa census data in order to avoid perpetual loss due to poor storage and also the need to encourage and enhance further analysis, dissemination and utilization of the massive census data that the African Census Analysis Project (ACAP) was initiated. In 1998, ACAP held its first African workshop in the Republic of South Africa, chaired by Professor Tukufu Zuberi, and hosted by Pali Lehohla the Focal Point for the SADC Population Census Project and Statistician General of the Republic of South Africa. Following this workshop, at the Southern African Development Community (SADC) meetings of the SADC Ministerial Committee held in Maputo, Mozambique in August 1999, a Memorandum of Understanding was established that formalized the collaboration between ACAP and SADC. This memorandum became the model of collaborative agreements between ACAP and over 30 other African governments and research institutions. This relationship has given ACAP the exclusive rights to develop a platform for archiving over 50 African census micro-data, and promoting that these data be made available and accessible to scholars, international agencies and researchers. After ten years of our work these data are finally viewed as one of the most important and unique source of population data in Africa.

In this paper, we examine some major aspects of the history of census data collection in Africa. In particular, problems of archiving, exploitation and utilization of the data will be highlighted. Presentation of the African Census Analysis Project (ACAP) will provide an idea of the strategy employed to preserve this wealth of information. It will equally give an idea of the tools developed and the various attempt made to maximize the utility of these data.

THE HISTORY AND FUTURE OF THE AFRICAN CENSUS ANALYSIS PROJECT (ACAP)

Tukufu Zuberi and Martin Bangha

Data and information are crucial elements in the development planning of any country. Ideally, reliable, consistent, accurate, and timely data are required for the purposes of policy formulation and monitoring of development programs in any democratic society. Monitoring of development programs allows for the identification of areas of sub standard performance that deserve more focused attention allowing for the limited resources and efforts to be channeled to the more deserving sectors. At independence and up to a decade post-independence, very few African countries had conducted a comprehensive census of its national population. The then scanty existing information on their populations was for the essential based on isolated administrative headcounts conducted by the various colonial administrations. Because these administrative headcounts were mainly to provide estimates of the residents of the colonies or “available manpower” for fiscal purposes and/or eventual military recruitment and due to the way such headcounts were conducted, the demography of African countries remained virgin territory. However, the past four decades or so have been associated with dramatic progress in the data collection environment in Africa. In effect, thanks to the combined efforts of various United Nations organs particularly the UNFPA and other international organizations, the continent has shifted from being “data-poor” to a ‘data-rich’ situation with most countries counting at least two censuses and several demographic surveys. In the particular case of censuses, which are by definition very heavy and intensive operations demanding a considerable amount of resources (including manpower and time), these necessitates a huge sacrifices and heavy investment in order to obtain exhaustive wealth of information on the population and its composition. The sad side of these efforts has been great underutilization of the massive information collected and the deterioration due to poor storage.

In effect, the history of census taking in most of Africa is characterized by gross under-utilization of census data, the primary preoccupation having been the execution of the processes of enumeration, data processing, the production and publication of census tables and in some cases the preparation, production and dissemination of population projections. Very little attention has been paid to systematic evaluation of the quality of census data, conducting in-depth or further analyses of data, including spatial analysis, and disseminating data in format and media that meet the diverse needs of users. Consequently, census results have not adequately informed policy formulation and program implementation, or socio-economic development planning in general. Not only have census data been under-utilized, but until recently they are poorly stored, resulting in considerable loss in a number of countries. The few countries that have managed to preserve a portion of their data tend to have poor meta-data on census codes and definitions, which has limited the utility of their census data.

In the past, access to the vast amount of data collected has often been a major problem at the local level due to the available level of information technology. Large data sets have often posed significant management, manipulation and analysis challenges to both the data producers and the potential users. In the particular case of census micro-data, these have often required serious computing support including large disk space for storage and substantial memory capacity for any analysis. It is partly for this reason that African censuses were neglected during the very period when there has been an enormous increase in the data-analytic capacity necessary to evaluate them. Census data are particularly useful because these allow for a clear picture of the target populations at lower administrative level. It is in this local use that the census, though a costly and heavy operation find most of its justification. Besides, there is obviously an important need to go beyond census analytical reports and the computation of basic demographic variables in order to develop key indicators that can be used in development planning from these rich datasets as well as further analysis for better understanding of the population health and society in general.

It was in recognition of the need to preserve Africa census data in order to avoid perpetual loss due to poor storage and also the need to encourage and enhance further analysis, dissemination and utilization of the massive census data that the African Census Analysis Project (ACAP) was initiated. In 1998, ACAP held its first African workshop in the Republic of South Africa, chaired by Professor Tukufu Zuberi, and hosted by Pali Lehohla the Focal Point for the SADC Population Census Project and Statistician General of the Republic of South Africa. Following this workshop, at the Southern African Development Community (SADC) meetings of the SADC Ministerial Committee held in Maputo, Mozambique in August 1999, a Memorandum of Understanding was established that formalized the collaboration between ACAP and SADC. This memorandum became the model of collaborative agreements between ACAP and over 30 other African governments and research institutions. This relationship has given ACAP the exclusive rights to develop a platform for archiving over 50 African census micro-data, and promoting that these data be made available and accessible to scholars, international agencies and researchers. After ten years of our work these data are finally viewed as one of the most important and unique source of population data in Africa.

In this paper, we examine some major aspects of the history of census data collection in Africa. In particular, problems of archiving, exploitation and utilization of the data will be highlighted. Presentation of ACAP will provide an idea of the strategy employed to preserve this wealth of information. It will equally give an idea of the tools developed and the various attempt made to maximize the utility of these data. We focus mainly on the evolution of census taking in Africa south of the Sahara and in terms of time span, we focus on the post-independence era or the period since the 1960s.

Brief History of Census-taking in Africa

Historically, isolated records indicate that population enumerations in Africa date way back to the mid 17th century, though Egypt is said to have a very long history of census taking. Some sources indicate counts in Egypt since 4000 BC. However, there are very few records of these counts and have been questioned (Domschke and Goyer 1986). No empirical data or records exist but the practice of headcount was apparently known to some African countries before the arrival of the European colonists. Similarly, the motives behind the headcounts were not new. Counts were historically carried out for military and taxation purposes at such times as the chiefs deemed necessary. The famous customary counting procedure involved a the heads of families dropping articles into receptacles such as grains of cereals, beads or cowries in function of the number of their dependents (Owusu 1968). Different items (or colors) were used for male and females.

