Table des matières

Contents

Editorial

Nous souhaitons la bienvenue aux lecteurs de ce troisième volume du Journal et remercions tous ceux qui nous ont envoyé des commentaires sur le deuxième volume. Ce volume présente des articles et autres informations qui sont importants pour la communauté statistique Africaine. Les articles présentent un gamme de sujets qui couvrent la coopération institutionnelle pour le développement de la statistique, l'amélioration de la collecte des données en utilisant la méthodologie du Questionnaire Unique d'Indicateur de Bien être de Base (QUIBB), les procédures d'estimation dans les petites unités géographiques et le Système d'Information Géographique (SIG), la validation et l'apurement des données pour les non-réponses, l'évaluation du cycle 2000 des recensements et l'utilisation des informations statistiques pour faire le suivi et mesurer le progrès accompli dans la mise en œuvre des objectifs nationaux et internationaux. Ces articles ont été écrits par des experts d'horizons divers : statisticiens, démographes, analystes du développement et experts en technologie de l'information. C'est ce qui fait la richesse intellectuelle de ce journal. Nous voudrions remercier les auteurs qui ont choisi de publier leurs articles dans le journal et les personnes qui ont passés en revue ces articles pour leur assistance.

Dans le deuxième et troisième trimestre de l'année, un certain nombre d'événements ont eu lieu. Ces événements vont profondément influencer le cours du développement de la statistique en Afrique. Le Conseil d'Administration de la Banque Africaine de Développement (BAD) a érigé la Division Statistique en Département. La Commission Economique des Nations Unies pour l'Afrique (CEA) a décidé de créer le Centre Africain de Statistique. La Division de la Statistique de la CEA avait été dissoute suite à la restructuration des années 1990. Ces actions louables de la BAD et de la CEA ont non seulement rehaussé le profil de la statistique dans ces institutions, mais ont aussi fourni le cadre institutionnel pour l'assistance au développement de la statistique dans les pays Africains. Nous voudrions feliciter les hautes directions de ces deux institutions régionales pour ces actions.

Egalement au cours de ces deux trimestres, la BAD a organisé quatre ateliers sous-régionaux pour les pays du Marché Commun de l'Afrique Orientale et Australe (COMESA), les pays anglophones de la Communauté Economique des Etats de l'Afrique de l'Ouest (CEDEAO), des pays de la Communauté de Développement de l'Afrique Australe (SADC) et les pays membres de l'Observatoire Economique et Statistique d'Afrique Subsaharienne (AFRISTAT). Ces ateliers avaient un triple objectif : (i) lancer officiellement l'assistance de la BAD aux pays membres régionaux pour élaborer ou mettre à jour leurs Stratégies Nationales de Développement de la Statistique (SNDS), (ii) renforcer les connaissances des acteurs-clés du système statistique national sur les processus de la SNDS et comment les mettre en œuvre et (iii) faciliter l'échange d'expérience entre les pays. Ces ateliers ont mis l'accent sur l'intégration des questions et préoccupations sectorielles dans la conception des SNDS. Suite aux ateliers du COMESA et de la CEDEAO, un forum de discussion par Internet a été établi pour échanger et partager les connaissances, l'expérience et les informations sur les statistiques sectorielles.

Finalement en Septembre 2006, la communauté statistique et les partenaires au développement se sont rendus à Dar es Salem, Tanzanie, pour célébrer 40 ans de service rendu à la région par le Centre de Formation en Statistique de l'Afrique de l'Est (EASTC). Au cours de ces années, le Centre a formé des statisticiens dont certains ont été promus comme directeurs des instituts nationaux de statistique et d'autres occupent des postes-clés dans des organisations internationales. Nous félicitons le Centre pour ce service exemplaire rendu à la sous région et au continent Africain en général.

Dr Charles Leyeka Lufumpa
Co-Président du Comité de rédaction

Professeur Ben Kiregyera
Co-Président du Comité de rédaction

Editorial

We welcome readers to Volume 3 of the journal and thank all those readers who sent us comments on Volume 2. In this volume of the journal, there are interesting articles and other materials of importance to the African statistical community. The articles range from institutional cooperation in statistical development to improving data collection using the Core Welfare Indicators Questionniare (CWIQ) technology; using small-area estimation techniques and Geographical Information System (GIS) technology; validation solutions and correcting survey data for non-response; assessment of the 2000 censuses and use of statistical information to monitor and measure progress towards national and international development goals. These articles have been written by experts with different backgrounds – statisticians, demographers, development analysts and ITexperts . This is what makes this a knowledge rich journal. We would like to thank the authors for choosing to publish their articles in the journal and the referees for their assistance.

In the second and third quarters of the year, a number of things took place that will, in a profound manner, influence the course of statistical development in Africa. The Board of Directors of the African Development Bank (ADB) upgraded the Statistics Division into a full Department of Statistics. At the United Nations Economic Commission for Africa (ECA), a decision was make to create an African Centre for Statistics. As part of the restructuring process in the mid-1990s, the Statistics Division at ECA had been scrapped. Among other things, these laudable actions have not only raised the profile of statistics in the respective organizations but have also provided the right institutional framework for assisting African countries with statistical development. We would like to commend the leaderships of the two regional institutions for these actions.

Also during the period covered by the two quarters, the AfDB conducted four sub-regional workshops for countries in the Common Market for Eastern and Southern Africa (COMESA), English speaking countries of the Economic Community of West African States (ECOWAS), countries in the Southern Africa Development Community (SADC) and member states of the Observatoire Economique et Statistique d'Afrique Subsaharienne (Afristat). The purpose of the workshops was three-fold i.e. to officially launch the AfDB assistance to regional member countries to design or update their National Strategies for the Development of Statistics (NSDS), empower key actors in national statistical systems with knowledge about the NSDS processes and how to go about them, and to facilitate sharing of experience among countries. These workshops placed special emphasis on integrating sectoral issues and concerns in the design of the NSDS. Following the COMESA and ECOWAS workshops, an e-Discussion platform was established to exchange and share knowledge, experience and information on sectoral statistics.

Finally in September 2006, the statistical community and development partners in statistical development in Eastern Africa converged on Dar es Salaam, Tanzania, to celebrate 40 years of service to the region by the Eastern Africa Statistical Training Centre (EASCT).

Over these years, the Centre has trained many professionals some of whom have risen to head their national statistical offices and others are now in key positions in international organizations. We would like to congratulate the Centre on this exemplary service to the sub-region and the African region in general.

Dr. Charles Leyeka Lufumpa
Co-Chair

Prof. Ben Kiregyera
Co-Chair

The 2000 Round of Censuses in Africa: Achievements and Challenges.

Jason O. Onsembe[1] and James P.M. Ntozi[2]

Summary

During the last three decades, population and housing censuses have been the most critical source of data for policy decisions and for formulation, monitoring and evaluation of national and sectoral development programmes in Africa. This was achieved through active participation of many countries in the decennial census programmes, commonly known as census rounds. The 2000 census round can be described successful in most cases. However, a number of countries experienced implementation problems and challenges, including inadequate funding and weak technical expertise in national statistics offices. To ensure successful implementation of the 2010 census round, it will be necessary to develop strategies that would address the identified problems of 2000 round. For instance, African governments and the international community should commit adequate resources to statistical programmes, including censuses and enhance capacities of the national statistics offices, especially through formal training of staff and provision of the equipment and materials.

Key Words

Census round and census programme.

Résumé

Au cours des trois dernières décennies, les recensements de la population et de l'habitat ont été la source essentielle de données pour les prises de décision politique ainsi que l'élaboration, le suivi et l'évaluation des programmes nationaux et sectoriels de développement en Afrique. Ceci a été réalisé grâce à la participation active de nombreux pays aux programmes de recensements réalisés tous les 10 ans, communément appelés cycles de recensements. La phase 2000 du recensement a été réalisée avec succès dans la plupart des cas. Cependant un certain nombre de pays ont rencontré des problèmes et des défis de mise en œuvre, y compris des financements insuffisants et une faible expertise technique au niveau des instituts nationaux de statistique. Afin de garantir la réussite de la mise en œuvre du cycle des recensements de 2010, il est nécessaire de développer des stratégies tenant compte

1: Adviser in Population Data, UNFPA/CST, Addis Ababa, Ethiopia
2: Professor, Department of Population Studies, Institute of Statistics and Applied Economics, Makerere University, P.O. Box 7062, Kampala, Uganda.

des problèmes identifiés durant le cycle de 2000. Par exemple, les gouvernements Africains et la communauté internationale devraient allouer des ressources suffisantes aux programmes statistiques, y compris les recensements, et renforcer les capacités des instituts nationaux de statistique, en particulier par des formation du personnel et la fourniture d'équipements et de matériels.

Mots clés

Cycles de recensements et programme de recensements.

I. Introduction

Census taking in sub-Sahara Africa (SSA) can be traced back almost two centuries in early nineteenth century. Lorimer (1968) reported that sporadic counts of total population in certain trading centres and special colonies by colonial powers took place. For example, the British conducted a series of annual enumerations of Sierra Leone Crown Colony from 1833 to 1851 to check on the alleged kidnapping and re-enslavement of Africans. In the early period of twentieth century, census taking in colonial Africa was mostly head counting where heads of households used to assemble at administrative headquarters and to be counted for the purpose of estimating people eligible for hut tax. This happened in the former British East Africa (now Kenya and Uganda), former British West Africa (now Ghana, Sierra Leone and Nigeria) and the Belgian colonies of Burundi and Rwanda.

In the late period of colonialism, modern censuses (those based on modern procedures/rules) were introduced to SSA for planning purposes. The SAA countries affected included the British colonies of Uganda in 1948 and 1959; Kenya in 1948 and 1962; Tanganyika in 1957; Ghana in 1948; Nigeria in 1952-53; Southern Rhodesia (now Zimbabwe) in 1948, and Northern Rhodesia (now Zambia) in 1950-51 and 1963. In the French speaking Africa sample censuses were conducted in Guinea in 1954-55, the Congo (Congo Brazzaville) in 1955-56, Rwanda-Burundi in 1952 and Congo Leopoldville (now the Democratic Republic of Congo) in 1955-57.

Later, in the post-colonial period, the modern censuses, which required substantial technical and financial resources, were supported by the multilateral and bilateral aid agencies. For instance, the international community mostly through UNFPA almost wholly supported the African Census Programme (ACP) in 1960s and 1970s. Secondly, with support of UNFPA and other development partners a section of the Statistics Division of the Economic Commission for Africa (ECA) known as the Region-

al Advisory Services in Demographic Statistics provided technical support to African countries to conduct the censuses of population and housing in the 1980 and 1990 rounds of censuses. These census programmes provided demographic and socio-economic information that became the main reference materials for evidence based development planning in the region.

However, the census programmes encountered a number of problems. The problems included limited technical capacities in statistical offices, poor infrastructure especially lack of adequate offices and equipment, technology in terms of computers being at the formative stages, limited support from some African governments, inadequate legal framework to support and protect the census taking, largely illiterate and ignorant population and existence of anti-census taking taboos.

With the restructuring of ECA in mid 1990s, UNFPA sub-regional Country Support Teams (CST) based in Addis Ababa, Dakar and Harare assumed the full responsibility of providing technical and advisory services to the region in the 2000 round of censuses which ended in 2004 and ushered in the 2010 round.

The purpose of this paper is to present the experience of the 2000 census round of censuses in SSA. Specifically, the objectives of the paper are to document the achievements of the 2000 census round; identify challenges and constraints of the programme, particularly as they contributed to availability of less data; give the current population data needs in the region, especially in response to the emerging issues; and suggest strategies for conducting better censuses in the 2010 round.

2. 2000 Round of Censuses in Africa

The 2000 census round kicked off in 1995. Before that time, the United Nations Statistical Commission and the Economic and Social Council (ECOSOC), in resolution 199/57 recommended that all member states carry out population and housing census during the period 1995 - 2004. Table 1 shows that 36 out of 51 SSA countries participated in the 2000 round of censuses. The 15 countries that did not participate in the census programme were Angola, Burundi, Cameroon, Chad, Congo Democratic Republic, Djibouti, Eritrea, Ethiopia, Liberia, Madagascar, Nigeria, Sierra Leone, Somalia, Sudan and Togo. From the list, it can be noted that two thirds of the countries were engaged in conflicts and hence were not stable enough to plan and execute a census programme.

Most of the funding of the 2000 round was from the development partners. A selected list of 18 SSA countries that received assistance from various multilateral and bilateral donors can be seen in Table 2. As expected, UNFPA was involved in the censuses of all the countries that participated. UNFPA Country Support Teams (CST) in Africa provided the technical support to NSOs and other government departments in SSA that planned and conducted the censuses.

Table 1: Census dates for countries in Africa: 1965 to 2014

Country or area	1970 Round (1965-74)	1980 Round (1975-84)	1990 Round (1985-94)	2000 Round (1995-2004)	2010 Round (2005-2014)
Algeria	4/4/66	1/12/77	4/20/87	6/25/98	--
Angola	12/15/70	--	--	--	2007 S
Benin	--	3/20-30/79 E	2/15/92	2/11/2002	2012 S
Botswana	8/31/71	8/12-26/81	8/14-23/91	8/17-26/01	2011 P
Burkina Faso	--	12/1-7/75 E	12/10-20/85	12/10-20/96	--
Burundi	--	8/15-16/79 E	8/15-16/90	--	2005 S
Cameroon	--	4/9/76 E	4/14-28/87	--	2007 S
Cape Verde	12/15/70	6/1-2/80	6/23/90	6/16-30/00	2010 P
Central African Republic	--	12/8-22/75 E	12/8/88	12/8/2003	2013 S
Chad	--	--	4/15/93 E	--	4/05 S
Comoros	7/66-9/66	9/15/80	9/15/91	9/03	2013 S
Congo (Republic of the)	2/7/74 E	12/22/84	--	6/6-7/30/96 7/00 A	-- --
Congo (Democratic Republic of)	7/1/70 A 7/1/74 A	7/1/84	--	--	2005 A S
Cote d'Ivoire	--	4/30/75 E	3/1/88	11/21/98	2008 P
Djibouti	1970-71 A 3/17/67	1/3/83 E	--	--	--
Egypt	5/30/66 SC	11/22-23/76	11/17-18/86	11/18-19/96	2006 P
Equatorial Guinea	9-10/71 A	7/4-17/83	9/9-10/5/94	2/02	--
Eritrea	--	5/9/84 E (by Ethiopia)	--	12/04 S	2014 S
Ethiopia	--	5/9/84 E	10/11-27/94	--	2007 S
Gabon	6/1/69-6/70	8/1-31/80	7/1-31/93	12/31/2003	2014 S
Gambia, The	4/21/73	4/15/83	4/15/93	4/15/2003	2013 S
Ghana	3/1/70	3/11/84	--	3/26/00	2010 S
Guinea	5/19-27/67 A 12/31/72 A	9/28/77 A 2/4-17/83 E	--	12/1-15/96	--
Guinea-Bissau	12/15/70	4/16-30/79	12/1/91	12/01	2011 P
Kenya	8/24-25/69	8/25/79	8/24/89	8/24/99	2009 P
Lesotho	4/14-24/66	4/12/76	4/14/86	4/14/96 5/13-31/01	--
Liberia	2/1/74	2/1-14/84	--	10/03 S	--
Libya	7/31/73	7/31/84	--	8/95	--
Madagascar	--	1/26-8/18/75 E	8/1-19/93	--	2005 S
Malawi	8/9/66 E	9/20/77	9/1-21/87	9/1-21/98	9/08 S
Mali	--	12/76 E	4/1-30/87	4/17/98	--

Country or area	1970 Round (1965-74)	1980 Round (1975-84)	1990 Round (1985-94)	2000 Round (1995-2004)	2010 Round (2005-2014)
Mauritania	--	12/12/76-3/77 [F]	4/5-20/88	11/00-3/01	--
Mauritius	6/30/72	7/2-3/83	7/1/90	7/2-3/00	2010 [P]
Mayotte	9/1-30/66	7/1/78	8/5/85 8/9/91	8/5/97	--
Morocco	7/20/71	9/3-21/82	9/4/94	9/04	--
Mozambique	12/15/70	8/1/80	--	8/1-15/97	2007 [S]
Namibia	5/6/70	8/26/81	10/21/91	8/27-9/11/01	2011 [P]
Niger	--	10/7-11/6/77 [F]	5/10-24/88	5/20-6/18/01	2011 [S]
Nigeria	--	--	11/27-29/91		2005 [S]
Reunion	10/16/74	3/9/82	3/15/90	3/8/99	--
Rwanda	--	8/15-16/78 [F]	8/15-16/91	8/16-30/02	2012 [S]
Saint Helena	07/24/66	10/31/76	2/22/87	3/8/98	--
Sao Tome and Principe	9/30/70	8/15/81	8/4/91	8/25-9/8/01	2011 [S]
Senegal	--	4/16/76 [F]	5/27/88	12/8-22/02	2012 [S]
Seychelles	5/5/71	8/1/77	8/17/87 8/26/94	8/22-26/02	--
Sierra Leone	12/8/74	--	12/15/85	12/04 [P]	2014 [S]
Somalia	--	2/7-20/75 [F]	11/86-2/87	--	--
South Africa	5/6/70	5/6/80	3/5/85 3/7/91	10/09/-10/10/96 10/9-10/01	2006 [S] 2011 [S]
Sudan	4/3/73	2/1/83	4/15/93	2003 [S]	--
Swaziland	5/24/66	8/25/76	8/25/86	5/12-12/97	--
Tanzania	8/26/67	8/26/78	8/28/88	8/25/02	2012 [S]
Togo	3/1-4/30/70 [F]	11/22/81	--	--	11/05 [S]
Tunisia	5/3/66	5/8/75	4/20/94	4/28/04	2014 [S]
Uganda	8/18/69 11/74	1/18/80	1/12/91	9/13/02	2012 [S]
Western Sahara	12/31/70 1974	9/3-21/82 (by Morocco)	9/4/94 (by Morocco)	--	--
Zambia	8/22-30/69 8/26-9/6/74 [SC]	7/25/80	8/20/90	10/25/00	2010 [P]
Zimbabwe	4/21-5/11/69	8/18/82	8/18/92	8/18/02	2012 [S]

-- No census listed in this round.

date: date Annual counts between dates

[A] Administrative census.

[S] Scheduled; not yet taken or known if taken.

[F] First full modern census taken.

[P] Projected based on pattern of census dates.

[SC] Sample census.

[R] National registry

Table 2: Multilateral and bilateral agencies providing assistance to the 2000 round of census and large scale surveys (for selected countries)

Country	Census Year	Local sources	Agencies/countries participating
Benin	2002	Government	UNFPA, World bank, UNICEF, Switzerland,
Cameroon	2003	Government	UNFPA, World Bank, PTTE*
Cape Verde	June 2003	Government	UNFPA, French Cooperation, European Union, UNICEF Portugal CESD**
CAR	December 2003	Government	UNFPA, Japan, European Union, UNICEF
Comoros	September 2003	Government	UNFPA, WHO, UNDP, European Union, UNICEF
Cote D'Ivoire	November 1998	Government	UNFPA, World bank, France
Gambia	April 2003	Government	UNFPA, World bank
Ghana	March 2002	Government	UNFPA, DFID, USAID, JICA, China
Kenya	August 1999	Government	UNFPA, DFID, USAID, UNDP
Mali	April 1998	Government	UNFPA, UNICEF, European Union, World Bank, Canada
Mozambique	August 1997	Government	UNFPA, Netherlands, Sweden, Denmark, European Union, USA and Norway
Namibia	September 2001	Government	IDA, UNFPA, Spain, Britain and France
Niger	June 2001	Government	UNFPA, World Bank, European Union, UNDP, UNICEF
Rwanda	August 2002	Government	UNFPA, European Union, DFID
South Africa	October 2001	Government	UNFPA, SIDA, USAID
Tanzania	August 2002	Government	UNFPA, DFID, SIDA, Japan, UNDP, USAID
Uganda	September 2002	Government	UNFPA, DFID, NORAD, Japan, European Union
Zambia	October 2000	Government	UNFPA, DFID, CIDA, Japan, Netherlands, UNHCR, Finland Germany

Source : Onsembe and Hie (2004)

** CESD = Center European la statistique et Développement

* PPTE= Pays Pauvre Très Endettes

CAR = The Central African Republic

3. Achievements

In order to appreciate the achievements of the 2000 round of censuses, it is important to focus the achievements on whether the census programme satisfied the data needs of SSA. The data is needed for sub-national level decision making, monitoring of indicators of poverty eradication goals, measuring progress towards Millennium Development Goals (MDGs), and estimating demographic change.

3.1 Data requirements for sub-national level decision-making processes

Many countries in SSA have been decentralized to promote good governance and

give power and decision making to the communities by devolving the political and administrative power from the centre to the lower government units. This decentralization policy requires reliable data detailed to the lowest administrative unit for planning and decision-making, which is hard to get without complete census enumeration. The 2000 census programme, starting at the time the decentralization policy in SSA was taking root, collected most of the data required for the purpose and it is already being used in many countries with decentralized government structure such as Uganda and Tanzania. Another use of the population census data in promoting decentralization policies has been to help in making requests for assistance from international community by local government units.

3.2 Monitoring poverty reduction programmes

Poverty reduction programmes have been implemented in many SSA countries for sometime now. Implementation of these programmes needs to be monitored to know the progress of implementation and how much has been achieved. As can be seen in Table 3 giving an example of Tanzania, 14 out of 19 indicators (74%) of poverty reduction programme in the country were provided by the 2002 census. Similarly, the census programmes of different SSA countries contributed to the poverty reduction programme. This is a huge achievement for the 2000 census round in the region.

Table 3: Sources of data for monitoring poverty indicators in Tanzania

Type of poverty/Indicators	Source of data
Income poverty	
Headcount ratio	Labor Force survey
Food poverty line	Household Budget survey
Proportion of working age population not currently employed	Population and Housing Census
Human capabilities	
Girl/boy/ration in primary education	Population and Housing Census
Girl/boy/ration in secondary education	Population and Housing Census
Literacy rate of population aged 15+	Population and Housing Census
Net primary enrollment	Population and Housing Census
Gross primary enrollment	Population and Housing Census
Prevalence of ARI	MOH Statistics
Population with access to safe water	Population and Housing Census
Infant mortality rate	Population and Housing Census
Under-five mortality rate	Population and Housing Census
Life expectancy	Population and Housing Census
Children under five immunized	MOH Statistics
Births attended by skilled health worker	MOH Statistics/Demographic and Health Surveys

Extreme vulnerability	
Proportion of orphaned children	Population and Housing Census
Proportion of child-headed households	Population and Housing Census
Proportion of children in labor force not going to school	Population and Housing Census
Proportion of elderly living in a household where no one is economical active	Population and Housing Census

3.3 Measuring achievement of MDGs indicators

The United Nations set some 8 goals to be used as targets for socio-economic development by 2015 and called them Millennium Development Goals (MDGs). The measurement of the achievement of these national and international development goals/targets depends on availability of reliable data. Table 4 shows an example of Mozambique with the MDGs and indicators to measure the achievement of related targets in the country using the census data. The table indicates that the 1997 census in the country was the main source of data for seven out of eight MDGs, with exception of goal number 1 on poverty and hunger. The experience in the region indicates that census data is required for events or characteristics, which, because of their rarity, cannot be captured through sample surveys. For example, information on maternal and adult mortality are better collected from censuses than sample surveys. The use of the census data in Mozambique for MDG tracking has been repeated in the whole region.

Table 4: Sources of MDGs Indicators for Mozambique:

Goal	Indicators
Goal 2: Achieve universal primary education	Number 8: Literacy rate of 15 –24 year olds
Goal 3: Promote gender equality and empower women	Number 9: Ratio of girls to boys in primary, secondary and tertiary education Number 10: Ratio of literate women to men aged 15 – 24 years Number 11: Share of women in wage employment in the non-agricultural sector
Goal 4:Reduce child mortality	Number 13: Under-five mortality Number 14: Infant mortality rate
Goal 5: Improve maternal health	Number 16: Maternal mortality ratio
Goal 6: Combat HIV/AIDS, malaria and other diseases	Number 20: Ratio of school attendance of orphans to school attendance of non-orphans aged 10-14 years
Goal 7: Ensure environmental sustainability	Number 29: Proportion of population using solid fuels Number 30: Proportion of population with sustainable access to an improved water source, urban and rural Number 31: Proportion of population with access improved sanitation, urban and rural
Goal 8: Develop a global partnership for development	Number 45: Unemployment rate of 15-24 year olds, male, female and total

3.4 Measuring the achievement of indicators of population programmes

Since 1994 International Conference on Population and Development (ICPD) held in Cairo, almost all SSA countries have population policies, which are implemented through various population programmes. These programmes have targets to be achieved at specified dates. The 2000 census round data was used to estimate different population measures, such as fertility, mortality, migration, marriage and growth rates that are compared with targets of population programmes. Several evaluations have been conducted to measure the progress of ICPD nationally and regionally and the indicators computed from 2000 census data have been used. Hence, various population programmes in the region have been tracked, discarded or improved using indicators of the 2000 censuses. For example, in 2004, ten years after ICPD, UNFPA and ECA undertook an extensive review of ICPD + 10, where data from the 2000 round of censuses was extensively used to determine the progress made in the African region for a number of areas/issues including: adopting and implementing reproductive health approach, strengthening efforts to improve gender equality and equity, addressing adolescent reproductive health, and promoting integration of population dynamics and trends in development planning.

3.5 Meeting the demand of other emerging issues of socio-economic concerns in SSA

At present there are many issues of great concern globally, especially in SSA. These include ageing, migration, religion, ethnicity and race. Census data in SSA has been used to inform decision makers, planners and programme managers about these issues.

3.6 Other achievements of the 2000 census programme.

Besides meeting the data needs of SSA, the 2000 census round has benefited the region in several ways. First, the United Nations Statistics Division (UNSD) has helped create a strong culture of census taking by making available necessary manuals and handbooks. For example, the publication of the Principles and Recommendations for Population and Housing Censuses, Revision 1 in 1998, provided countries with guidance on the use of new developments and techniques, as well as accumulated knowledge and experiences of census operations. The majority of SSA countries have extensively used the document as a reference material for census planning and implementation.

In the 2000 round of censuses, more SSA countries than before used cartographic maps in their censuses and found them to be useful tools. The census cartography should sufficiently delineate the country into small, well-identified and homogeneously sized enumeration areas, which enhances the role of the census in promoting development of an integrated national statistical system. Over a period of time, cartographic base maps will also contribute to significant reduction of costs by sharing census cartography with other data collection activities such as agriculture census and national master sample for integrated programme of household surveys. In addition, a cartographic base will address the issue of small area statistics for which there is an ever-growing demand.

Another achievement was some use of Geographical Information System (GIS) technology in the 2000 round of censuses in sub-Saharan African countries, such as Uganda 2002, South Africa 2001, Namibia 2001, Tanzania 2002, Rwanda 2002 and Sierra Leone 2004. Although the use of GIS technology was limited in 2000 round of censuses, the fact that it was used at all was a great improvement since the censuses in the region before did not use it at all. GIS has advantages that sub-Saharan African countries have to benefit from, including reduced duration for enumeration areas (EA) demarcation, better control of EA demarcation, and less expensive and less time consuming EA demarcation exercises.

The fourth achievement of 2000 round of censuses in SSA was capacity building. During the 2000 Round of Census, a number of international community agencies including DFID, USAID, NORAD, UNICEF, WHO, UNDP, SIDA, JICA, DANIDA, the World Bank, European Union, CIDA, French Cooperation, Development Cooperation of Ireland, AusAID, Netherlands Interdisciplinary Demographic Institute, etc., assisted different countries in SSA in capacity building within the national statistical offices by providing necessary equipment and required technical support. With regard to staff capacity development, a number of approaches were deployed including undertaking technical backstopping missions at critical stages of census planning and implementation, organizing technical workshops for nationals, imparting skills through the on-job-training, and conducting formal training programmes. For example, for the Kenya census of 1999 and the 2002 Tanzania census, several senior professionals from National Bureaux of Statistics were sent to US Bureau of Census for training in data processing and analysis.

Another achievement was about the census products. For most censuses held in the 2000 round, statistical tables were produced within a relatively short time, within 12 months after enumeration as compared to the practice in the past. Analytical re-

ports, illustrated by GIS output and other new thematic areas, were available within two years of enumeration. User-friendly databases were also available and accessible to the general public later.

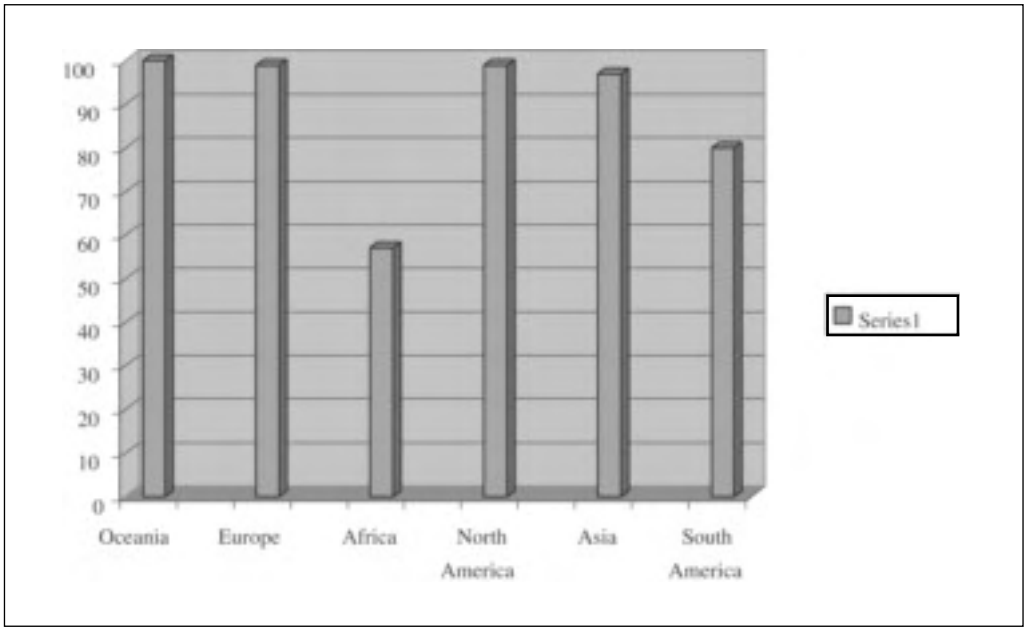
4. Limitations and Constraints of 2000 Round of Censuses in SSA

Although many sub-Saharan African countries successfully participated in the 2000 round of censuses, there were some countries, which, for one reason or another, did not participate. Furthermore, not all countries, which participated in the census, were able to carry out the exercise to its logical conclusion, namely; producing and disseminating census results. The planning and implementation of the 2000 round was therefore not without a number of challenges and constraints. These include: inadequate participation; weak organizational and managerial skills at NSOs; weak technical capacities; lack of sufficient government commitment; inadequate and untimely funding; non-conducive political environment and lack of gender data.

4.1 Inadequate participation in the census programme

The participation of sub-Saharan Africa in the 2000 round of censuses was the lowest in the world. Table 1 above shows that the participation of sub-Saharan Africa in terms of its population was only 57% of the total population in the region, much lower than the overall world participation, which was a huge 91%. Compared to other world regions: Oceania – 100%, Europe and North America – 99%, Asia – 97% and South America – 80%, the participation of sub-Saharan Africa was the lowest (Zewodi 2002). The wide difference between sub-Saharan Africa and other regions is clearly demonstrated in Figure 1.

Figure 1: Percentage of Population that participated in the 2000 Round of Census by Various Regions of the World



4.2 Weak organizational and managerial skills

Most censuses were planned and conducted by national statistics offices (NSO). These statistics offices (or other government departments) charged with the responsibility of planning and conducting censuses faced a number of challenges and constraints. The first constraint was lack of comprehensive programmes on the census. Some of the NSOs and departments charged with the responsibility of conducting censuses on behalf of the NSO did not plan the organization of the census well. For instance, there were no clear guidelines for implementation, leading to piecemeal activities with no clear vision for the whole process. There was a tendency of planning activities up to the enumeration day and excluding the important post-enumeration activities including data processing, analysis, report writing and dissemination. In other cases, censuses were only undertaken at the instigation of development partners when governments were not committed to the exercise.

Secondly, some countries lacked institutional capacity in terms of equipment and staffing to conduct the census. Since some NSOs do not maintain cartographic section during the inter-censal period, they found it difficult to revive the cartographic activities close to the census date and do a good job.

Thirdly, there are structural difficulties of census offices. A number of African countries do not consider census activities as an integral part of the national statistical system, but as a separate chain of operations. This had the effect of negatively affecting the coordination of the census work and reducing the commitment of staff of NSOs to the exercise. For instance, in a few countries such as Kenya, South Africa and Senegal where census cartography sections were permanent units within national statistics offices, mapping activities were completed on time leading to the enumeration to be done on schedule. In contrast, countries such as Tanzania, Uganda and Ethiopia which set up cartographic sections rather late could not complete EA demarcation in time and had to postpone enumeration dates from 1998 to 2002, 2001 to 2002, and from 2004 to 2007, respectively.