Perhaps due to accessibility, available records indicate that the first systematic counts were conducted in the Island states or countries with coastal access (see Table 1). In Southern Africa, enumerations started with the settlement of the Dutch colonist at the Cape of Good Hope in 1652 (Zuberi and Kalfani 1999). The first population count in Reunion was conducted in 1690. Between the 17th to the late 19th century, this practice of population counts seems to have witnessed a rather relatively slow spread across the continent with sporadic counts of colonists and the population of special colonies. For instance, there was a remarkable series of annual enumerations of the Sierra Leone Crown colony¹ from 1833 to 1851, to check on the alleged kidnapping and re-enslavement of Africans (Lorimer 1968) but little or no interest on the population residing on the mainland of the continent. This practice of population counts really began its spread more rapidly across the continent towards the late 19th century following the famous Berlin Conference in 1884 when the European nations sat down and agreed on the partitioning of continent between themselves. This conference gave rise to the proliferation of overseas colonial territories and as the colonial administrations started to move into, and settled to administer their new found territories, there was evidently some great need for manpower as well as resources to run these administrative structures. Of course, they needed to tap into the local resources and for a better idea of the resources and manpower at their disposal, there was no better way² than to carry out their own assessment².

¹ Clarke (1968) though, reports that in the Colony of Sierra Leone, irregular censuses were held from 1802 until 1860 and decennially from 1871 until 1931. There are also indications of population counts conducted in various countries like Angola in 1777, Mauritius 1735, Seychelles 1791, etc (see Domschke and Goyer 1986) but these were mostly restricted to the European or 'civilized' populations.

² For instance, the current French institute for development research (institut de recherche pour le developpement—IRD) is a legacy of such colonial administration which until very recently used to go by the colonial acronym ORSTOM translated as scientific research office for the overseas territories. As Lorimer (1968) noted, the division of all Africa into spheres of European influence generated investment in mines, railways, and plantations that gave rise to labor supply problems.

Table 1: Some of the earliest population enumerations in Africa

<i>Country</i>	<i>First census</i>	<i>Second</i>
Algeria	1856	1861
Botswana (Bechuanaland)	1904	1911
Cape Verde	1878	1890
Egypt	1882	1897
Equatorial Guinea	1900	1910
The Gambia	1881	1891
Ghana (Gold Coast)	1891	1901
Lesotho (Basutoland)	1875	1891
Malawi (Nyasaland)	1901	1911
Mauritius	1848	1851
Nigeria	1901	1911
Seychelles	1851	1871
Sierra Leone	1901	1911
South Africa	1865	1875
Tunisia	1891	1896
Zimbabwe	1901	1911

Sources: Compiled from CSO websites; <http://www.acap.upenn.edu/Data/navigator/>; Domschke and Goyer 1986; Groupe de Démographie Africaine 1988, UN 1992.

As such, the frequency and spread of population counts across Africa (or into the hinterlands in especially) only began at the turn of the 20th century. Even then, the European population and national populations were treated differently in the procedures. Details were only collected for the European populations while the enumerations for the national populations were restricted to simple or rough estimates. As mention previously, these enumerations were mainly administrative censuses or simple headcounts that provided estimates or assessment of the residents of the colonies for fiscal purposes and/or eventual defense needs targeting military recruitment of troops. These were apparently based on arbitrary personal observations by the colonial administrator or group-assembly estimation procedure. In short, these were essentially a little more than guesstimates involving the establishment of local registers of adult males subject to taxation (see Lorimer 1968, Clarke 1968). These local tax rolls in turn provide the basis for periodic estimation of the indigenous populations by the administrators. The estimations were largely based on the hut tax method or on taxes received; calculated as the number of hut, tents or taxpayers (per district or local are) adjusted by a

factor of assumed number of persons per hut, tent or taxpayer³. This factor varied from country to country, from 3.5 in Zimbabwe through 5 in Tanganyika, 6 in Algeria to 7 in Sierra Leone and within each country by district (see Domschke and Goyer 1968, Clarke 1968) from one enumeration to another

For obvious reasons, the group-assembly method of enumeration, borrowing from a customary practice was also widespread and evidently questionnaires were pretty uncommon until towards the mid 20th century. For instance in the first population enumerations in Botswana, Ghana, at the request of the Governor, the village headmen, leaders, local teachers or clerks organized the counts in the hinterlands. The customary method of counting was essentially adopted where an urn was set in the village and each head of household or family deposited a cereal corresponding to each member. Also, in order to differentiate between males, females and children the head separated beans, groundnuts and stones. In the case of Swaziland beads of different color and sizes were used to represent male and females above and below ages 15 years (Owusu 1968, Domschke and Goyer 1986, GDA 1988).

It is therefore obvious that only basic information was collected like age and sex, etc. In line with the taxation and labor force motivation only very broad age classifications were used. The counts were very simple with only totals of men, women and children and at times, manpower or taxpayers. This notwithstanding, the methods laid the foundation for modern African demography. As these head counts and enumerations became more regular (decennial) during the first half of the 20th century, they were expanded in depth, coverage and focus. Also new methods were adopted. The modern census in Africa, which is a more complex operation compared to these earlier pre-colonial or colonial head counts, acquired prominence more recently in the latter half of the 20th century (during the post-independence era). And aside for the frequent instability and hot spots in certain parts of the continent, a relatively stable African country has conducted at least 3 censuses since independence (see table 2). Our conservative estimate suggests that there have been well over a 200 census operations conducted within Sub Saharan Africa over the last 4 decades. Of this total only few have survived.