4.3 Inadequate funding

Censuses are the largest, most elaborate and costly data collection activity that any country can undertake, and there is a tendency for the costs to continue rising. In many countries, census expenditures were in excess of 10-15 per cent of the budget of Statistical Offices over an entire decade. For example, estimated census costs for the censuses of Burundi, Uganda, CAR, Ghana were US$ 3.5 millions, US$ 17 millions, US$ 3.5 millions and US$ 21, respectively. These costs excluded substantial indirect costs governments spent on paying salaries of permanent employees, use of government buildings and telephones, which are usually paid centrally and are difficult to separate for the census. Outside South Africa, these expenditures can not be afforded by sub-Saharan countries.

In sub-Sahara Africa, both national governments and the donor community provided census funding but the contributions did not satisfy the demand mainly due to a number of reasons including: increasing costs of censuses as a result of both inflation and sophistication of methodologies, rapid population growth, limited national budgets and waning international support. For example, due to rapid population growth, the recently completed census in Nigeria of 2006 is estimated to have cost more than US$ 300 million, compared to US$ 88 used on the 1991 census.

Technical co-operation and assistance from international agencies and the donor community played a major role in providing funds for the 2000 round. With census donor fatigue, it is not guaranteed that this massive support from development partners will continue. Recent unsuccessful experience of fund raising for the UNFPA CST regional survey shows that funding the next round of censuses may be a problem.

4.4 Non-conducive political environment

Weak political commitment and lack of appreciation of the usefulness of census data has contributed to the apparent lack of support by the national government. For example, the Mozambique Government only became actively aware of the importance of the census after the devastating floods that engulfed the country in 2000. Although there are general claims that governments do support censuses, actions indicate otherwise. One way of measuring such support is through assessing census budget allocation. It has been observed that in many countries in the region budget support to the census programme is quite low compared to other national exercises, such as voter registration and parliamentary elections. For instance in Uganda, while the government contribution to the budget of the census exercise was less than US$15 million, the recent general elections consumed several times this figure.

4.5 Weak technical expertise in NSOs

Technical assistance from multilateral donors led by UNFPA and bilateral donors, played a major role in the success of censuses in sub-Sahara Africa in the last three to four decades. For instance, recent experience has shown that technical involvement of regional advisers (UNFPA/CST) has contributed effectively to the development of national technical expertise through on-the-job training in census organization, cartography, data processing, sampling and advocacy. After providing such technical assistance for a long time, the donors expect the African countries to be self-sufficient. However, given the large staff turnover over in most NSOs, there is continuing need for such assistance even for the near future.

4.6 Lack of data for gender indicators

Managing gender issues is very critical to the development of sub-Saharan Africa. To do it effectively and efficiently, indicators of gender equality and equity have been developed and they need appropriate data for measuring them. The data from the 2000 round of censuses does not satisfy most of the needs of indicators on gender. Except for female illiteracy rate and ratio of girls to boys in primary and secondary schools, other indicators of progress on gender equality and equity, namely HIV/AIDS prevalence among the female population, gender empowerment, seats held by females in national Parliaments, and gender-based violence could not be measured from the 2000 round of census data.

5. Conclusions and Proposed Strategies for 2010 Census Round

As indicated in earlier sections of this paper, there were commendable achievements in the 2000 census round, which ensured adequate availability of demographic and socio-economic data in a number of countries. On the other hand, there were challenges and constraints, which imply that there is room for improvement. With this wealth of experience, organizers and planners of the 2010 census programmes should use strategies that should build on the strength and minimize the challenges and constraints of the 200 census round. The strategies should adress the following:

5.1 The need for government commitment to census programmes

It is important that census organizers should get commitment of government before starting to plan a census programme. Although most governments appreciate the importance of up-to-date and reliable census data in the overall development planning, they lack serious commitment. Both political and civil service leaders at all levels of government must be consulted and educated on the benefits of census data to their work through an advocacy programme and their commitment secured. Census legal frameworks in form of statistical laws must be strengthened and updated to smoothly facilitate the implementation of the census programme. Governments should also commit funding to the programme to ensure that the activities are financed up to the end, before inviting the development partners to supplement their funding commitments.

5.2 Basic organizations/administrative aspects of the census

Lessons learned from the 2000 census round show that there were several organizational and administrative problems in the conduct of censuses across the region. The NSOs in the region should, therefore, put in place comprehensive census organizational structures that will ensure successful implementation of the programme. Countries with ad hoc cartographic units should set up permanent units to ensure that census mapping is fully done and on time. It is important that census cartography and use of GIS are promoted and encouraged because they facilitates accurate enumeration of populations by better demarcating the enumeration areas (EAs). Furthermore, NSOs will need to strengthen collaboration with development partners

and identify one lead development partner to coordinate the rest of the development partners in mobilizing funds. Pooling of donor resources, which is a cost-effective strategy for meeting the diverse demands expected in these countries and ensuring that the funds are used for the intended purpose, should be emphasized.

5.3 The need to satisfy technical requirements in census planning and implementation

Lessons learned indicate that the greatest bottleneck of the census programme is in the delay of release of census results, which is related to data processing. New technologies that could improve the speed in data capture and reliability in data processing should be introduced in all census operations. It is expected that such technologies will also reduce the number of staff to be involved in data processing activities. Recent innovations in this regard include: the use of OMR/OCR for scanning, the application of computer assisted codification techniques, the use of local computer network in order to increase the storage capacity and the distribution of data files. The use of new technology to scan the raw data and input it direct into computers in Tanzania reduced the period between the census enumeration (September 2002) and release of full results (March 2003) from several years in the past to only 6 months.

Another important strategy is to heavily involve the analytical team members from the conception of census operations to the end. These specialists should participate in drafting census questionnaires and formulating analytical plan for data analysis and write-up.

5.4 Effective census data utilization

Although censuses are expensive, in relative terms, they are not expensive in terms of the potential value. There is concern that censuses are often not fully utilized because the data is not always analyzed and turned into information for use by policymakers. One reason why census data is not fully utilized in some African countries is because NSOs do not have effective dissemination strategies in place. UNFPA and other international agencies are paying increasing attention to the need of supporting the use of census data for poverty mapping and promoting thematic policy dialogue.

There is need to make more extensive use of data for national planning and development, including the use of census data for small-area estimates for the benefit of local governments. This means that NSOs should conduct training workshops for local government and community officials on how to use such data without breaching confidentiality.

There are also problems associated with the use of census data for non-statistical purposes, for example for election voter registers and updating of administrative registers. While these demands for census data should be met, NSOs need to review the procedures on what data to release to the national electoral commissions and administrative registrars without compromising the confidentiality of individuals.

References

Dackam R. (2006), Paper Presented at the International Meeting Towards the Successful Implememation of the 2010 Round of Population and Housing Census in Developing Countries, New York, 24 – 25 February 2006.

Dackam R. 2002. Role of Census Data in Monitoring National Poverty Targets and Tracking the Millennium Development Goals in Population Census, New Directions and Cost Saving Strategies In-house Capacity Building Meeting (21 – 23 October 2002)

Kent, M. Mary, Kelvin M. Pollard, John Haaga and Mark Mather, 2001. First Glimpses from the 2000 U.S. Census, Population Bulletin, Vol. 56, No. 2.

Lorimer 1968. The Present Situation of Demography in Africa South of the Sahara in Brass W. A.J. Coale et. al. (editors): The Demography of Tropical Africa. Princeton University Press.

Onsembe J.O. and Hie J-M, Nature and Magnitude of UNFPA's Assistance to Statistical Development in Africa. Paper Presented at the First Forum on African Statistical Development, Addis Ababa, 12-13 May 2004

UNFPA 2000. Report of the Joint Interagency Coordinating Committee on Censuses for sub-Sahara Africa and PARIS 21 Task Force Meetings. Eurostat, Luxembourg, 26-27 October 2000.

UNFPA 2000: A Quick Survey on the 2000 Census Round, in Background Paper for the International Meeting on Advocacy and Resource Mobilization for the Successful Implementation of the 2010 Census Round, New York, February 24-25, 2005- Nott.

UNFPA 2000: Funding Crisis in the 2000 Round of Population Censuses; UNFPA/ PARIS21 International Expert Group Meeting Mechanisms for Ensuring Community of 10- Year Population Censuses; Strategies for Reduction of Census Costs.

United Nations 2003: World Population Prospects, the 2002 Revision , Volume 2002, vision Volume 1, Comprehensive Tables, United Nations Publications , Sales, No. E.03, New York, 2003.

Zewodi Y. 2002. Utilization and Analysis of 2000 Round, a Paper Presented at 20th Population Conference, Ulaabaatat, Mongolia, 21 October 2002.

Institutional cooperation for statistical development: the case of twinning in Mozambique – a model for other countries?

Hans Erik Altvall[1] and Ms. Destina Uinge[2]

Summary

Donor support to statistical development has different designs. Donors with different approaches can be involved in projects to the same statistical bureau in the same country. This often creates problems in coordination with implications for a consistent, systematic and sustainable statistical development.

The ongoing Scandinavian project to strengthen the National Statistical System (SEN) in Mozambique attempts to avoid these risks by joining the three Scandinavian donors in co-financing one and the same institutional cooperation and under a common project set- up.

The cooperation is based on a twinning arrangement between the National Statistical Institute (INE) and a consortium of three Scandinavian national statistical bureaus. INE has the overall responsibility for the organization, management and administration of the project and to initiate, control, run, implement and review project supported activities. Project activities are integrated with INE's regular plans and reports.

Key Words

Integration; umbrella; cooperation; synergy; harmonization

Résumé

L'assistance de donateurs en matière de développement de la statistique s'opère de différentes façons. Des donateurs ayant des approches différentes peuvent se retrouver engagés dans des projets au sein d'un même institut de statistique dans un même pays. Ceci crée des problèmes de coordination qui ont un impact sur la cohérence, la régularité et la durabilité du développement statistique.

1: Hans Erik Altvall, Team Leader, Scandinavian Program, 01-05; Statistics Sweden. hanserik.altvall@scb.se
2: Destina Uinge, Project Director, Scandinavian Program, National Statistical Institute (INE), P. O. Box 493, Maputo, Mozambique. destina.uinge@ine.gov.mz

Le projet scandinave de renforcement des capacités du Système Statistique National en cours au Mozambique vise à éviter ces risques en s'associant à trois bailleurs de fonds scandinaves pour co-financer une et même coopération institutionnelle dans le cadre d'une mise en œuvre commune.

La coopération se base sur un jumelage de l'Institut National de Statistique (INE) à un consortium de trois instituts nationaux scandinaves de statistique. L'INE a la responsabilité générale de l'organisation, la gestion et l'administration du projet et initie, contrôle, dirige, met en œuvre et passe en revue les activités du projet. Les activités du projet sont intégrées dans le programme et le reporting régulier de l'INE.

Mots clés

Intégration; Parapluie; Coopération; Synergie; Harmonisation

1. Introduction

Various modalities are practiced in donor-supported activities for statistical development. For many years, the main modality was technical assistance to a specific project such as a labor force survey with little if any allowance for related support to formal training activities or to activities of crucial importance for institutional development, for instance management, accounting or organizational issues.

Although nowadays there is more of combined support, it is still rather frequent that one donor or a small donor group supports one or a couple of specific projects in a national statistical bureau. One example is that a country receives donor support to cover the various requirements that are necessary to carry out one specific survey. Procedures for the survey are created without considering that these should be seen as an integrated part of other activities within the same national statistical bureau and hence with the institutional development of the agency.

Different donors may support the same bureau. This means that several donors are involved in the statistical development in the same country. The recipient institution may have to deal with different partners with different working practices, work processes and methodologies with an evident risk for disharmony and inconsistency in the institutional development.

In order to achieve improved harmonization and consistency in the supporting activities and to reduce the need for support from many donors, the concept of institu-

tional cooperation based on a twinning arrangement was created. This means a direct cooperation between two institutions, which is opened for any area or activity of importance for statistical development, subject matter areas as well as issues on management, administration and procedures for budget and accounting. It also implies that technical assistance is integrated with capital assistance and with support to formal long and shortterm training.

2. The Philosophy of Institutional Cooperation

The principle idea behind institutional cooperation is to make use of relevant parts of the competence, methodology and institutional arrangements, work processes, etc. in a developed institution to support another institution in its development. The support should focus on activities that can be sustained and be further developed after the end of the cooperation.

The two institutions have in general similar professional references and areas of responsibility with similar demands on knowledge and competence. This promotes a good and natural climate from start. The lessons learnt in achieving and maintaining institutional competence at the more developed institution can be used to assist in developing human resources of the other institution.

An ideal twinning arrangement covers all areas of work and does not leave any area or any staff outside. It is especially important to include horizontal functions, such as management, systematic quality work, accounting, etc. since these are crucial for the institutional work and development.

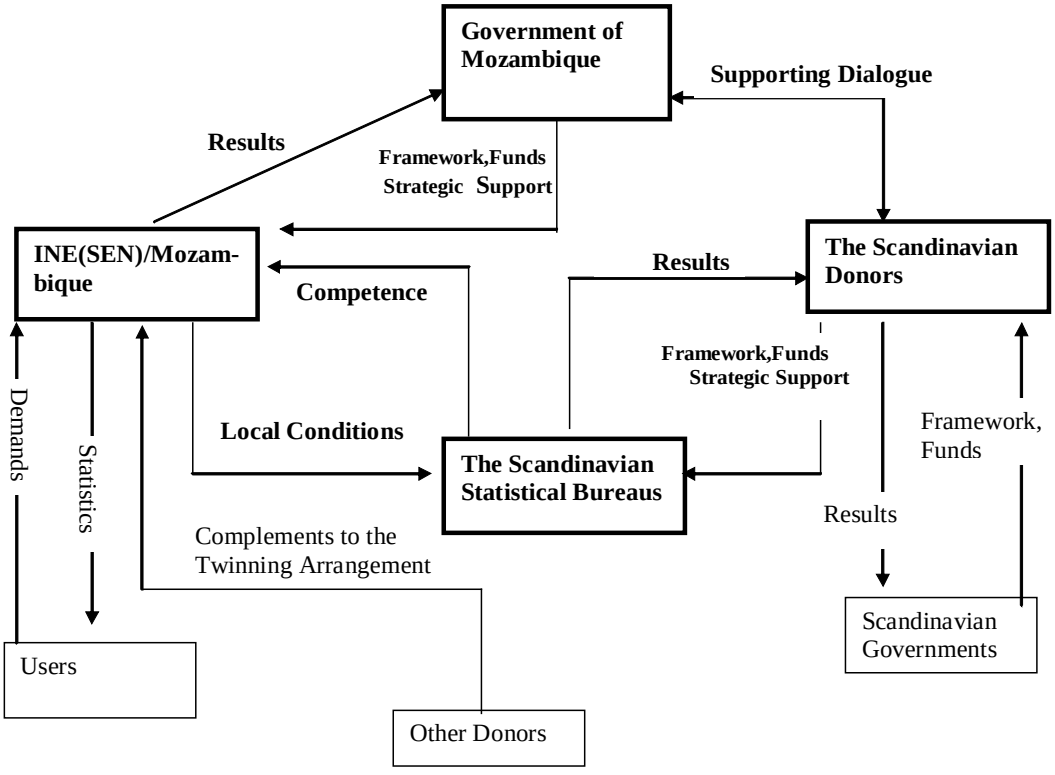
The donor support to this kind of cooperation must be of a long-term nature to enable the two partners to thoroughly and systematically plan for lengthy activities and reduce the need for ad hoc crisis intervention. It provides an important sense of security in the long-term plans and reduces the risk for involving other donors and other technical assistance with different working practices, methodology and work processes.

Long-lasting trusting relationships between the respective staff can be created for the promotion of long distance contacts, which can easily continue after the termination of formal cooperation.

3. The Actors in an Institutional Cooperation

A twinning arrangement has, apart from the two participating institutions, two other major actors, all with evident and heavy responsibilities for a successful cooperations can be seen in Figure 1.

Figure 1. Actors in the Institutional cooperation between INE (SEN) / Mozambique and the three Scandinavian Statistical Bureaus



One major actor is the Government of the supported institution. The Government sets the objectives, designs the framework and decides the plan for the development of the institution. This warrants that the development strategy is consistent with strategies for related institutions and for the country.

The Government also provides the local input for the institution in its development. It must recognize that the cooperation constitutes a process during which the insti-

tution develops a new and dynamic role with increased efficiency and expectations, which generates new development. This development normally implies an increase in the national budget and a change of authority and institutional arrangements.

The same Government has a key role in particular during the phasing out and termination of the donor support to ensure that there is sufficient local input to maintain the achieved sustainability for further institutional development.

The other major actor is the donor (and it's Government). This sets another framework for the cooperation in terms of objectives and length of the support and the scope of the finances. A donor that supports a twinning arrangement must in its contacts with other donors be in the front-line for the coordination of donors so as to promote that other support, if any, complements and not create disharmony in the twinning arrangement.

4. The Case of Mozambique

In 2002 a project started on Scandinavian assistance to strengthen the institutional capacity of The National Institute of Statistics in Mozambique (INE). The project is a continuation of previous support from Scandinavian countries to INE's institutional development, though using a rather different modality. The new project joins the three Scandinavian donors and the three national statistical bureaus in these countries in the same project that forms an umbrella for all donor support to INE.

The project seeks to strengthen the National Statistical System (SEN) in Mozambique through support to INE and its provincial branches so that it will efficiently generate reliable, relevant and timely statistical information. The development objectives are to support the implementation of SENs 5-year plan 2003-07 – guided by a strategic plan for the same period-and thus enable government, private sector, researchers, donors and the general public able to monitor social and economic development in the country. One specific focus is the implementation of the strategy of Mozambique for reducing poverty.

The High Council for Statistics that is chaired by the Prime Minister decides both the strategic and the operational plans.

The project budget for 2003–07 is 10 680 000 EURO that corresponds to about 50% of INEs total budget for the same period. Forty one percent of the Scandina-

vian support is for long and short-term technical assistance, 47% for surveys, operational activities and equipment and 12% for formal and in-house training courses.

The project is based on a twinning arrangement between INE and all the three Scandinavian national statistical bureaus. A consortium of the three bureaus is supporting INE in implementing the project with Statistics Denmark as the lead partner.

The project seeks to have a set-up which puts special emphasis on:

- Donor coordination
- A Twinning arrangement with three national statistical bureaus
- INEs role as fully responsible for the project activities
- Integration of project activities with INEs regular plans and reports
- Implementation of a comprehensive human resource development plan 2003-07

5. Donor Cooperation

INE is supported by a fairly large number of donors with an evident risk for lack of harmonization and coordination in the support that may jeopardize a systematic institutional development of INE. In order to reduce these risks, to achieve synergy and to increase the efficiency and effectiveness of their support, the three Scandinavian donors – DANIDA, NORAD and SIDA - agreed to co-finance a single project. This is by far the largest and most important support project to INE and serves as the umbrella for all donors to INE so as to avoid cross-purposing support.

The three donors have a joint responsibility for their total support and each provides about one third of the funds. They all have their own coordinated bilateral agreements with Mozambique on their respective contributions. DANIDA retains its part in Denmark to cater for the costs of technical assistance. The other two donors transfer their contributions to INE to support a number of subprojects that are outlined (together with appropriate budgets) in the Project Document.

DANIDA is the lead donor and ensures regular contacts and coordination between the three donors, with INE and with the Consortium. This implies that INE can conduct all its business through one Scandinavian donor without needing to spend time and resources to discuss with the individual donors. This has proved to be of a significant advantage from the viewpoints of both INE and the Consortium.

DANIDA in particular should also have a key role in contacts with other INE donors so as to promote the Scandinavian support to INE and the umbrella concept. It is difficult to conclude if this has had any specific advantage so far.

On the other hand it is of course INE itself that has the overall responsibility for all support and to ensure necessary coordination under the umbrella of the Scandinavian support. It is the responsibility of INE to ensure transparency in all matters connected with the Scandinavian Program and to keep all partners informed about other donor initiatives.

One activity at an overall level in this direction is that the major donors attend meetings with two Steering Committees, one for the Scandinavian Program and one for the Italian Program. The Italian program, covering the period 2004-05, is designed to complement the support from the Scandinavian Program to all INEs offices in the Provinces. The transparency for day-to-day activities is improved by an efficient use of Intranet.

Discussions during the preparatory stage amongst the various partners regarding the format for the contract and the modalities for the execution of the project were extensive and very time consuming.

6. TwinningArrangement with 3 National Statistical bureaus in a Consortium.

The three Scandinavian bureaus have very similar statistical systems, which is the foundation for the Twinning arrangement in the project. The three bureaus have formed a consortium that provides technical assistance and support to INE to implement project activities.

The main contract includes a specific agreement that formalizes the consortium, and outlines its obligations and role in the project; it specifies for instance how the recruitment of advisers is done.

The consortium has a coordination office in Statistics Denmark – the lead agency – that is the focal point for general backstopping and home office management of the project. The lead agency is the consortium's representative towards DANIDA (the lead donor), INE and other donors. It has regular contacts and meetings with the other two institutions in the consortium.

The consortium monitors the performance of the project through an internal project-monitoring group. The objective of the group is to ensure relevance, unity, quality, co-ordination and momentum of the project. Its task is to observe and continuously examine the quality of performed work.

There are also separate external evaluations of project achievements as well as regular external financial audits.

The consortium is represented in Mozambique through a Team Leader (TL), who co-ordinates the technical assistance from all the three bureaus. The TL is the spokesperson for the consortium in discussions with INE and with the donors and is also the main link between INE and all the three Scandinavian bureaus.

The TL lends strategic support to the development of INE towards a modern statistical bureau. This requires that the person is highly knowledgeable about modern statistical systems as well as institutional performance.

With three bureaus, a broader and stronger platform is provided for recruitment of long and short-term consultants than would be the case in twinning projects with only one bureau. The project has, apart from the TL, two to four consultants contracted from all the three bureaus on long-term assignments of at least 2 years.

The advisers on short-term assignments of two to three weeks are usually from the Scandinavian bureaus but may also be persons with relevant skills from outside or from other countries. Hence, the consortium and INE make efforts to establish contacts with an even broader resource base relevant to INE.

An important component in the Twinning arrangement is that the project has strategic support from the top managers and the Directors General (DG) at the three Scandinavian bureaus. The DGs have visited INE to transfer to the INE managers their experiences from Management and Strategic Planning and from International Statistical Cooperation.

Another principal component is study tours to the three Scandinavian bureaus. The tours broaden and deepen the contacts and are of specific importance for areas that are crucial for institutional development. Hence, there have been tours, for example, on systematic quality work and on strategic aspects concerning competence and information technology (IT), respectively.

The experience of the three bureaus from the set-up is positive; this is especially the case when advisors are given the necessary language training and have had an opportunity to become familiar with the new working environments and local conditions. INE's contact net has also become broader and deeper. The knowledge of and interest for statistics in Mozambique is now more spread and deepened through the web pages at the Scandinavian bureaus, the many advisors to INE and the study tours.

7. Project Setup at INE.

INE has the overall responsibility for the organization, management and administration of the project, charged with the responsibility to initiate, control, run, implement and review project supported activities.

This is a very important basis of the cooperation and the Twinning arrangement. It is also essential for integrating supported activities with INE's plans and regular activities and for supporting activities that can continue and be further developed after the termination of the project. However, it puts specific requirements on INE.

INE has a Project Director (PD) who is the Director for Integration, Coordination and External Relations, including planning, IT and dissemination. The contract between the Partners includes Terms of Reference specifically for the Project Director. The Project Director acts on behalf of the President of INE and has weekly discussions with the President on achievements and problems in the project.

The Director for Administration and Human Resources serves as Deputy Project Director. The Deputy in particular keeps an eye on issues of administration and accounting. Other key persons are a full-time senior accountant and a full-time administrative assistant both supported by the donor.

The Project Director is overall responsible for the day-to-day work assisted by the Team Leader, a Steering Committee and an internal Coordination Group.

The PD has weekly meetings with the *Project Management Group* that consists of herself, the deputy PD, the accountant, INE's training coordinator and the TL.

The Steering Committee includes the 3 top managers at INE, with the President of INE as chairman, all the three donors, the Project Director and the TL. It promotes an

efficient management and follow-up of project activities. It could be argued to have only the lead donor in the Committee to represent all the three donors. The Committee discusses achievements and accounts, can adjust planned activities and can reallocate funds. Other major INE donors are invited to the meetings to promote donor coordination. There have been discussions to reformulate the ToR for the group to enable all main donors to attend the meetings on equal terms.

The Committee convenes at least 3 times annually – in November to discuss plans for the following year, in April to discuss reports for the previous year and in September to discuss the mid term report.

The internal *Coordination Group* is chaired by the Project Director and includes all the Directors at INE – a level one step lower than top management – INE training coordinator, all Long Term Advisers (LTA) and the Project Accountant. The purpose of the group is to promote an efficient and smooth coordination and implementation of Project Activities and their integration with INE's regular activities. It usually meets once a month.

The group is also crucial for deliberations on activities that are common for all the Directorates at INE and strongly related to INE's institutional development such as IT, quality, planning, competence, etc.

The project set-up functions reasonably well. However, taking into account the scope and magnitude of the Scandinavian program, one of the major issues of concern is the workload of the Project Director. Given that the INE Director covers perhaps the most difficult area at INE, namely the major horizontal functions that are crucial for INE's institutional development, the combination of this with the responsibility of overseeing the daily activities within the Scandinavian Program, is perhaps a little too much. In fact, it was envisaged during the planning of the project that the Project Director would have to work almost full-time with the Scandinavian Program.

8. Integration with INEs Activities, Plans and Reports

INE has the responsibility to implement the activities and monitor the achievements. The activities form a natural part of INEs activities, in terms of both current and development activities. They are integrated with INEs daily activities and with various plans and reports. This means that most staff members at INE profit from the support and have a role in the institutional development of INE.

All Directors and Heads of Departments are responsible in their respective fields for the day-to-day functioning of the project, assisted by the LTAs. The LTAs provide advice in relation to the implementation of the activities and their work must be well integrated with INE's activities.

The activities are discussed at the Directorate or a Department's regular monthly meetings. The concerned long-term adviser from the Scandinavian Program is invited to attend the meetings and the minutes are circulated within the program. This procedure was agreed between the partners but is not always followed, raising serious concerns regarding the transparency within daily activities.

INE's counterparts for a short-term mission (STA) prepare terms of reference for the mission according to agreed format. Work on quality and on competence development that are essential issues for institutional development are always covered by an STA. The recommendations from an STA must be well established within INE and be feasible to implement.

The counterparts have responsibility to follow up the results of a mission and should provide explanations in cases where a recommendation is not or will not be implemented.

The Directors at INE are responsible for plans and budgets for their respective areas. They prepare activity plans with detailed annual and quarterly budgets for about 20 different subprojects according to the Project document for the cooperation (PRODOC). The plans and budgets as well as the follow up of the results and costs are focal points for the meetings with the internal coordination group. Proposals on changes of plans and reallocations of funds are presented to the Steering Committee for action.

A specific issue of major concern in INE's institutional development and in the Scandinavian project is to improve management performance and cost awareness amongst the staff at INE and in particular amongst the various Heads of Departments. For this purpose, there is ongoing work to develop a set of simple indicators that will show the annual change in relation to quality, user satisfaction, staff satisfaction, cost and competence on product level. All these annual indicators must show improvements to reflect better management.

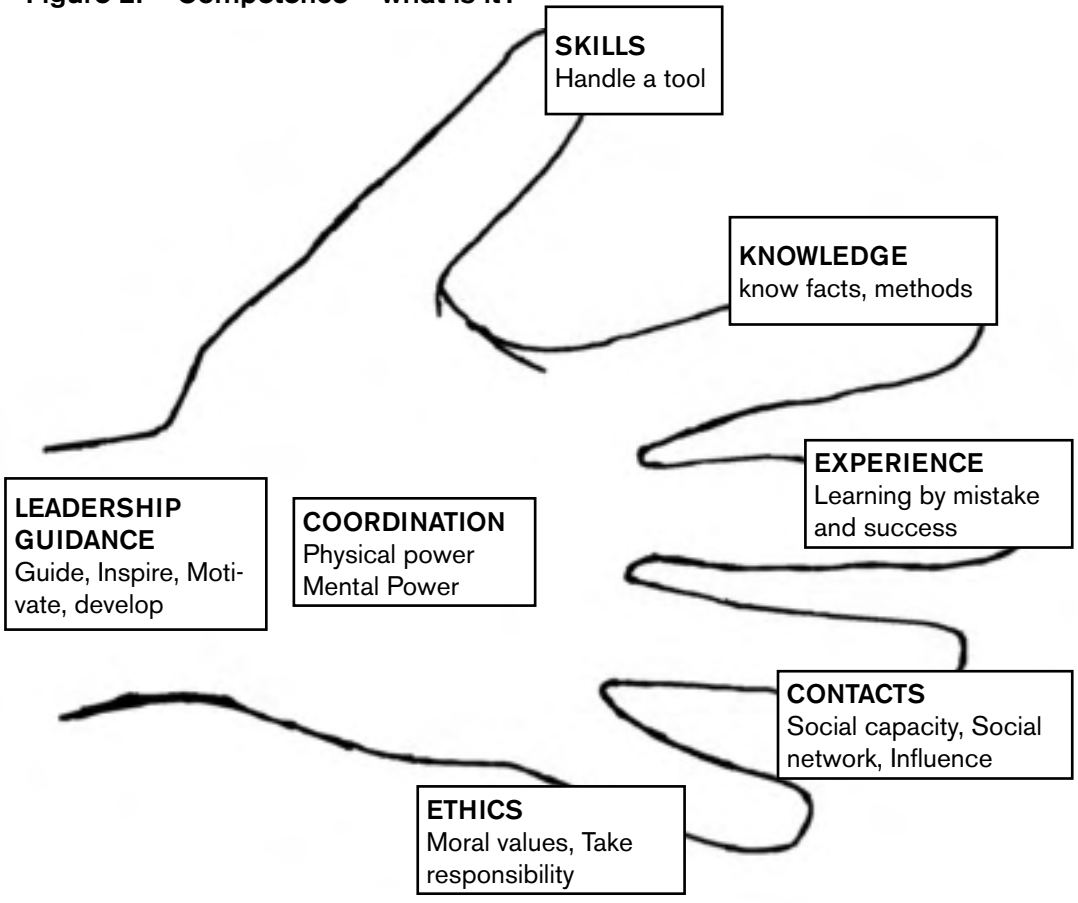
There are also specific courses for Heads of Departments in management and planning. Another issue which impacts on institutional development and institutional com-

petence concerns work on horizontal issues such as quality, in project groups that cross the line organization.

9. Competence Development Activities

The Scandinavian Program aims to transfer and develop competence at INE in order for INE, as a learning institution, to achieve a competence for self-development. The concept of competence is explained in Figure 2. This is done through various means, such as formal long and shortterm training, in-house tailor-made training, workshops, study tours, etc., and perhaps the most important means, on-the-job training with technical assistance.

Figure 2. Competence – what is it?



Source: The toolbox on Competence Development in Learning Organizations; Boliden Contech 2000; Illustration: Kerstin Keen

The competence concept covers 5 areas: Skills, Knowledge, Experience, Social Capacity and Ethics. The training places special emphasis on institutional competence and on issues that are crucial for institutional development.

The role of the consultants is to be advisers and contribute to competence development at INE. Hence, they should not be gap-fillers that do the actual work or present ready-made solutions. Instead they are competence developers that advise on methods or on procedures of problem-solving.

INE has developed a strategy for Human Resource Development (HRD) Activities to be implemented in 2003-07 with the assistance from the Scandinavian Program.

INE's training coordinator has a key role in coordinating HRD activities that are supported by the program with training activities that are supported by other donors. The role of this coordinator is also to administer assessments and evaluations of the training and to develop INE's own in-house training capacity with own trainers. One step in this direction has been to designate INE officer as assistant trainer to a consultant during an in-house training course with the aim to prepare the officer to take over the training later.

A matter of major concern is related to the vast training support to INE from several donors. This is difficult to coordinate and harmonize unless the training coordinator is consulted and has a clear mandate to co-ordinate all training activities. Experience shows that often new training activities take place on a very short notice without any information being given to the Scandinavian Program which is supposed to be the umbrella for all support.

10. A Concluding Remark

The three external reviews that have been carried out all give a relatively positive picture regarding this initiative. The latest evaluation which took place as a mid term review in April 2005 was very positive in it's report dated 17 June 2005. It highlights the main reasons for the success as capacity building, long-term engagement, assured finance and quality. 0.