A distinct pattern is observable between Francophone countries and the rest. Almost all of the Francophone countries gained independence in 1960 but virtually all these ex-French colonies in the immediate post-independence era, did not embark on an assessment of their population via census fashion. They continued with an inherited tradition from the pre-independence era of administrative censuses or population surveys⁴ that had been introduced to them by the French

³ As reports indicate, the taxation system (or poll tax) was initially per hut or tent. A case in point is Zimbabwe (see Domschke and Goyer 1986, GDA 1988) in 1894 where a decree established a poll tax in all administrative districts on the basis of number of huts. In 1901, this poll tax became based instead on the number of male adults.

⁴ Unlike modern population surveys, this was an approach whereby populations of various segments of the acquired territories were assessed separately at different times with particular attention to the densely populated areas.

colonial administrators. As they were being repeated, they become more systematic and were termed systematic sample surveys (see Brass 1968, Domschke and Goyer 1986). It was not until the mid 1970s that most francophone African countries switch to the census.

Table 2a: Inventory of post-independence censuses in Africa: Anglophone countries

<i>Country/Yr of independence</i>	<i>First census</i>	<i>Second</i>	<i>Third</i>	<i>Fourth</i>	<i>Fifth</i>	<i>Average Intercensal</i>
Angola	1960	1970	--	--		10
Botswana 1966	1964	1971	1981	1991	2001	<10.0
Cape Verde	1960	1970	1980	1990	2000	10.0
Gambia 1965	1963	1973	1983	1993		10.0
Ghana 1960	1960	1970	1984	--	2000	13.3
Guinea Bissau	1960	1970	1979	1991	2001	10.0
Kenya 1964	1969	1979	1989	1999		10.0
Lesotho 1966	1966	1976	1986	1996	2006*	10.0
Liberia 1892	1962	1974	--	--		12
Malawi 1964	1966	1977	1987	1998		10.6
Namibia	1960	1970	1981	1991	2001	10.2
Nigeria 1960	1963	1973	--	1991	2005	14
Sierra Leone 1961	1963	1974	1985	--	2004	13.7
South Africa	1960	1970	1980/85	1991/96	2001	<10.0
Sudan 1956	1955-56	1973	1983	1993		~10
Swaziland 1968	1966	1976	1986	1997	2006*	10.0
Tanzania 1964	1967	1978	1988	2002		11.6
Uganda 1962	1969	1980	1991	2002		11.0
Zambia 1964	1969	1980	1990	2000		10.1
Zimbabwe 1965	1962	1982	1992	2002		10.0

Source: Compiled from CSO websites; <http://www.acap.upenn.edu/Data/navigator/>; <http://www.census.gov/ipc/www/cendates/cenafric.html>, Domschke and Goyer 1986; Groupe de Démographie Africaine 1988, UN 1992.

Enumeration method: de jure vs de facto

The universally used unit of enumeration, even implicitly in historical population listings has been the household (van de Walle 2006). This use of the household as the elementary unit of enumeration is associated with two basic concepts: *de jure* and *de facto* populations. In the latter case, individuals are enumerated at the place where they are actually found on the census night as

opposed to where they usually reside. As van de Walle (2006) suggest, most French-speaking countries in Africa adopt the *de jure* approach while most English-speaking countries go by the *de facto* mode of enumeration. In effect, it appears most immediate post-independence censuses or most the first in the series of modern censuses in Africa were conducted on both the *de jure* and *de facto* basis. But as the concept of household has become rigorously defined, most recent censuses tend to collect data on population and housing characteristics on a *de facto* basis. However, where there is the existence of a permanently high level of mobility or a considerable moving population like nomads both approaches are employed jointly. Apparently most of the earlier or pre-independent censuses were conducted on a *de jure* basis (see Domschke and Goyer 1986). Besides unlike the pre-independence censuses where there was frequent resort to the system of collective or group enumerations, the enumerations in the post-independence era have adopted essentially the house-house canvassing method.

Census questionnaires

Certain questions especially those that provide for the estimation of population count and measurement of changes in basic composition are considered elementary in censuses and so must appear in all census questionnaires. Questions on supplementary topics are often modified or dropped in some censuses while new topics are frequently introduced. Earlier censuses were very limited in subject coverage and mainly collected or targeted information for estimating the total number, broad age groups, and sex composition. As the main aim was to assess the tax base and resources, some of these censuses did include information of the livestock. As mentioned earlier, over the years the operation has become heavy and complex along with the tools and instruments. With increasing ability to handle more complex data collection operations as well as data needs, there is also an increasing tendency to load the questionnaire with questions are being included in the census. For instance data on the stock and condition of housing as well as ownership of some durable goods and access to facilities have now almost acquired the rank of elementary questions. Meanwhile a few recent censuses have gone the extra mile of combining the operation with what used to be previously an agricultural census (e.g. Benin 2001 & Uganda 2002). In addition to traditional data, some recent censuses now include information on the orphanhood status of children. Census taking has expanded to be more comprehensive & wider coverage of topics. Census cartography has equally improved and with GIS databases being mounted, the likelihood of omissions of localities should be minimized.

In terms of questionnaire types, the censuses distinguish between the individual household questionnaire and the collective one. However there are several countries that have frequently used what is termed the long and forms (or detailed and general). The long or detailed in this case is often administered to a sub-sample of the population while the short or general questionnaire usually contains the first few basic questions (4-8) of the general questionnaire and is administer to the entire population. Apparently this might be a strategy to reduce cost of census operation.

Availability of the censuses

Of all the censuses taken pre- and immediate post-independence in Africa very few records and documentation have been kept let alone the huge amount of data so collected. Indeed very little of virtually no efforts were made to preserve these data. Though earlier operations might have focused on headcounts, for obvious reasons, African census history is pretty rich. It is unacceptable to sacrifice so much resource to collect data and little efforts to preserve the data so collected. Having taken a census is good, two is better and three even more better for planning and assessment of welfare program. However, having taken a census and losing the data is worse than not taking at all. The two Congos are in this situation of having lost all their censuses.