Experiencing the Semper Validation Software: Genuine African solutions for data validation within the International Comparison Program

Mathieu B. Djayeola and Roland Rittenau[1]

Summary

Since the first African participation in 1970, this is the first time an African institution has taken the lead to manage the International Comparison Program and provide support to more than forty countries in the region. This support covered the design of price survey instruments for price data collection and office editing tools for data validation at the national and regional levels. The development of the Semper Validation Software has provided user-friendly tools to assist countries in their price survey data validation activities. Beyond the technicalities of computer software, the Semper could be appreciated as a communication vehicle, a code, and a language between the national ICP teams and the regional coordinator. The main purpose is to question the quality of the data provided for international validation. An answer is provided within minutes with graphical tables on which action is required from the user.

Key words

International comparison - price survey - data validation

Résumé

Depuis la première participation africaine en 1970, c'est la première fois qu'une institution africaine est à la tête du Programme de comparaison internationale et fournit un appui à plus de quarante pays de la région. Cet appui a porté sur la conception d'instruments d'enquêtes sur les prix destinés à la collecte de données et sur des outils d'édition en vue de la validation des données sur le plan national et régional. La mise au point du logiciel Semper se situe dans le cadre d'un processus visant à fournir des outils conviviaux destinés à aider les pays dans leurs activités de validation des données issues des enquêtes sur les prix. Outre sa technicité en tant que logiciel informatique, le Semper peut être aussi considéré comme un moyen de communication, un code et un langage entre les équipes nationales du PCI et le coordon-

1: This paper was prepared by Mathieu B. Djayeola (b.djayeola@afdb.org), Statistician Economist and Roland Rittenau (r.rittenau@afdb.org), Principal Statistician, both members of the Regional Coordination Team of ICP-Africa, African Development Bank, Tunis, Tunisia.

nateur régional. Il sert surtout à tester la qualité des données en vue de leur validation internationale. L'utilisateur est assuré d'une réponse en quelques minutes dans un tableau à thèmes graphiques sur lequel il devra intervenir.

Mots clés

Comparaison internationale – enquête sur les prix – validation des données.

1. Background

The International comparison Program (ICP) is a global statistical initiative established to produce internationally comparable price level, expenditure values, and Purchasing Power Parity (PPP) estimates, with the objective of facilitating cross-country comparisons of economic aggregates and price levels for GDP and its sub-aggregates. The genesis of the ICP was an early recognition that measures of economic aggregates based on exchange rates do not reflect differences in price levels between countries, and as such they are unstable for policy decisions which in principle relate to volumes only, free of price and exchange rate distortions.

The prices collected for the ICP can be used for other comparative purposes such as poverty-specific price comparison[2]. Whatever the purpose, it is important that countries collect prices for products that are comparable both within the country and between countries. More generally, the quality of international price comparisons, like the quality of national consumer price indices (CPI), is highly dependent on the quality of the basic input data. The validation process described in this paper is designed to ensure that countries provide the ICP regional coordinator with good quality price data for a selection of comparable products.

This paper is concerned with the intra-country validation of prices collected for the ICP – that is, the validation of its price data that a country carries out before reporting them to the ICP regional coordinator. The process of inter-country validation of prices subsequently carried out by the ICP regional coordinator, and its supporting software are described elsewhere[3].

2: Based on a poverty-specific sub-set of products and a specific weighting pattern with the aim to recalculate international poverty lines into local currencies in a much more appropriate way than it is currently done.
3: See for example, Annex IV Quaranta editing procedure, Eurostat-OECD Methodological manual on purchasing power parities, Luxembourg, 2005

2. Terms of Reference for the Development of a Validation Software

The development of Semper stems from the need to verify, clean and validate price data in the context of CPIs and ICP. In most of the countries, this validation process is carried out manually or using non integrated and specialized tools. Therefore, there is no assurance of consistency in methods within a country or over time. This is compounded when comparing validation practices between countries.

The product list was the main ICP survey instrument provided to the countries. It was developed through an extensive participatory process that involved the regional member countries and the African Development Bank. To check the quality of prices collected for these products, the regional coordination team of the ICP-Africa entrusted Mathieu B. Djayeola, Statistician Economist, to develop a specific validation tool[4] according to the following terms of reference:

1. Verify the consistency between reference and observed unit of measurement types[5], e.g.: litre is incompatible with kilogrammes and vice-versa;
2. Check the relationship between observed quantity and unit of measurement, like entering 800 kilogrammes instead of 800 grammes;
3. Process and flag prices reported in different currency units, e.g.: South African Rand and its cents;
4. Generate average prices for each product;
5. Compute price dispersion indicators such as minimum, maximum and standard deviation;
6. Identify outliers and potential mistakes;
7. Analyze price escalation over time to control trends;
8. Identify wrongly reported or inputted price quotations: e.g. misplaced decimal point, confusing use of sub-multiple of currency unit, etc.

The Semper software was named after the Latin dictum, *Semper aliquid novi Africa affert*[6] to underline the novelty as well as the African peculiarity of the application. This Microsoft Excel Visual Basic procedure facilitates data validation at the country level. It aims at verifying the consistency of units of measure, quantities and prices, as well as whether price variations are within acceptable limits.

The quest for data quality and the increased awareness for product comparability have positive impact on national statistical data gathering systems, especially the CPI. As the ICP surveys cover a twelve-month period in Africa, a time series analysis

4: The software was developed under the overall supervision of Michel Mouyelo-Katoula, ICP-Africa Regional Coordinator, and the technical guidance of Roland Rittenau, Principal Statistician.
5: Unit of measurement types refer to weight, capacity, linear scales, units, etc.
6: "There is always something new out of Africa": Author unknown, recorded by Pliny the Elder.

tool was essential. Functionalities of the Semper Time Line are specifically relevant for CPI-type data validation.

What is the Semper validation software? What are the targeted tasks to be accomplished by the software? How does it operate? What are its specific features? What are the advantages and limits of this application? This paper elaborates on one year of experience with ICP data collection and validation in forty-eight African countries.

3. What is the Semper Validation Software?

The Semper validation software builds on the concept of price quotation, which, in the context of the International Comparison Program, includes the following minimum information:

1. To which group of products, or in ICP terminology, to which basic heading does the product belong? According to international classifications, a bottle of orange drink purchased in a city shop is different from the same bottle purchased for immediate drinking in a restaurant. From a national accounts point of view, these bottles of orange drink do not belong to the same classification category as they do not serve the same consumption purpose and the service element is different. For instance, in the context of the ICP, the bottle bought in a shop is a beverage pertaining to the basic heading "Soft drinks and concentrates" while the bottle bought in a restaurant relates to the basic heading "Catering in hotels and restaurants".
2. The identification of the product: A code is necessary for automatic processing of information related to products. The name of the product is needed for understanding what a specific product code refers to. This is also a classification issue but may also be relevant when making economic analysis of data collected in the field. All characteristics and modalities of the products are important identification factors. For example, "Tinned Pineapple" is fully defined by the following product specification: [Basic heading: Preserved and processed fruits; Product code: 025.07; Reference quantity & unit of measurement: 850 Milliliters; Product presentation: Tin; 800-900 Milliliters; Type: Pineapple; Form: Chunks; Juice: Syrup (water + sugar)].
3. The observed unit of measurement: It is necessary to ensure that the observed unit of measurement is identical or at least convertible into the reference unit of measurement. Example, if the reference unit of measurement for "Tinned Pineapple" is milliliter, the actually observed unit can be any multiple of a milliliter, whereas units of weight (such as gram, pound, etc.) will not be accepted.

4. The observed quantity: the quantity observed must be reported to estimate the unit price (i.e. related to the reference quantity) known as the "recalculated price" in the application.
5. The observed price.
6. The place of observation: the outlet, the location, the country where a price is collected are important price determining factors. The price of mineral water in a tourist area will not be the same as in a neighborhood shop of a residential area.
7. The time of data collection may also be critical for some products. Seasonal product prices vary over the months and prices of perishable goods may vary according to the time of observation in the day. For example, prices of fresh seafood may go down in the evening.
8. Comments and remarks can provide additional information on a particular price quotation, in terms of survey circumstances or deviation from the product specification.

The Semper validation software is an integrated application designed for office editing of field data collected in the context of the ICP-Africa. Its three-stage procedure is as follows:

1. Check survey constants: Conformity of reported product codes, name, quantity, unit of measurement, and other characteristics to the reference specification in the product list. Any product not included in the ICP-Africa product list is immediately rejected by the Semper.
2. Establish the dual mathematical relationship between observed and reference quantity and unit of measurement and effect necessary conversions to rescale observed prices.
3. Analyze rescaled product prices to identify outliers and potential errors.

4. How Does it Operate?

The Semper Validation Software analyses "*country data*" in the context of the ICP-Africa. It needs clean and standardized files, structured and verified database to run properly. The application first checks these requirements before processing. Like a factory, the Semper works with an Input-Process-Output system (Figure 1).

The Input – What the Semper requires.

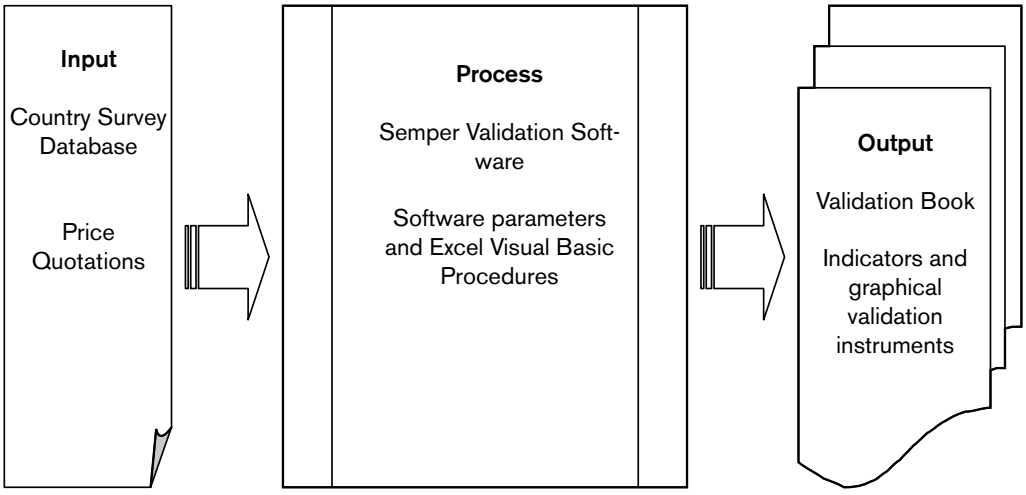
The Semper validation tool generates its input from the ICP-Africa Data Entry Sheets. The Data Entry Sheets are specific devices for data entry with a high degree of se-

curity, protection of cells and scroll functions to guide users during the data entry process, guarantee a homogenous structure of the data sets and avoid typing mistakes as far as possible. These sheets are typically structured by price collection center in a particular country and are relatively heavy in size, due to the big number of control functions. They are merged into one single input file referred to as "*Country Data*" through an automatic Excel macro[7].

The Process – What does the Semper do?

The Semper program file contains five worksheets and a built in Visual Basic module. The first sheet (the PS or Product Specification Sheet) contains general information on the product specification and individual data validation parameters such as the reference quantities, the reference unit of measurement and the acceptable outlier percentage. The second sheet (the UoM or Unit of Measurement sheet) contains information on compatible unit of measurement and subsequent conversion factors, the fourth sheet (the User Information Sheet) contains basic information on the way the price database is built and how the outlier indicators are computed, and the fifth sheet (the Comments Sheet) is reserved for the user's comments and remarks. The third sheet (the Program Info Sheet) is the core part of the Semper. It contains the Start button to launch the application and the processing log.

Figure 1: Structure of the Semper software



7: This macro named "*MergeCountryDataMacro*" is necessary for fast and automatic preparation of a national price database for validation.

The Output: What do you get out of the Semper validation software?

The Output (or "Validation Book") is a specific Excel workbook containing three sheets: (1) the "*Computing Area*" Sheet, (2) the "*Summary*" Sheet to store temporarily processed information, and (3) The "*Miscellaneous*" Sheet, meant to serve as user notebook (Table 1).

Using converted prices (to the reference quantity) for a particular product, Semper computes selected indicators for survey data analysis. These are:

(1) Averages prices
(2) Count of price quotations
(3) Minimum price
(4) Maximum price
(5) Standard deviation of prices
(6) Count of representative quotations
(7) Count of poverty relevant quotations
(8) Count of poor location quotations

Once all these indicators are processed, the Semper starts the individual data validation by generating, simultaneously, an outlier indicator for individual quotations and minimum - maximum indicator for a group of quotations related to each product. Then, it highlights the isolated quotations with a range of colors: no color when the price is close to average, yellow when the price is within an acceptable distance from the average, pink to indicate a likely outlier and red to highlight obvious outliers and errors (Table 2).

Users can save the output generated by the software in different names or specific locations for future consideration. Any attempt to corrupt the structure of the file or inappropriate database configuration can run into errors and stop the program.

In the following example (Table 1) we see information at the level of the overall price report (Grand_Aggregation): this specific country has reported 4732 price quotations. The overall average price is statistically irrelevant, but shows that no technical problem occurred, as all individual data could be treated.

One level below, the product level is "00104_Aggregation", referring to product 00104. In this example three products were observed: 00104, 00105 and 00107, all of them belonging to the first basic heading 001 – Rice. At the level of the product,

the table shows the number of price quotations (Price_Count) and the average price. For instance, for product 00104, this country reported 6 quotations with an average price of 553 currency units. All this information is called "recalculated", because the observed prices are recalculated to the reference quantity. The outlier indicator for product 00104 is red, as it is below - 30 meaning that the minimum price is much less than half of the maximum price quotation.

At the level of individual price quotations, index numbers identify each quotation easily in the country reports. This is important when correcting errors. Outlier indicators for individual quotations are highlighted when they are too far from the average price for the selected product.

Table 1: A screenshot of the validation book (computing area)

Index	Product Name [02]	Outlier_ Indicator	Requested Outlet type [03]	Product Code [01]	Recalculated Price_Count	Recalculated Price_Average	…
				Grand_Aggregation	4732	1292	
		-32		00104_**Aggregation**	6	553	
01225	Long-grained rice	8	Supermarket	00104	1	632	
01277	Long-grained rice	-4	Supermarket	00104	1	340	
03873	Long-grained rice	-16	Supermarket	00104	1	898	
04097	Long-grained rice	6	Supermarket	00104	1	452	
04140	Long-grained rice	12.5	Supermarket	00104	1	582	
04370	Long-grained rice	3	Supermarket	00104	1	416	
		30		00105_**Aggregation**	5	1740	
00687	Long-grained rice	9	Neighbourhood shop, Market	00105	1	1960	
00688	Long-grained rice	13	Neighbourhood shop, Market	00105	1	1666	
01721	Long-grained rice	9.5	Neighbourhood shop, Market	00105	1	1540	
02119	Long-grained rice	7.5	Neighbourhood shop, Market	00105	1	1470	
03661	Long-grained rice	6	Neighbourhood shop, Market	00105	1	2065	
		-30		00107_**Aggregation**	32	70	
00124	Medium-grained rice	8	Neighbourhood shop	00107	1	60	
00280	Medium-grained rice	8.5	Neighbourhood shop	00107	1	80	
00480	Medium-grained rice	1	Neighbourhood shop	00107	1	90	
00642	Medium-grained rice	11.5	Neighbourhood shop	00107	1	65	
00982	Medium-grained rice	13.5	Neighbourhood shop	00107	1	68	
…	….	…	…	…	…	…	…

For more detailed explanation, a user-guide is available.

5. Special Features

Unlike most of common Microsoft Visual Basic applications, the Semper builds on both internal and external procedures of the mother software – Microsoft Excel. Internal procedures are usual Excel actions such as formula, copy and paste tasks, as well as sub grouping procedures. Beyond the regular Visual Basic, Semper interacts with the operating system to gather information on language specific parameters: e.g. where does Excel store sub-group information when computing subtotal? In the original English Excel version, the sub-group identification is on the left of the string, while the identification is on the right in some French versions. By querying the computer system files, the Semper locates these information items and takes them into account while running.

As the application intends to collect all necessary parameters on its own, Semper runs specific procedures that fit the size of the price database. This re-scaling procedure aims at allowing the application to self-extract information on the number of price quotations, the number of products surveyed and the number of quotations per product.

Strict compliance to the common ICP data validation standards was the core objective of the Semper. The price database input to the Semper ("*Country Data*") is derived from the products list and the data entry form, which were originally compiled in strict compliance with the World Bank requirements for data transmission. The Semper has an extension for time series analysis called Time Line. An interesting point from a CPI perspective is that this component is capable of validating data over time.

6. Advantages

Semper generates a central matrix (located in the *Validation Book*) which can be used for further graphical data analysis: navigation, filtering and grouping. There are three levels of analysis in the Validation Book. The lower the level, the finer the analysis:

(1) The overall price report level,
(2) The product level, and
(3) The price quotation level,

It is easy to navigate through the Validation Book: the user can concentrate on one product, select a range of products or select price quotations within a product to be analyzed together. Semper also allows for the selection of price quotations on the basis of their outlier indicator value, the level of representativity and/or the relevance of the product to poverty analysis.

Fatal errors are clearly flagged with two specific codes as outlier indicator: the software creates a specific "outlier code" when either of the key input to the computation of the recalculated price (price, quantity, or unit of measurement) is missing [Code 9696]; and when, despite the availability of all these variables, the recalculation of prices is not possible [Code 9898]. This case occurs if the observed unit of measurement is not compatible with the requested unit of measurement (Figure 3).

The outlier indicator is set to positive values for "non problematic" price observations and negative values for "problematic" price observations. It is calculated following a formula, which is determined by the setting of an acceptable degree of price variation to the average for a particular product. These settings are variable and can be modified by the user. It is also possible to set different tolerance limits for different type of products to reflect the heterogeneity of the markets in the validation procedures. The most problematic items are attributed a higher number in absolute terms. This way, Excel filters will highlight these items and direct the user to them before showing less problematic price observations and those for which the Semper did not identify any anomaly.

The color-coding of outliers has proved to be very practical and is highly rated by most users. However, in few cases, there is the psychological danger of avoiding colors by simply deleting highlighted quotations. This is an inappropriate data validation practice and must be avoided. It has been noticed that some of these observations depict the market reality of the country so they have to be maintained in the data set.

When processing the survey data, Semper preserves the original copy of the "*Country Data*" file to allow exclusively human intervention in the validation process. The software does not take any step towards the correction or deletion of any "problem-

atic" price quotation. It does not directly point out any record as unsuitable and does not generate any outside table summarizing the information in the price database.

Table 2: Overview of color codes in the Semper

Absolute deviation of price quotations	Mini/Max ratio of product prices	Color coding
0-30% to average	< 0.5 – Ration is less than 1 to 2	None
30-40% to average	10% less than a 1 to 2 ratio	Yellow
40-50% to average	20% less than a 1 to 2 ratio	Pink
More than 50% to average	30% less than a 1 to 2 ratio	Red
Errors cases ("9696" or "9898")	#N/A, #DIV/0! Error types	Red

Instead, the Semper is restricted to the role of a data analysis tool, providing easy-to-read indicators to help the user in the identification of problematic items and unclean price quotations.

The underlying philosophy is that the validation and the cleaning of the price data base cannot be automated. Any automatic deletion of outliers could lead to biased results. Outliers need to be verified and only deleted or corrected if found to be wrong. Reality is sometimes surprising and different from expectations. In other words, outliers can reflect market conditions.

Therefore, the task of a software shall not be the automatic cleaning of a price report, but a highlighting of potentially problematic cases. Flagged observations have to be verified and corrected in case of errors or confirmed if reflecting reality.

7. Discussing Recurrent Issues

One specification of the Semper lies in its nature as an intra-country validation tool, consequently the possibilities of analysis in the case of just one price quotation for a certain product description are limited. In such cases no meaningful analysis on the basis of average prices and variation can be done. However, cases with just one price-quotation per product description can be analyzed in the framework of time series or through cross-country analysis of price data. Tools for analyzing these cases are included in other software and data analysis procedures such as the Quaranta tables and the Time Line analysis, an extension of the Semper validation software.

It should be highlighted that the core Semper software makes the strong assumption that each monthly price database is self-determined. It is assumed that apart from the base parameter such as the conversion factors, the outlier tolerance ratio limit, and the minimum-maximum price ratio limit, all other data analysis features are included in the database. Obviously, there is no way to validate isolated price quotations.

Other features of the software are of a conceptual nature: (1) the exclusive use of graphical instruments to highlight problematic items, instead of words and letters (2) the software runs from the Random Access Memory (RAM), and therefore requires a minimum capacity of 256 Kilobyte, (3) the ignorance of the time dimension within the country by the core Semper application is fixed by adding a Time Line component.

The first feature stems from a technical choice: for instance, it would have been possible to exclude "problematic" items from the computation of recalculated prices and outlier indicators. Instead, the Semper made the option of contaminating all quotations of a product if one of them is a fatal error. Unless all compatibility problems are resolved, the product level indicator will not be processed, and till the last product is analyzed, the overall group indicators will not be computed. One can argue about this strategy of contamination as it puts emphasis on the errors until the user solves all highlighted errors.

Secondly, the Semper is designed to run using exclusively the RAM. It does not manage internal resources of the computer. The memory is heavily used and information is stored in the RAM each time the application runs within one working session. As the Semper is not a resident installed application, it runs directly from the RAM, where the temporary information, generated when the application is running, is stored. Unfortunately, freeing the RAM will mean removing the Semper from it. There is a design rule in programming software stated as follows: "you can not cut a branch on which you are installed". This rule limits the performance of the Semper and after the application processes approximately 50,000 price quotations, it is recommended to start a new session by restarting the computer, depending on which Operating System is used. However, in standard country-validation procedures, the number of 50,000 price quotations for one survey is rarely ever reached, but can be exceeded in case of treating a huge survey several times a day.

Concerning the third feature, some room for future improvement is with the consideration of the time dimension within countries. Especially for bigger and culturally di-

verse countries, it makes a difference, if price variations occur within a certain region or between regions. The same way, in a country with high inflation, variation can occur from one month to the other. In the core Semper application, all price-quotations are treated on the national level for calculating the averages and outlier indicators. During the validation process, country experts may even now follow the regional origin of each quotation also in the Semper output; this is an important aspect when judging about plausibility of variation.

8. Practical Country Experiences

The Semper is now used in more than 40 countries on a monthly basis for one year and after some "teething problems" (mainly referring to surrounding software versions and language settings) having been fixed, it runs without any technical difficulty all over the continent. Originally, the Semper was meant to have the following main characteristics:

- Using standard software with easy or no installation procedure;
- Performing the required tasks in a tailor-made way to satisfy ICP-Africa needs;
- Working also on older models of computers;
- Being operational in different language settings;
- Being modifiable for future developments and modified uses.

The philosophy of the software creators was not to do data validation automatically, as this is considered to be a genuine human task, needing individual interpretation of specific data situations. The aim of the software is to support and facilitate this human task by highlighting dubious situations, but never to "correct" applying a standard rule.

Monthly data collection started in most African countries for the main household consumption component in June–July 2005. Three countries have started earlier in January and few launched the ICP surveys in April–May 2005. Participating countries have submitted Semper validated data from the start of data collection up to June 2006. The editing and validation of field data was organized in five sub-regions managed by sub-regional organizations namely (1) Afristat, comprising Francophone West and Central African countries; (2) COMESA, composed of Eastern African countries; (3) ECOWAS handling activities in Anglophone West African countries; (4) MAGHREB for Northern African countries directly managed from the African Development Bank, and (5) SADC for southern African countries.

Periodic sub-regional workshops have been organized in all sub-regions, with the objective of giving practical training on how to use the Semper and sharing experience on the quality of the data collected in the participating countries. The international validation was mainly based on the analysis of Quaranta tables generated by other software specifically designed for the ICP.

The Semper provided necessary tables to assess (1) the coverage in terms of the number of products covered out of the 853 items of the regional products list of the main household consumption component for urban and rural areas, (2) the number of quotations per products, (3) the average, minimum and maximum prices for reference quantities of each product specification, (4) obvious errors in terms of unit of measures, product specifications, data entry errors etc., and (5) variation in view of identifying outliers and reducing them.

9. Conclusion

The Semper Validation Software is an integrated application designed to facilitate office editing of price data collected during field surveys in the context of the International Comparison Program for Africa (ICP-Africa). The development of the application was inspired by the Input – Process – Output approach used in the industrial community.

Since its release in July 2005, the application has helped to carry out intra-country data validation in more than 40 countries participating in the ICP for the Africa region. After twelve month of experience, it is important to stop over and share about the experience reported by the countries.

As the software highlights potentially problematic price observations in various colors (red being strongest), some countries showed initially the tendency to eliminate the respective observations just because of their marking. This was subject to discussions in several workshops and has now been overcome by most countries. Countries report consistently that the Semper is a very appropriate tool to separate errors from pure price variation.

The agreement is that good quality data do reflect the reality. If reality sees pure

price variations due to transport problems, climate, culture and other influences, a high quality data set will still have some cases highlighted.

10. References

African Development Bank. (2005). Training of Field Supervisors and Price Data Collectors, A Trainer's Guide, ADB, Tunis, Tunisia, 1-77.

Adam, A. (2005). Quality Assurance Guidelines, ADB, Tunis, Tunisia, 1-18.

Astin, J. (2004). ICP 2004 Operational Manual, What National Coordinators Need to Know, Canterbury, UK, 48 - 50.

Eurostat-OECD (2005). PPP Methodological Manual - Annex IV Quaranta Editing Procedure, Luxembourg, Luxembourg, 1-6.

Kokil, B. et al. (2005). Data Validation with the Semper Excel based Software, ADB, Tunis, Tunisia, 1 - 5.

Rittenau, R. (2005). The Quality of Data in ICP-Africa, ADB, Tunis, Tunisia, 1-6.

Sergeev, S. (2003). Description of the VBA Program for the Computation of the EKS-PPP at the Basic Heading Level and the "Quaranta" Tables, Vienna, Austria, 1 - 27.

The World Bank. (2004). Price Collection ToolPack, Price Collection Module User Guide, Washington, USA, 1 - 62.

A la découverte du logiciel de validation Semper

Solutions authentiquement africaines à la validation des données dans le cadre du Programme de comparaison internationale

Mathieu B. Djayeola et Roland Rittenau[1]

Résumé

Depuis la première participation africaine en 1970, c'est la première fois qu'une institution africaine est à la tête du Programme de comparaison internationale et fournit un appui à plus de quarante pays de la région. Cet appui a porté sur la conception d'instruments d'enquêtes sur les prix destinés à la collecte de données et sur des outils d'édition en vue de la validation des données sur le plan national et régional. La mise au point du logiciel Semper se situe dans le cadre d'un processus visant à fournir des outils conviviaux destinés à aider les pays dans leurs activités de validation des données issues des enquêtes sur les prix. Outre sa technicité en tant que logiciel informatique, le Semper peut être aussi considéré comme un moyen de communication, un code et un langage entre les équipes nationales du PCI et le coordonnateur régional. Il sert surtout à tester la qualité des données en vue de leur validation internationale. L'utilisateur est assuré d'une réponse en quelques minutes dans un tableau à thèmes graphiques sur lequel il devra intervenir.

Mots clés

Comparaison internationale – enquête sur les prix – validation des données.

Summary

Since the first African participation in 1970, this is the first time an African institution has taken the lead to manage the International Comparison Program and provide support to more than forty countries in the region. This support covered the design of price survey instruments for price data collection and office editing tools for data validation at the national and regional levels. The development of the Semper Validation Software participates in the process of providing user-friendly tools to assist countries in their price survey data validation activities. Beyond the technicalities of computer software, the Semper could be appreciated as a communication vehicle, a code, and a language between the national ICP teams and the regional coordinator. The main purpose is to question the quality of the data provided for international

1: Le présent document a été élaboré par Mathieu B. Djayeola (b.djayeola@afdb.org), Statisticien-économiste et Roland Rittenau (r.rittenau@afdb.org), Statisticien principal, tous deux membres de l'équipe de Coordination régionale du PCI-Afrique à la Banque africaine de développement, Tunis (Tunisie).

validation. Answer is provided within minutes with graphical tables on which action is required from the user.

Key words

International comparison - price survey - data validation

1. Contexte

Le Programme de comparaison internationale (PCI) est une initiative statistique mondiale destinée à disposer à l'échelle internationale des prix comparables, des niveaux de dépenses et des estimations de Parité de pouvoirs d'achat (PPA) en vue de permettre des comparaisons transnationales d'agrégats économiques et de prix entrant dans l'estimation du Produit intérieur brut (PIB) et de ses sous-agrégats. Le PCI est né de la prise de conscience, très tôt, du fait que la détermination des agrégats économiques reposant sur les taux de change ne reflète pas les variations de prix d'un pays à l'autre et, en tant que tels, ces agrégats sont instables et ne sauraient orienter, seuls, la prise de décisions politiques qui, en principe, devraient se baser sur les volumes, à l'abri de toutes distorsions de prix et de taux de change.

Les prix collectés pour le PCI peuvent être utilisés à d'autres fins de comparaison comme les comparaisons spécifiques à la pauvreté[2]. Quel qu'en soit l'usage, il est important que les pays collectent des prix pour des produits comparables à la fois à l'intérieur du pays et entre les pays. Plus généralement, la qualité des comparaisons internationales, comme la qualité de l'indice des prix à la consommation (IPC), est tributaire de la qualité des données de base. Le processus de validation décrit dans ce papier est conçu pour assurer que les pays fournissent des données de prix de bonne qualité pour une sélection de produits comparables.

Ce papier se concentre sur la validation des prix collectés dans les pays pour le PCI, en d'autres termes il traite de la validation des prix qu'un pays collecte sur le terrain avant de les présenter au coordonnateur régional. Le processus de comparaison internationale entrepris ensuite par le coordonnateur régional du PCI, ainsi que les logiciels qui le supporte sont décrits dans d'autres documents[3].

2:En se fondant sur un sous-ensemble de produits entrant dans le panier de pauvreté, avec un schéma de pondération particulier en vue d'un nouveau calcul plus fiable des seuils de pauvreté internationaux dans les monnaies locales.

3:Se référer par exemple au manuel méthodologique sur les PPA édité par l'Ocde and collaboration avec Eurostat en 2005, en particulier l'annexe IV.

2. Termes de référence pour la conception du logiciel

La mise au point du Semper procède de la nécessité de vérifier, d'affiner et de valider les données sur les prix dans le cadre des Indices des prix à la consommation (IPC) et du Programme de comparaison international (PCI). Dans la plupart des pays, ce processus de validation se fait manuellement ou en faisant usage d'outils non intégrés et non spécialisés. En conséquence, il n'existe aucune garantie quant à la cohérence des méthodes utilisées dans un pays donné ou au fil du temps. Cette situation se complique davantage lorsque l'on compare les pratiques de validation d'un pays à l'autre.

La liste des produits a été l'instrument majeur d'enquête du PCI mis à la disposition des pays. Elle a été conçue grâce à un processus participatif intense auquel ont pris part les pays membres régionaux et la Banque africaine de développement. Afin de vérifier la qualité des prix relevés pour ces produits, l'équipe de coordination régionale du PCI-Afrique a confié à Mathieu B. Djayeola, statisticien-économiste, la mission de mettre au point un outil destiné à la validation des données[4]. Cette mission reposait sur les termes de référence suivants :

1. vérifier la cohérence entre le type de l'unité de mesure observée et le type de l'unité de mesure de référence[5], par exemple, le litre est incompatible avec le kilogramme et vice-versa,
2. vérifier le rapport existant entre la quantité observée et l'unité de mesure, par exemple une saisie de 800 kilogrammes au lieu de 800 grammes,
3. traiter et indiquer les prix exprimés dans diverses unités monétaires à l'exemple du Rand sud africain et de ses centimes,
4. calculer des prix moyens pour chaque produit,
5. calculer des indicateurs de dispersion des prix tels que le minimum, le maximum et l'écart type,
6. identifier les observations isolées ainsi que les erreurs potentielles,
7. analyser la variation des prix en vue d'en contrôler la tendance au fil du temps, et
8. identifier les prix mal relevés ou mal saisis : par exemple les cas de mauvais emplacement du point décimal, l'utilisation inappropriée de sous-multiples de l'unité monétaire, etc.

4: La mise au point du logiciel était entièrement placée sous la supervision de Michel Mouyelo-Katoula, Coordonnateur régional du PCI-Afrique, et sous la direction technique de Roland Rittenau, Statisticien principal.
5: Les types d'unités de mesure se réfèrent au poids, à la capacité, à l'échelle linéaire, aux unités, etc.,

Le logiciel Semper tire son nom du dicton latin *Semper aliquid novi Africa affert*[6] afin de souligner à la fois l'innovation et la spécificité africaine de l'application. Cette procédure mise au point à l'aide de *Microsoft Excel Visual Basic* permet la validation des données au plan national. Elle vise à vérifier la cohérence des unités de mesure, des quantités et des prix, et à s'assurer que les variations de prix restent dans des limites tolérables.