Table 2b: Inventory of post-independence censuses in Africa: Francophone countries

<i>Country/Yr of independence</i>	<i>First census</i>	<i>Second</i>	<i>Third</i>	<i>Intercensal Interval</i>
Benin 1960	1979	1992	2002	11.5
Burkina Faso 1960	1975	1985	1996	10.5
Burundi 1962	1979	1990	2005- TBC	13
Cameroon 1960	1976	1987	2005	14.5
Central African Rep. 1960	1975	1988	2003	14
Chad 1960	N/A	1993	2005-TBC	12?
Comoros 1975	1980	1991	2003	11.5
Congo Brazza 1960	1974	1984	1996	11
Cote D'Ivoire 1960	1975	1988	1998	11.5
Gabon 1960	1980	1993	2003	11.5
Guinea 1958	N/A	1983	1996	13
Madagascar 1960	1975	1993	2005-TBC	15
Mali 1960	1976	1987	1998	11
Mauritania 1960	1976-77	1988	2000-01	12
Niger 1960	1977	1988	2001	12
Rwanda 1962	1978	1991	2002	~16
Sénégal 1960	1976	1988	2002	13
Togo 1960	1970	1981	2005-TBC	~16

Source: Compiled from CSO websites; <http://www.acap.upenn.edu/Data/navigator/>; <http://www.census.gov/ipc/www/cendates/cenafric.html>, Domschke and Goyer 1986; GDA 1988, UN 1992

In terms of analysis we must recognize that there have been considerable improvements over time. The censuses now include a more elaborate publication plan and the reports have moved a long way from the listing of statistical tables to being more analytical as well as thematic. Nevertheless, these censuses have remained relatively underutilized and consequently, census results have not adequately informed policy formulation and program implementation, or socio-economic development planning in general. Besides, until recently they are poorly stored, resulting in considerable loss in a number of countries.

Metadata situation:

Of course, census data collection skills have been sharpened and attempts are frequently being made not only to improve on quality but also to address (collect) data to answer many innovative questions, census reports have progressively become somewhat relatively analytical. However, there are several aspects of censuses in Africa where there has been little improvement. Key among such areas is the metadata situation. While technology has improved dramatically and storage is becoming increasingly better than before, it appears the metadata culture is still below expectation. It is still a major problem trying to figure out how most computations and imputation were made in the data even for recent censuses because these are rarely preserved. On the contrary it appears some considerable efforts were made to preserve computations in some earlier censuses than in the recent ones⁵. Knowing how computations and imputations were made is crucial for understanding and analyzing the data. It also allows one to check whether the recovered data are correct and thereby save time.

There are frequent cases of located codebooks with little or no background information as to how variables were coded and/or how the computations were made. And with the high mobility or circulation of census bureau personnel, it is usually time consuming trying to locate persons with some basic knowledge on how the computations were made. Besides, even when we manage to locate such a person as it happens in some rare cases, memory problems would have set in. This is further compounded by the fact that technology has improved dramatically. This creates some real difficulties in revisiting the data and making any meaningful analysis. Even, with the efforts by ACAP and the systematic attempts to archive documentation of the various censuses, the process is still somewhat slow.

In sum, we note that modern censuses of national populations have been conducted more systematically and regularly since independence across Anglophone Africa, except for countries that have experienced some amount of political instability. This does not seem to have been the case with much of francophone Africa where the conduct of censuses since independence appears to be

⁵ A case in point is Cameroon: for the 1976 (first nationwide census) the computation programmes are available in the census documentation and provides clear idea of how certain variables were obtained. On the contrary, the most recent census has no such information for any of the computations.

more erratic. For instance, most countries of Anglophone Africa have conducted at least 4 censuses and these are generally within close range of the recommended interval of 10 years. On the contrary, the most stable francophone counterpart has conducted 3 censuses and they frequently tend to drift towards intercensal intervals of 15 years. In the section that follows, we describe the ACAP initiative, our noble efforts to preserve this important part of African heritage. Our conservative estimates suggest that there have been well over a 200 census operations conducted within Sub Saharan Africa over the last 4 decades. Of this total only few have survived. Thanks to ACAP efforts, some considerable fragments of the 1970s and the 1980s censuses can now be accessible. In what follows we present the ACAP initiative. This presentation of ACAP and its major accomplishment over the last decade will provide an idea of the strategy employed to preserve this wealth of information as well as the tools developed to maximize the utility of these data.

The ACAP Initiative:

In recognition of the need to preserve Africa census data in order to avoid perpetual loss due to poor storage and also the need to encourage and enhance further analysis, dissemination and utilization of the massive census data, the African Census Analysis Project (ACAP) was initiated as a joint initiative of the Population Studies Center, University of Pennsylvania and African research and governmental institutions. This was to allow for collaboration with various African governments and research institutions at archiving and analyzing African census data, both at national and sub-national level, in order to inform appropriate policy interventions on the continent.

At the onset, ACAP pursue basically three main objectives:

- Development of the Pan-African Census Explorer (PACE) for the purpose of archiving and analyzing African census micro-data.
- Demographic capacity strengthening in Africa.
- And research collaboration with African researchers.

In 1998, the African Census Analysis Project held its first African workshop in the Republic of South Africa. Following this workshop, at the Southern African Development Community (SADC) meetings of the SADC Ministerial Committee held in Maputo, Mozambique in August 1999, a Memorandum of Understanding was established that formalized the collaboration between ACAP and SADC. This memorandum became the model of collaborative agreements between ACAP and over 30 other African governments and research institutions. This close working relationship with African governments has allowed ACAP to develop a platform for archiving over 50 African census micro-data, and making these data available and accessible to scholars, international agencies and researchers. After ten years of our work these data are finally viewed as one of the most important and unique source of population data in Africa. Most African countries do not have effective

national vital registration systems, and census data is the only empirically based source of estimates for the entire population.

Accomplishments Phase I

PACE development and data recovery efforts: One of the key accomplishments and attractions of the initial phase of ACAP is the creation of a unique data bank that helped avoid the destruction of many of the 1970 and 1980 rounds of the data alongside the development of the Pan-African Census Explorer. In fact, we continue to be the repository of the data for several countries as they experience the “disappearance” of their national censuses. The recovery and transfer of past censuses constitutes an important accomplishment of the first phase because most of these censuses had been classified as irretrievable and unusable. ACAP has been quite successful with its traditional role of negotiating and recovering census micro-data from various African nations. These data are then processed (i.e. made machine-readable, cleaned, and harmonized) for integration into PACE. We usually proceed by trying to read the data tapes in the countries and if not possible at that level of the country, they are then brought over to the US for further attempts.