La quête de la qualité des données et le plaidoyer pour la comparabilité des produits ont un impact positif sur le système national de collecte d'information statistique et sur l'indice des prix à la consommation en particulier. Puisque les enquêtes du PCI se passent sur une période de douze mois en Afrique, un instrument d'analyse de séries temporelles était nécessaire. Les fonctionnalités du Semper Time Line ont été spécialement conçues à cette fin.

Qu'est-ce que c'est que le logiciel Semper? Quelles sont les tâches qu'il cible, ses caractéristiques particulières, ses avantages et ses limites? Le présent document se propose de présenter dans le détail l'expérience partagée en une année de collecte et de validation des données du PCI dans quarante-huit pays africains.

3. Qu'est-ce que c'est que le logiciel de validation Semper?

La mise au point du logiciel Semper repose sur le concept de prix observé qui, dans le cadre du Programme de comparaison internationale, comporte au minimum les informations suivantes :

1. A quel groupe de produits, ou dans la terminologie du PCI, à quelle position élémentaire le produit appartient-t-il ? Selon les classifications internationales, une bouteille de boisson à base d'orange achetée en ville dans un magasin n'a pas la même signification que la même bouteille achetée dans un restaurant dont le contenu est destiné à la consommation immédiate. Du point de vue de la nomenclature de la comptabilité nationale, ces deux bouteilles n'appartiennent pas à la même catégorie dans la mesure où leurs finalités diffèrent vis-à-vis de la consommation. Par exemple, dans le cadre du PCI, la bouteille trouvée dans la boutique appartient à la position élémentaire « Boisson non alcoolisée et concentrée », alors que l'autre rentre dans la position élémentaire « Approvisionnement dans les hôtels et restaurants ».

6:" En Afrique, il y a toujours quelque chose de neuf à découvrir ": Auteur inconnu, transcription de Pliny the Elder.

2. L'identification du produit : un code est nécessaire pour le traitement automatique des informations liées aux produits, tout comme le nom d'un produit est nécessaire pour cerner la signification du code spécifique qui lui est affecté. C'est là aussi une question de classification qui pourrait s'avérer pertinente dans le cadre de l'analyse économique des données collectées sur le terrain. Toutes les caractéristiques et modalités des produits constituent des facteurs importants dans leur identification. Par exemple, la "Conserve d'ananas" est entièrement définie à l'aide de la spécification suivante : [Position élémentaire : Conserves de fruits; Code du produit : 025.07 ; Quantité de référence et unité de mesure : 850 Millilitre(s) ; Mode de présentation du produit : Boîte ; Quantité : 800-900 Millilitres ; Type : Ananas ; Forme : Morceaux ; Jus : Sirop (eau + sucre)].
3. L'unité de mesure observée : il est nécessaire de s'assurer que l'unité de mesure observée est identique à l'unité de mesure de référence ou, tout au moins, qu'il est y convertible. Par exemple, si l'unité de mesure de référence de la "Conserve d'ananas" est le millilitre, l'unité observée effective peut être un multiple du millilitre, tandis que des unités de poids (telles que le gramme ou la livre, etc.) ne seront pas admises.
4. La quantité observée : pour permettre l'estimation du prix unitaire, la quantité observée doit être signalée (en relation avec la quantité de référence). Le prix unitaire est aussi appelé prix recalculé dans le logiciel.
5. L'indication du prix observé.
6. Le lieu d'observation : le point de vente, la localité, le pays où a été relevé un prix sont des facteurs importants dans la détermination des prix. Le prix de l'eau minérale dans une zone touristique ne sera pas le même que dans un magasin de quartier dans une zone résidentielle.
7. Le moment de la collecte des données pourrait être aussi un facteur crucial dans le cas de certains produits. Les prix des produits saisonniers varient d'un mois à l'autre, et les prix des produits périssables peuvent varier selon le moment de la vente dans la journée. Par exemple, les prix des fruits de mer frais peuvent baisser le soir.
8. Les commentaires et remarques peuvent fournir des informations supplémentaires sur un prix, notamment sur les conditions de l'enquête ou des écarts enregistrés par rapport aux spécifications de produit.

Le logiciel de validation Semper est une application intégrée destiné à l'édition, au bureau, de données recueillies sur le terrain dans le cadre du Programme de comparaison internationale pour l'Afrique. Sa procédure en trois phases se déroule comme suit :

1. Vérification des constantes de l'enquête : conformité du code du produit, nom, quantité, unité de mesure et autres caractéristiques, en rapport avec la spécification de produit. Tout produit ne figurant pas sur la liste des produits du PCI-Afrique est immédiatement rejeté par le Semper.
2. Détermination du double rapport mathématique existant entre la quantité observée, la quantité de référence et l'unité de mesure, et exécution des conversions requises en vue de réajuster les prix observés.
3. Analyse des prix ajustés des produits en vue de l'identification des observations isolées et les erreurs potentielles.

4. Comment fonctionne-t-il?

Le logiciel de validation Semper analyse la base de données des prix des pays appelée "*Country Data*" dans le cadre du Programme de comparaison internationale pour l'Afrique. Pour son fonctionnement normal, le logiciel a besoin de fichiers propres et normalisés, et d'une base de données structurée et vérifiée. L'application procède tout d'abord à la vérification de l'existence de ces conditions requises avant tout traitement. A l'instar d'une usine, le Semper fonctionne selon un système Entrée-Traitement-Sortie. (Figure 1).

A l'entrée – Ce dont le Semper a besoin.

L'outil de validation Semper génère ses inputs à partir des masques de saisie du PCI-Afrique. Ces masques représentent des moyens particuliers de saisie des données avec un degré élevé de sécurité et de protection des cellules; ils sont dotées de fonctions de défilement permettant de guider l'utilisateur au cours du processus de saisie des données, de garantir l'homogénéité de la structure des ensembles de données, et d'éviter autant que possible des erreurs typographiques. Ils sont généralement structurés dans un pays donné par centre de collecte et sont de taille relativement importante du fait du nombre de fonctions de contrôle qu'ils contiennent. Fusionnés à l'aide d'une macro automatique Excel[7], un fichier unique de base de données dénommé "*Country Data*" est obtenu. Il servira d'input au Semper.

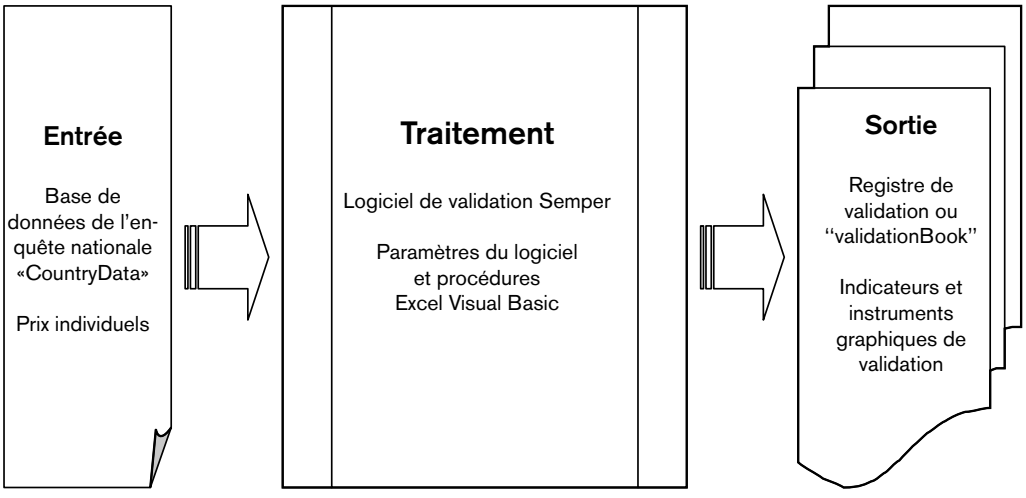
Le processus – Que fait le Semper?

Le fichier du programme Semper contient cinq feuilles de calcul et un module *Visual Basic* intégré. La première feuille (SP lire « Spécification de produit ») contient des

7: Cette macro dénommée "*MergeCountryDataMacro*" est requis pour la mise en place rapide et automatique d'une base de données nationale à valider.

informations générales sur la spécification des produits et les paramètres de validation de données individuelles tels que les quantités de référence, l'unité de mesure de référence, et le pourcentage tolérable d'écart d'un prix individuel par rapport à la moyenne. La deuxième feuille contient des informations sur les unités de mesure compatibles et les facteurs de conversions subséquentes, la quatrième (Informations destinées aux utilisateurs) des informations de base sur la façon dont est structurée la base de données des prix et les modalités de calcul des indicateurs d'écart, et la dernière feuille est réservée aux commentaires et remarques de l'utilisateur. Quant à la troisième feuille (Informations sur le programme), elle constitue l'élément central du Semper, et contient un Bouton de démarrage servant à lancer l'application ainsi que la séquence de traitement.

Figure 1 : Structure du logiciel Semper



Sortie : Que produit le logiciel de validation Semper?

L'output, dénommé registre de validation ou "*ValidationBook*", est un classeur Excel particulier à trois feuilles : (1) la feuille zone de calcul ou "*Computing Area*", (2) la feuille Récapitulatif ou "*Summary*" pour le stockage temporaire des informations traitées, et (3) la feuille Divers ou "Miscellaneous", servant de carnet de notes pour l'utilisateur (Tableau 1 – *zone de calcul*).

En utilisant les prix convertis d'un produit donné (en relation avec la quantité de référence), le Semper calcule des indicateurs sélectionnés destinés à l'analyse des données d'enquête. Ce sont :

(1) les prix moyens,
(2) le nombre de prix observés ou nombre de relevés pour un produit,
(3) le prix minimum,
(4) le prix maximum,
(5) l'écart type des prix,
(6) le nombre de relevés marqués représentatifs,
(7) le nombre de relevés indiquant un élément du panier de pauvreté,
(8) le nombre de relevés dans des zones dites de pauvreté.

Après avoir traité tous ces indicateurs, le Semper entame la validation de chacune des données en générant simultanément un indicateur d'écart pour des prix individuels observés et un indicateur minimum - maximum pour un groupe de prix observés pour chaque produit. Ensuite, il met en évidence les prix isolés à l'aide d'une gamme de couleurs : absence de couleur lorsque le prix est proche de la moyenne, couleur jaune lorsque le prix marque un écart acceptable par rapport au prix moyen, couleur rose signalant l'éventualité d'un écart considérable, et couleur rouge pour mettre en relief de grands écarts et des erreurs évidentes (Tableau 2).

Les utilisateurs peuvent sauvegarder les résultats obtenus par le logiciel sous diverses dénominations ou à des endroits précis en vue de futures consultations. Toute tentative visant à corrompre la structure du fichier ou toute configuration inappropriée de la base de données peut entraîner des erreurs et l'arrêt du programme.

Dans l'exemple ci-dessous, (Tableau 1) l'information est présentée d'abord au niveau global de la base de données des prix (Agrégation globale): le pays concerné a effectué au total 4732 relevés de prix. Même si le prix moyen n'a pas de signification matérielle dans cette présentation, le fait qu'il soit calculé indique que le logiciel n'a rencontré aucun problème technique dans la mesure où toutes les observations individuelles ont pu être traitées.

Un niveau plus bas, le niveau du produit est indiqué par 00104_Agrégation, en relation avec le produit 00104. Dans l'exemple, trois produits ont été observés, notamment 00104, 00105 et 00107, tous appartenant à la position élémentaire 001 – Riz. A ce niveau, la table montre le nombre de relevés et le prix moyen recalculé. Par exemple, pour le produit 00104, ce pays a effectué 6 relevés avec un prix moyen recalculé de 553 unités monétaires. Tous ces indicateurs sont dits *recalculés parce* qu'ils sont rapportés à la quantité de référence. L'indicateur d'écart pour le produit 00104 est rouge parce qu'il est inférieur à -30, valeur correspondant à un prix minimum inférieur à la moitié du prix maximum pour ce produit.

Au niveau des relevés individuels, des indices identifient chaque relevé dans la base de données des prix. Ceci est particulièrement important durant la phase d'édition ou de correction des données. L'indicateur d'écart pour les relevés individuels change de couleur lorsque ces derniers s'éloignent du prix moyen pour ce produit.

Tableau 1 : Visualisation du Registre de validation ou ValidationBook (Zone de calcul ou Computing Area)

Indice	Nom du produit [02]	Indicateur d'écart	Type de point de vente requis [03]	Code de produit [01]	Nombre de prix recalculés	Moyenne des prix recalculés	...
				Agrégation globale	4732	1292	
		-32		00104_**Agrégation**	6	553	
01225	Riz long grain	8	Supermarché	00104	1	632	
01277	Riz long grain	-4	Supermarché	00104	1	340	
03873	Riz long grain	-16	Supermarché	00104	1	898	
04097	Riz long grain	6	Supermarché	00104	1	452	
04140	Riz long grain	12.5	Supermarché	00104	1	582	
04370	Riz long grain	3	Supermarché	00104	1	416	
		30		00105_**Agrégation**	5	1740	
00687	Riz long grain	9	Magasin de quartier, Marché	00105	1	1960	
00688	Riz long grain	13	Magasin de quartier, Marché	00105	1	1666	
01721	Riz long grain	9.5	Magasin de quartier, Marché	00105	1	1540	
02119	Riz long grain	7.5	Magasin de quartier, Marché	00105	1	1470	
03661	Riz long grain	6	Magasin de quartier, Marché	00105	1	2065	
		-30		00107_**Agrégation**	32	70	
00124	Riz mi-long	8	Magasin de quartier	00107	1	60	

00280	Riz mi-long	8.5	Magasin de quartier	00107	1	80	
00480	Riz mi-long	1	Magasin de quartier	00107	1	90	
00642	Riz mi-long	11.5	Magasin de quartier	00107	1	65	
00982	Riz mi-long	13.5	Magasin de quartier	00107	1	68	
…	….	…	…	…	…	…	…

Pour de plus amples détails, un manuel de l'utilisateur est disponible.

5. Caractéristiques particulières

Contrairement à la plupart des applications *Microsoft Visual Basic*, le Semper repose à la fois sur des procédures internes et externes du logiciel de base - *Microsoft Excel*. Les fonctions habituelles d'Excel telles que les formules, copier-coller ainsi que les procédures de création de sous-groupes constituent des procédures internes au Semper. Au-delà du Visual Basic ordinaire, le Semper a des interactions avec le système d'exploitation en vue de recueillir des informations sur des paramètres dont l'emplacement est fonction de la langue de conception du système : par exemple, où Excel stocke-t-il les informations sur les sous-groupes lorsqu'il procède au calcul des sous-totaux ? Dans la version anglaise d'Excel, l'identification des sous-groupes se trouve à gauche de la chaîne, alors qu'elle est à droite dans certaines des versions françaises. En interrogeant les fichiers système de l'ordinateur, le Semper arrive à localiser ces éléments d'information et à en tenir compte dans ses opérations.

Etant donné que l'application recueille les paramètres nécessaires sur une base autonome, le Semper exécute des procédures particulières en vue de s'adapter à la taille de la base de données des relevés de prix. Cette procédure d'adaptation est destinée à permettre à l'application d'extraire, toute seule, des informations relatives au nombre de prix observés, au nombre de produits sur lesquels a porté l'enquête, et au nombre de relevés par produit.

L'objectif fondamental de la conception du Semper est le respect rigoureux des normes de validation actuelles du PCI. L'input du logiciel ou la base de données sur les prix ("*Country Data*") provient de la liste des produits et du formulaire de saisie des données, tous compilés au préalable en stricte conformité avec les critères de la Banque mondiale en matière de transmission de données. Un des prolongements du Semper, appelé *Time Line* (analyse de séries temporelles), analyse les séries chronologiques, c'est-à-dire qu'il permet de valider des données en ajoutant une compo-

sante temporelle. Dans une perspective de l'indice des prix à la consommation, cette composante paraît novatrice.

6. Avantages

Le Semper génère une matrice centrale dans le *Registre de validation ou "ValidationBook"* qui peut être utilisée pour d'autres analyses graphiques des données : navigation, filtrage et regroupement. Il existe trois niveaux d'analyse dans le Registre de validation. Plus le niveau baisse plus l'analyse est raffinée :

(1) le niveau des informations générales sur les prix relevés,
(2) le niveau des produits, et
(3) le niveau des prix observés.