The difficulty of this recovery process usually starts with the attempts at locating the census tapes on which the past censuses are stored. Based on our observation and past experience, it is often the case that once a new census is conducted the preceding census is relegated to the background. It not uncommon to arrive a country to find no one even recalls where the tapes containing the previous censuses were stored and kept let alone talk of reading them. However, our multiple trips across the continent and relentless efforts in collaboration with the censuses offices in the countries have been quite rewarding. We managed to track most old tapes of past censuses and have been particularly successful in recovering a good number of the data stored on such tapes and transferring them to modern computer storage media. Examples of recent successful recovery efforts include the Ghana 1984 census, Benin 1979 census, Cameroon 1976 census, and Malawi 1977 census.

For an idea of the challenges, time and resources involved in this recovery exercise, it is perhaps worthwhile to illustrate this with the most recent case of the successful recovery of the 1977 Malawi census. A total of fourteen (14) 9-track magnetic tapes that held this census (originally created on ICL 1900 series mainframe computer) have been in our keeping for over five years. During this period the tapes have made several rounds to commercial data recovery firms with little success at recovering and transferring the data to modern media. Most of the data recovery firms after diagnoses declared these tapes in pretty bad shape and we had almost given up any little hopes of ever reading these census data. For instance, one of the data recovery services upon receipt of the tapes sent us this message along with the quote for services:

...The tapes are in pretty bad shape and will require treatment for striction to loosen up emulsion ...and cleaning so they can be read by our drives...

But after treatment and diagnoses, our little hopes were dashed when we received the following from same firm:

I have researched the job sent to us from the U. of Pennsylvania to copy tapes. The tapes were in bad shape and could not be copied. They are returned to the U. and invoiced for...

Following similar results from other attempts we then became a little reluctant to continue spending on tapes with little hope of success. Somehow we did not write off on this and one of the attempts was fruitful when the tapes contents were converted into sets of 'bitstream' files which were subsequently converted to ASCII characters. However, we were not able to make anything of the resulting information because this did not match the census data expected and appeared to have some form of compressed data embedded in it. In effect, some sort of memory-saving compression⁶ routine had been used to store the data, the details of which had been lost. As part of the continuous efforts to recover these data, our Systems Specialist had concurrently posted, on the internet; some request for information on the ICL 1900 series machines. In the course of numerous exchanges with various experts on these machines, we came across a British programmer who had spent his entire career (in application and systems programming) on these machines. It was thanks to his assistance in November 2004 that a program was finally prepared that pad out the records with spaces and successfully produced data that lined up correctly as expected.

It is thanks to such relentless efforts and continuous resourcefulness that we have made tremendous progress in helping avoid the destruction of many of the 1970 and 1980 rounds of African censuses. The expanding PACE collection currently consists of census micro-data of 55 censuses from close to 26 African nations. ACAP has the largest collection of complete census data in the world (see table 3).

Research capacity strengthening: ACAP has promoted the use of census micro-data for a better understanding of African health and society. ACAP has helped to strengthen the research capacity of over 100 African researchers in several ways: granting support to students for training in Africa; collaborating with and supporting selected population- and health-oriented scholars working in Africa; bringing over 50 African students and researchers to the University of Pennsylvania for Masters and PhD degree training, or postdoctoral research; organizing over 10 workshops for scholars using census data in their research; and co-organizing or regularly participating in seminars, conferences, and census-related training workshops or activities in Africa.

A major part of the training support covered graduate studies in demography at the University of Pennsylvania where most students are encouraged to work on the archived census data as part of ACAP research activities. Under this training support program, ACAP has produced

⁶ As explained earlier in the paper, we could have been saved this time, resources and energy if there was some metadata indicating clearly how these data were stored or compressed. But there is still a lot ground work needed to encourage the metadata spirit.

Over 20 PhDs and Masters' graduates during the 10 years of its existence. The most recent PhD graduates who have completed their studies and successfully defended their dissertations over the last five years are from Ghana, Malawi, Sierra Leone, South Africa, Tanzania and Uganda. ACAP has also hosted about seven postdoctoral fellows from Sierra Leone, Nigeria, Tanzania and Zimbabwe who worked on various aspects of African demography as part of ACAP research activities using the archived census data.

Table 3: African Census Data Collection (ACAP Archive) till date

One year of census data			Multiple years of census data		
Country	Year	Number	Country	Years	Number
Chad	1993	1	Benin	1979; 1992	2
Central African Rep.	1988	1	Botswana	1981; 1991	2
Cape Verde	1990-part	1	Burkina Faso	1985; 1996	2
Mali	1987-part	1	Cameroon	1976; 1987	2
Mozambique	1997	1	The Gambia	1973; 1983; 1993	3
Niger	1988	1	Ghana	1971-PES; 1984; 2000-5%	3
Nigeria	1991-PES	1	Guinea	1983; 1996	2
Former SA*			Kenya	1969; 1979; 1989; 1999	4
Ciskei	1991	1	Lesotho	1986; 1996	2
Venda	1991	1	Liberia	1843; 1974	2
Sudan	1973	1	Malawi	1977; 1987; 1998	3
Swaziland	1986	1	Mauritius	1990; 2000	2
Tanzania	1988-20%	1	Senegal	1976; 1988	2
			South Africa	1970; 80; 85; 91; 96 & 01	6
			Bophuthatswana*	1980; 1991	2
			Uganda	1980-part; 1991	2
			Zambia	1980-part; 1990	2
Total		12			43

* Former South Africa Homelands

The recently completed PhD dissertations based on the archived census data at Penn include:

- Differentials in infant and child mortality in East Africa (Kenya, Tanzania, Uganda)
- Health, well-being and mortality in Africa (Botswana, Lesotho, Zambia)
- Mortality in twentieth-century Malawi
- The Demography of migrant population in South Africa
- Grandparents and the mortality of their grandchildren in the Gambia: Findings from the 1993 census

- Mortality in the era of HIV/AIDS in Tanzania
- Poverty in Post-Apartheid South Africa: Measurement, Trends and the Demography of the Poor

Upon completion of their studies most of these students returned to their home countries and are currently working with their respective institutions or universities in Africa. They also retain the status of ACAP research collaborators for Africa. The return of most of these students is a demonstration of ACAP's commitment to demographic capacity strengthening in Africa. Besides these ACAP-supported African students who recently completed their doctoral studies, ACAP is currently supporting three African students from Kenya, Ghana and Cameroon to pursue graduate/doctoral studies in demography at the University of Pennsylvania, one of whom is currently at the dissertation stage. Meanwhile, to date, ACAP has also supported several other African students for short internship programs at the University of Pennsylvania. Meanwhile some of the training support was provided to students based at African institutions.

During the last five years, ACAP has been hosting on average 8-10 visiting scholars per year, most of them from African countries. The countries cover by these visits during the couple of years include Benin, Burkina Faso, Cameroon, Guinea, Mozambique, Ghana, Kenya, South Africa, Swaziland and Uganda. The estimated number of such visitors and the person-months duration of stay are presented in Table 4.

Table 4: Estimated African Scholars hosted at ACAP from 1998-2005

Country	Scholars hosted	Person-months of stay
Benin	1	1
Burkina Faso	1	2
Cameroon	1	1
Ethiopia	1	8
The Gambia	1	2
Ghana	3	69
Guinea	1	1
Kenya	2	29
Malawi	2	64
Mali	1	3
Mozambique	3	4
Nigeria	4	24
Senegal	2	2
Sierra Leone	2	52
South Africa	11	115
Swaziland	1	4
Tanzania	2	51
Uganda	3	68
Zimbabwe	1	37
Total	43	537

Publications: One of the main aims of ACAP is to contribute toward making census data more useful and relevant to scholarly pursuits, policy formulation and development planning in Africa. This is accomplished in part through promoting extensive research using census micro-data for a better understanding of African health and society. In line with this, we now have in place a multi-volume series with already two edited volume to our credit. Moreover, over the past years some of the research results have been presented at various international conferences and seminars, most of which have been published in peer review journals. The topics covered by the recent journal publications using the ACAP archived data include but not limited to (see appendix 1 and 2 for the exact titles):

- Racial classification and the modern census in South Africa,
- Polygyny and Fertility
- Marriage influences on individual behavior in urban Africa
- Assessing the reliability of census data
- Migrants and labor market
- Influence of individual and household characteristics on schooling attrition
- Living standards of the elderly
- Parental survival, living arrangements, and school enrolment of children in the era of HIV/AIDS;
- The effect of living standards on childhood mortality;
- Socioeconomic status and child mortality: Analysis using housing and household characteristics from African census data
- Disability in general and disability among the children of migrants
- Impact of external causes on expectation of life
- Childbearing and schooling
- The HIV/AIDS epidemic, kin relations, living arrangements and the elderly

Collaborative Efforts at Using Census Data at Regional Level

In collaboration with various regional and national institutions in the countries, ACAP has initiated various research programs aimed at enhancing the use of census data in development planning. Among such efforts, are four major ones worthy of mention: the collaborative project with the Southern African Development Community (SADC) Statistical Committee program on the further analysis of the 2000 round of censuses; the collaborative program of research (with the International Network of field sites for the continuous Evaluation of Population and Their Health in developing countries-INDEPTH) on population and health in Africa, the program to strengthen the information base for proactive development planning and research program on the situation of orphans in African Societies.

- ***SADC Program on Analysis and Utilization of Census Data***

The Southern African Development Community (SADC) countries like most African have each conducted at least 2 population censuses. However, data from these censuses have been grossly under-utilized and not adequately informing socio-economic development in the SADC region. In a bid to address the need for policy oriented demographic and spatial analysis in the region, ACAP spearheaded this project that was approved by the SADC Council of Ministers. The main aim of the Analysis and Utilization of Census Data Project was to facilitate strengthening the capacity of SADC member countries in demographic, social and spatial analysis in order to promote the understanding and utilization of census data by government policy makers and planners.

- ***ACAP-INDEPTH Research Program on Population and Health in Africa***

Recognizing that African censuses archived by ACAP and the Demographic Surveillance Systems (DSS) coordinated by the INDEPTH Network both produce vast and complementary kinds of demographic data for Africa that offers an exciting potential to examine African population and health, a collaboration between the two institutions was consolidated. The main objective of this collaboration was to pool the resources and expertise of ACAP and INDEPTH to undertake high-quality joint research projects in order to inform demographic and health policy in Africa. The first step of the collaboration has started with a few INDEPTH sites that had the capacity to contribute the basic data needed for investigating common research topics. Based on mutual understanding between INDEPTH and ACAP, Burkina Faso, Ghana, Mozambique, Senegal, South Africa, and Tanzania were selected for this initial phase and will serve as demonstrative examples of the usefulness of the collaboration. The second phase of the collaboration is expected to involve the majority of the INDEPTH sites and ACAP collaborators.

List of Selected Pilot Countries and Partner Institutions (ACAP-INDEPTH)

Census Bureaus	DSS Sites
<ul style="list-style-type: none"> ■ INSD, Burkina Faso ■ Statistical Service, Ghana ■ INE, Mozambique ■ DPS, Senegal ■ Statistics South Africa ■ NBS, Tanzania 	<ul style="list-style-type: none"> ■ Nouna (CRSN), Burkina Faso ■ UERD, Burkina Faso ■ Navrongo (NHRC), Ghana ■ Manhica, Mozambique ■ Niakhar DSS ■ Agincourt, South Africa ■ Ifakara, Tanzania ■ Rufiji, Tanzania

- ***Data needs for Proactive Development Planning***

The aim of this collaborative activity is to enhance the process of planning and the implementation of democratically based proactive development programs at the sub-national level by making census data easily accessible and available at the level of districts for analysis. This activity is the

practical demonstration of the usefulness of PACE in enhancing the capacity to manage, analyze and disseminate census data. The pilot phase of this program was planned for Uganda in 2004 in collaboration with the Uganda Bureau of Statistics (UBOS) with the hope of eventual replication elsewhere. Indeed, the District Planning units (DPUs) are the focus of this program. With decentralization of the planning function, the DPUs need to guide sub counties on how to develop their sub county plans of action. This guidance has to be supported by the relevant data. Hence the DPUs need the ability to provide information for planning disaggregated by sub-county and lower levels.

- ***Situation of Orphans in African Societies: Case of the Sahel.***

This collaborative effort with the University of Laval, Québec (which coordinates the activities of the demography network of the francophone universities agency (*Réseau Démographie Agence Universitaire de la Francophonie RD-AUF*) is not only intended as a way of expanding the geographical scope and coverage of the project. It was so elaborated to enable the francophone community in Africa to equally benefit more from our activities. It is in this light that our joint research project to examine the *Situation of Orphans in Africa Societies* focuses essentially on the Sahel countries. In effect, orphanhood is not only a common phenomenon in sub-Saharan Africa, but recent evidence suggests a dramatic increase. While this is quickly attributed to the AIDS pandemic coupled with the persistence of other numerous pandemics, weaknesses and crises of the health/social system, there is little understanding of the support mechanisms available to this vulnerable group and the likely effect on their life chances. This project proposes to document the phenomenon of orphanhood in Africa and to examine their lifetime achievements so as to assess the potential effect of parental loss on their life chances. By exploring the different sources of data available on orphans, our research program will attempt to highlight three important aspects of their lives: education, economic insertion and entry into adulthood that are most likely to be affected by the death of one or both parents. A number of francophone West/Central Africa countries have been selected for the current phase of the program. This is part of the concrete plans for the production of a volume.

The Future of ACAP

We believe that making African census data available to scholars' world wide is important; however, availability alone will not ensure that it is utilized on the African continent by African scholars. Hence, the second phase in the activities of ACAP has two major aims:

Aim 1: To create a mirror of this data bank on the African continent. We have taken the preliminary steps to accomplish this objective in several African nations.

Aim 2: Facilitating the collaboration of scholars to produce a multi-volume series, *A General Demographic of Africa* that will improve our knowledge of African population and society.

The Pan-African Census Explorer

The Pan-African Census Explorer (PACE) is an analysis engine that is aimed at promoting high quality research in a user-friendly interface. PACE is very important for data implementation and facilitates easy access to census data for demographic and statistical analysis. To date, ACAP has assembled a unique census micro-data collection. The expanding PACE collection currently consists of census micro-data collection of 55 censuses from close to 26 African nations. PACE requires serious computing support. Integrating data in a database server, with mirrors in selected African countries, will allow access to the data with a variety of point and click interfaces and guided applications of demographic methods to census data. PACE also facilitates easy distribution of census data by government agencies, while respecting the issue of confidentiality. It provides for the cataloging of databases and metadata. The user can browse datasets and retrieve documentation.

It is important that we not only archive these data, but provide access for analysis to scholars and researchers. Hence, one of the ongoing objectives of ACAP is to use this platform to make census data available and easily accessible to scholars, international agencies and researchers, thereby rendering this wealth of information more useful and relevant to policy formulation and development planning in African countries. In this light, we plan to have mirror sites of the data at the level of the countries. As part of the preliminary steps to accomplish this, we successfully set up and launched PACE South Africa. While this is currently available only to a restricted number of users (who are evaluating the product), this will eventually allow scholars, researchers and policy makers in South Africa and elsewhere to access their censuses, surveys and the corresponding metadata. We intend to consolidate this process and make it more publicly available.

Also, we intend to progressively create the DVD versions for the collaborating countries for them to try out (disseminate) and provide us with feedback for further enhancement of the explorer before it will be made publicly available. While our project has encouraged African nations to make their data available to researchers and scholars in the Europe and the United States via the internet, such efforts do not result in increased African use of these data. The cost for internet access remains prohibitive for most African institutions. African governments have yet to settle the regulation of their continents' international gateways, be they proposed East African Submarine cable system (EASSy) or SAT 3, the barriers to internet access for African researchers, health workers, and policy makers will continue to be slow and ineffective. Therefore we intend to provide computer and internet-based access to these data on the African continent.

A General Demography of Africa

One of the main aims of ACAP is to contribute toward making census data more useful and relevant to scholarly pursuits, policy formulation and development planning in Africa. This is accomplished in part through promoting extensive research using census micro-data for a better understanding of African

health and society. As a culmination of the efforts made by ACAP over the years, we now have in place a multivolume series. This series titled *A General Demography of Africa* is being published by M.E. Sharpe Press. As part of the practical arrangement and to ensure the quality of the series, we have formed an international panel of population and health experts who will be serving as the series *editorial advisory board*. Over 20 renowned population-health scientists were identified and invited to serve on the editorial advisory board; an overwhelming majority of them have already confirmed their willingness and enthusiasm to be part of this noble effort. In the summer of 2005, during a general meeting of the editorial board in Tours, France, members deliberated and agreed on the best strategy for the successful realization of this series.

The first volume in the series, *The Demography of South Africa*, edited by Tukufu Zuberi, Amson Sibanda and Eric Udjo was published in 2005. The second volume on *African Households: Censuses and Surveys* was edited by Etienne van de Walle and published in 2006. Volumes on poverty; evaluation of censuses and migration are under preparation and we anticipate additional volumes in the coming years. Some of the volumes will have a regional focus while others will be country specific. These volumes will be an essential output of our future activities alongside activities to concretize the establishment of the data bank and protocol for accessing African census data both on the continent and via the internet from ACAP.

DRAFT

References

Brass, W. et al (eds.) 1968. *The Demography of Tropical Africa* Princeton University Press, Princeton

Clarke, J.I. 1968. "Population Distribution in Sierra Leone" in J.C. Caldwell and C. Okonjo (eds.) *The Population of Tropical Africa* Longmans: 270-277.

Domschke, E. and D.S. Goyer 1986. *The Handbook of National Population Censuses: Africa and Asia*. Greenwood Press, Westport, CCT.

Groupe de Demographie Africaine (GDA) 1988. *Population Size in African Countries: An Evaluation* Vol. II. IDP-INED-INSEE-MINCOOP-ORSTOM, Paris.

Lorimer, F. 1968. "Introduction" in Brass et al (eds.) *The Demography of Tropical Africa* Princeton University Press, Princeton: 3-11.

Owusu, D.J. 1968. "Taking a Population Census in Tropical Africa" in J.C. Caldwell and C. Okonjo (eds.) *The Population of Tropical Africa* Longmans: 135-141.

United Nations, 1992. *Demographic Handbook for Africa 1992* Population Division, Economic Commission for Africa (ECA) Addis Ababa.

Van de Walle, E. 2006. "Introduction" in E. van de Walle (ed.) *African Households: Censuses and Surveys*, M.E. Sharpe Inc., New York: xxi-xxxix.

Zuberi, T. and Akil K. Khalfani. 1999. "Racial Classification and the Census in South Africa, 1911-1996." *ACAP Working Paper No 7*. The African Census Analysis Project (ACAP), Population Studies Center, University of Pennsylvania, Philadelphia.

Appendix 1: Recent PhD Dissertations at the University of Pennsylvania Based on the Census Micro-data

PhD Dissertations by African Graduates

Gideon Rutaremwa (2002): Differentials in infant and child mortality in East Africa (Kenya, Tanzania, Uganda)

Ayaga Bawah (2002): Health, well-being and mortality in Africa (Botswana, Lesotho, Zambia)

Henry Doctor (2003): Mortality in twentieth-century Malawi

Kevin Thomas (2004): The Demography of migrant population in South Africa

Claire M Noel-Miller (2005): Grandparents and the mortality of their grandchildren in the Gambia: Findings from the 1993 census

Robert Mswia (2006): Mortality in the era of HIV/AIDS in Tanzania

Sandile Simelane (2006): Poverty in Post-Apartheid South Africa: Measurement, Trends and the Demography of the Poor

Dissertation Chapters by other Graduates:

Ria Baker (2003): Marriage Forgotten or Foregone? An examination of the transformations of marriage in Botswana

Solene Lardoux (2005): Polygyny and fertility in rural Senegal

Appendix 2: Some Recent Publications Analyzing the Census Micro-data at ACAP

Khalfani, Akil Kokayi and Tukufu Zuberi. 2003. "Racial classification and the modern census in South Africa, 1911-1996." *Race and Society* 4(2): 161-176.

Zuberi, Tukufu, Amson Sibanda, Ayaga Bawah and Amadou Noubbissi. 2003. "Population and African Society." *Annual Review of Sociology* 29:465-86

Lardoux, Solène and Etienne Van de Walle 2003. "Polygyny and Fertility in Rural Senegal" *Population E*. No. 6. [Also in French as « Polygamie et fécondité en milieu rural sénégalais. » *Population F*, No. 6]

Luke, N. and Munshi, K. 2003. "Does marriage influence individual behavior in urban Africa? Evidence from a high HIV area in Kenya", MIT *Department of Economics Working Paper 03-20*, Cambridge, MA.

Mba, Chuks J. 2003. "Assessing the reliability of the 1986 and 1996 Lesotho census data" *Journal of Social Development in Africa* 18(1): 111-128.

Zuberi, Tukufu and Amson Sibanda 2004. "How Do Migrants Fare in a Post-Apartheid South African Labor Market?" *International Migration Review* 38(4): 1462-1491.

Sibanda, Amson. 2004. "Who Drops Out of School in South Africa? The Influence of Individual and Household Characteristics" *African Population Studies* 19(1): 99-117.

Noubbissi, Amadou. 2004. "Living Standards of the Elderly in South Africa" *African Population Studies, Supplement* 19(a): 201-219.

Doctor, Henry 2004. "Parental survival, living arrangements, and school enrolment of children in Malawi in the era of HIV/AIDS" *Journal of Social Development in Africa* 19(1): 31–56.

Doctor, Henry 2004. "The effect of living standards on childhood mortality in Malawi" *African Population Studies, Supplement* 19(a): 241–263.

Bawah, Ayaga and Tukufu Zuberi 2005. "Socioeconomic status and child mortality: An illustration using housing and household characteristics from African census data" *African Population Studies, Supplement* 19(b): 9–29.

Phillips, Heston and Amadou Noubissi 2005. "Disability in South Africa" *African Population Studies, Supplement* 19(b): 115–138.

Thomas, Kevin J.A. 2005. "Disability among the children of migrants in South Africa" *African Population Studies, Supplement* 19(b): 139–164.

Mba, Chuks J. 2005. "The impact of external causes on South Africa's expectation of life" *African Population Studies, Supplement* 19(b): 165–177.

Mba, Chuks J. 2005. "Racial Differences in Marital Status and Living Arrangements of Older Persons in South Africa" *Generations Review* 15(2): 23–31.

Bawah, Ayaga and Tukufu Zuberi 2005. "Socioeconomic status and child mortality in Southern Africa" *Genus* 61(1–2): 55–83.

Doctor, Henry and Sandile E Simelane 2005. "The impact of living standards on childhood mortality in South Africa: Evidence from cross-sectional data" *Journal of Social Development in Africa* 20(2): 7–38.

Sangeetha Madhavan & Kevin J. A. Thomas 2005. "Childbearing and Schooling: New Evidence from South Africa" *Comparative Education Review* 49(4): 452–467.

Merli, M.Giovanna and Alberto Palloni 2006. "The HIV/AIDS Epidemic, Kin Relations, Living Arrangements and the Elderly" *Aging in Sub-Saharan Africa: Recommendations for Furthering Research*. National Academies Press: 113–162.

Thomas, K.J.A. Child Mortality and Socioeconomic Status: An Examination of Differentials by Migration Status in South Africa." *International Migration Review* forthcoming.

Books:

The Demography of South Africa. Tukufu Zuberi, Amson Sibanda and Eric Udjo (eds.), 2005.

African Households: Censuses and Surveys (Botswana, Burkina Faso, Cameroon, Cote d'Ivoire, The Gambia, Kenya, Senegal, South Africa and Tanzania). Etienne van de Walle (ed.) 2006.