Il est facile de naviguer dans le fichier d'analyse Registre de validation ou "ValidationBook" : l'utilisateur peut se focaliser sur un produit, sélectionner une gamme de produits, sélectionner des prix observés pour un produit, ou analyser l'ensemble des relevés. Le Semper permet aussi la sélection de prix observés sur la base de la valeur de leur indicateur d'écart, le niveau de représentativité et/ou la pertinence du produit à l'analyse de la pauvreté.

Les erreurs de calcul sont clairement signalées à l'aide de deux codes spécifiques qui constituent l'indicateur d'écart : le logiciel crée un code d'écart spécifique "9696" si l'un des paramètres permettant l'estimation du prix recalculé (prix, quantité ou unité de mesure) venait à manquer. Si, en dépit de la présence de toutes ces variables, l'estimation du prix recalculé est impossible le code "9898" est renvoyé. Cette situation se présente en général lorsque l'unité de mesure observée est incompatible avec l'unité de référence (Figure 3).

L'indicateur d'écart est exprimé en valeurs positives pour les relevés « non problématiques » et en valeurs négatives pour les relevés « problématiques ». Il est calculé sur la base d'une formule définie en fixant un degré tolérable d'écart dans le prix d'un produit donné par rapport à la moyenne. De tels écarts tolérables sont variables et peuvent être modifiés par l'utilisateur. Il est par ailleurs possible de fixer diverses limites de tolérance pour divers types de produits, afin de refléter l'hétérogénéité des marchés au niveau des procédures de validation. Un nombre plus élevé en terme absolu est attribué à un relevé problématique si ce relevé a un écart plus important à la moyenne. Ce faisant, les filtres d'Excel mettront en relief les éléments

les plus suspects et attireront l'attention de l'utilisateur sur eux avant de passer à la présentation des relevés moins problématiques et ceux pour lesquels le Semper n'a décelé aucune anomalie.

Le codage des relevés isolés en couleur s'est avéré très pratique et est bien apprécié par la plupart des utilisateurs. Cependant, dans un certain nombre de cas, un risque psychologique subsiste, qui consiste à vouloir éviter les couleurs en effaçant purement et simplement les prix observés mis en évidence par le Semper. Cette pratique est inopportune et devrait être évitée dans la mesure où certaines de ces données reflètent la réalité du marché du pays concerné. En conséquence, ces relevés devraient être maintenus pour l'analyse de la situation du marché.

Dans le traitement des données d'enquête, le Semper conserve l'original du fichier "*Country Data*" pour ne permettre que l'intervention humaine dans le processus de validation. Le logiciel ne prévoit aucune mesure allant dans le sens de la correction ou de l'élimination d'un prix 'problématique'. Il ne signale pas directement le caractère inapproprié d'une information et ne génère aucun tableau séparé récapitulant les informations contenues dans la base de données de prix.

Tableau 2 : Aperçu général du codage en couleur utilisé par le Semper

Ecart absolu des prix observés	Rapport minimum/maximum des prix des produits	Codage en couleur
0-30% par rapport à la moyenne	< 0,5 – Rapport inférieur à 1 sur 2	Aucune couleur
30-40% par rapport à la moyenne	10% inférieur au rapport 1 sur 2	Jaune
40-50% par rapport à la moyenne	20% inférieur au rapport 1 sur 2	Rose
Plus de 50% de la moyenne	30% inférieur au rapport 1 sur 2	Rouge
Cas d'erreurs ("9696" ou "9898")	#N/A, #DIV/0! Types d'erreurs	Rouge

Par contre, le rôle du Semper se limite à celui d'un outil d'analyse de données en mettant à la disposition de l'utilisateur des indicateurs faciles à lire en vue de l'aider à identifier les éléments problématiques ou les prix isolés.

La philosophie qui sous-tend la démarche est que la validation et le nettoyage de la base de données ne saurait être automatisés. Toute suppression automatique des observations isolées pourrait mener à des résultats biaisés. Les observations isolées doivent être vérifiées puis, seulement si elles s'avèrent erronées, elles pourront être corrigées ou supprimées. La réalité est quelquefois surprenante, c'est-à-dire diffé-

rente de ce que l'on pourrait espérer. En d'autres termes, des points isolés peuvent être une réflexion des conditions du marché.

En conséquence, le rôle du logiciel n'est pas le nettoyage automatique d'un relevé de prix, mais la mise en évidence de cas problématiques potentiels. Les observations signalées devront faire l'objet de vérification et de correction en cas d'erreurs, ou de confirmation lorsqu'elles reflètent la réalité.

7. Quelques questions conceptuelles

L'une des caractéristiques particulières du Semper est liée à sa nature, en qualité d'outil de validation dans le pays. Lorsqu'un seul prix est observé, les possibilités d'analyse des relevés pour le produit concerné sont limitées. Dans de tels cas, aucune analyse significative axée sur les prix et écarts moyens n'est possible. Cependant, lorsqu'il n'existe qu'un seul prix pour une description de produit, l'analyse des séries chronologiques ou l'analyse transnationale de données apportent une solution. Ces possibilités sont prises en compte par d'autres logiciels et procédures d'analyse de données tels que les tableaux de Quaranta et l'analyse des séries temporelles par le Time Line, une procédure complémentaire au logiciel Semper.

Il faut signaler que le logiciel Semper de base s'appuie sur l'hypothèse que chaque base de données mensuelle de relevés de prix est autodéterminée. L'on suppose qu'à part les paramètres de base tels que les facteurs de conversion, le ratio de tolérance des écarts, et le ratio minimum/maximum, toutes les autres fonctions d'analyse de données sont intégrées dans la base de données, ce qui, naturellement, ne permet aucunement la validation de relevés pris isolément.

D'autres caractéristiques du logiciel sont de nature conceptuelle et concernent : (1) l'utilisation exclusive d'instruments graphiques en vue de mettre en évidence des éléments problématiques, en lieu et place chiffres ou de mots, (2) le fonctionnement du logiciel à partir de la Mémoire vive, ce qui exige au minimum 256 Kilo octet de capacité, (3) l'abstraction faite par l'application Semper de base de la dimension temporelle dans le pays est compensée par la mise à disposition d'une composante temporelle appelée Time Line.

La première de ces caractéristiques découle d'un choix technique : par exemple, il aurait été possible d'exclure les éléments 'problématiques' de l'estimation des prix recalculés et des indicateurs d'écart. Mais le Semper a fait le choix de contaminer

tous les prix observés d'un produit si l'un d'eux venait à générer une erreur de calcul. Le produit ne sera pas traité tant que tous les problèmes de compatibilité ne seront pas résolus pour tous les relevés y afférant. De même, tant que le dernier produit ne sera pas analysé, les indicateurs généraux de groupes ne seront pas calculés. Cette stratégie de contamination est discutable dans la mesure où les erreurs sont remises en évidence jusqu'à ce qu'elles soient entièrement corrigées par l'utilisateur.

En second lieu, le Semper a été conçu pour fonctionner exclusivement sur la Mémoire vive. Il ne gère pas les ressources internes de l'ordinateur. La mémoire est sollicitée de manière intensive, et des informations sont stockées dans cette mémoire chaque fois que l'application fonctionne au cours d'une session de travail. Etant donné que le Semper n'est pas une application résidente, il fonctionne directement à partir de la mémoire vive où sont stockées temporairement les informations générées au cours des opérations de l'application. Malheureusement, libérer cette mémoire reviendrait à la défaire du Semper. Une des règles conceptuelles de programmation du logiciel stipule que l'on 'ne doit pas scier la branche sur laquelle on est installée'. Cette prescription limite la performance du logiciel, et après le traitement d'environ 50 000 relevés, il est recommandé, en fonction du Système d'exploitation utilisé, d'ouvrir une nouvelle session ou de redémarrer l'ordinateur. Il est à noter qu'il est rare d'atteindre ce seuil dans le cadre d'une seule enquête dans un pays. Ceci n'exclut pas qu'il puisse être dépassé dans le traitement d'une enquête d'envergure plusieurs fois par jour.

Quant à la troisième caractéristique, il conviendrait de noter que la prise en compte de la dimension temporelle au niveau des pays laisse entrevoir quelques possibilités d'amélioration. Particulièrement dans les vastes pays pluriculturels, les variations de prix au sein d'une région donnée ou entre deux régions importent. De même, dans un pays où sévit une forte inflation, des variations peuvent être enregistrées d'un mois à l'autre. Avec l'application du logiciel Semper de base, tous les prix observés sont traités à l'échelle nationale en vue du calcul des moyennes et des indicateurs d'écart. Au cours du processus de validation, les experts nationaux sont capables de retracer l'origine des données dans les fichiers d'analyse du Semper, ce qui est important lorsque l'on veut se prononcer sur la plausibilité d'une variation.

8. Expériences pratiques dans les pays

Pour le moment, le Semper a été utilisé chaque mois pendant un an dans plus de 40 pays. Après que des solutions aient été trouvées aux "défauts de jeunesse" (concer-

nant les versions du logiciel dans les différents environnements linguistiques), l'application fonctionne partout sur le continent, sans aucune difficulté technique.

Il n'était pas de l'intention des concepteurs du logiciel de créer un outil de validation automatique de données, tout en sachant que ce travail, qui requiert l'interprétation des situations particulières de collecte des données, revient en priorité au personnel des services statistiques responsables de la collecte et de la validation des données. Le logiciel est destiné à appuyer et faciliter cette tâche humaine en mettant en évidence les situations douteuses plutôt que de les 'corriger' en appliquant une règle automatique.

A l'origine, le Semper a été conçu pour avoir les caractéristiques majeures suivantes :

- Etre d'installation facile ou sans procédure d'installation,
- Etre capable d'accomplir les tâches requises en recourant à des solutions sur mesure en vue de la satisfaction des besoins du PCI-Afrique,
- Etre utilisable sur les anciens modèles d'ordinateurs,
- Etre utilisable dans divers environnements linguistiques,
- Etre modifiable pour permettre des améliorations et servir différents objectifs.

La collecte mensuelle des données de prix a commencé dans la plupart des pays africains pour la composante principale des biens de consommation des ménages en juin – juillet 2005. Trois pays avaient commencé plus tôt en janvier et un petit nombre les avaient rejoints en avril – mai 2005. Les pays participants ont soumis des données validées avec le Semper depuis le démarrage des activités d'enquête jusqu'à juin 2006. L'édition et la validation des données ont été organisées suivant cinq sous régions d'opération gérées par des organisations sous-régionales à savoir : (1) Afristat comprenant les pays francophones d'Afrique de l'ouest et du centre, (2) Comesa, composée des pays d'Afrique de l'est, (3) Cedeao avec la gestion des pays anglophones d'Afrique de l'ouest, (4) Maghreb pour les pays d'Afrique du nord gérés directement depuis la Banque africaine de développement et (5) Sadc pour les pays d'Afrique australe.

Des séminaires sous-régionaux ont été organisés dans toutes les sous régions avec pour objectif de former les experts nationaux à l'utilisation du Semper et partager leurs expériences en matière de validation des données collectées dans les pays. La validation internationale s'est basée sur l'analyse des tables de Quaranta générées avec des logiciels autres que le Semper et conçues spécialement pour le PCI.

Le Semper a fourni les tableaux nécessaires à l'évaluation (1) de la couverture en termes de nombre de produits observés dans un pays sur les 853 produits de la liste régionale des biens de consommation des ménages pour les zones urbaines et rurales, (2) du nombre de relevés par produit, (3) des prix moyen, minimum et maximum pour les quantités de référence et pour chaque spécification de produit, (4) des erreurs évidentes d'unité de mesure, spécification de produit et des erreurs de saisie et (5) de la variation en vue de l'identification des relevés isolés et sa réduction.

9. Conclusion

Le logiciel de validation Semper est une application intégrée destinée à faciliter l'édition au bureau des données recueillies sur les prix au cours d'enquêtes sur le terrain dans le cadre du Programme de comparaison internationale pour l'Afrique – PCI-Afrique. Sa conception s'est inspirée de l'approche Entrée – Traitement – Sortie du processus industriel.

Depuis sa mise en service en juillet 2005, l'application a permis la validation des données nationales de plus de 40 pays participant au PCI-Afrique. Après douze mois d'expérimentation, il est important de marquer un arrêt et de partager ce qu'en disent les collègues des pays qui l'ont utilisé.

Etant donné que le logiciel met en évidence les prix observés potentiellement problématiques en leur affectant diverses couleurs (le rouge étant réservé aux cas extrêmes), au départ certains pays ont eu tendance à éliminer les observations respectives uniquement à cause de leur couleur, ce qui a fait l'objet de discussions au cours de plusieurs ateliers. Aujourd'hui, ces pays ont dépassé ce stade de compréhension de l'utilisation de l'outil. Ils rapportent fréquemment que le Semper est un outil très efficace servant à faire la part entre les erreurs et les variations absolues de prix.

L'idée essentielle à retenir est que des données de bonne qualité doivent effectivement refléter la réalité du marché. Si cette réalité est empreinte de variations absolues de prix liées aux difficultés de transport, au climat, à la culture et à d'autres influences, il est tout à fait naturel d'obtenir, même dans une base de données de qualité, des éléments colorés.

10. Références

Adam, A. (2005). Quality Assurance Guidelines, BAD, Tunis, Tunisie, 1-18.

Astin, J. (2004). ICP 2004 Operational Manual, What National Coordinators Need To Know, Canterbury, Angleterre, 48 - 50.

Banque africaine de développement. (2005). Training of Field Supervisors and Price Data Collectors, A Trainer's Guide, BAD, Tunis, Tunisie, 1-77.

Banque mondiale. (2004). Price Collection ToolPack, Price Collection Module User Guide, Washington, Etats-Unis, 1 - 62.

Eurostat-OCDE (2005). Methodological Manual - Annex IV Quaranta Editing Procedure, Luxembourg, Luxembourg, 1-6.

Kokil, B. et al. (2005). Data Validation with the Semper Excel based software, BAD, Tunis, Tunisie, 1 - 5.

Rittenau, R. (2005). The Quality of Data in ICP-Africa, BAD, Tunis, Tunisie, 1-6.

Sergeev, S. (2003). Description of the VBA Program for the Computation of the EKS-PPP at the Basic Heading Level and the "Quaranta" tables, Vienne, Autriche, 1 - 27.

Correcting Survey Non-Response With Census Data

Johannes G. Hoogeveen[1] and Youdi Schipper[2]

Summary

Household size related non-response occurs because the enumerator fails to find someone at home, or because information on household members is not captured. Using census and survey data from Uganda we show how such non-response leads to substantial bias in the survey distribution of household size and how, with the use of commonly available census information, the non-response bias can be corrected.

Keywords

Poverty and inequality measurement.

Résumé

La non-réponse liée à la taille du ménage se produit quand l'agent recenseur ne trouve personne à la maison ou quand l'information sur les membres du ménage n'a pas été saisie. En utilisant des données de recensement et d'enquête de l'Ouganda, nous montrons comment ces non-réponses occasionnent des biais importants dans la distribution de la taille des ménages issue des enquêtes et comment corriger ces biais en utilisant des informations du recensement communément disponibles.

Mots clés

Pauvreté et mesure d'inégalités

1. Introduction

This paper considers how to correct for non response in a sample survey when non-response is related to household size. Household size related non-response may occur when an enumerator fails to find any respondent at home. The probability of

1: Johannes Hoogeveen is with the World Bank, C. O. World Bank, P.O.Box 2054, Dar es Salaam, Tanzania, jhoogeveen@worldbank.org.
2: Youdi Schipper works with the Vrije Universiteit Amsterdam, FEWEB-4A29 De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, yschipper@feweb.vu.nl. We are grateful to the Uganda Bureau of Statistics for their help with the provision of survey and census data. We also very much appreciate useful help and suggestions from Johan Mistiaen. The findings, interpretations and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent.
3: It would appear that the probability of not finding the household head at home is much more equally distributed over household size classes. However, if the head of household is not at home, typically another household member will respond.

finding no one at home in a one person household is larger than in a multiple person household.[3] Small households are therefore likely to be underrepresented in the survey. A second type of household size related non-response results from the non-recording of individual household members. In surveys where information has to be collected for each individual in the household (as is typical for living standards type of surveys) not all individuals may be captured. This is more likely to happen in very large households. Both types of household size related non-response lead to an under-representation of both small and very large households and an over-representation of medium sized households.

If non-response is correlated with household size and if non-responding households are replaced in the sample without taking this correlation into account, substitute households are likely to differ systematically from the non-respondents they replace (Kish, 1965, Lessler and Kalsbeek, 1992). As a result, survey estimates may be biased. This paper presents an ex-post approach to correcting for household size related non-response by using an external source of data, the population census.

This paper follows a tradition in which adjustments for non-response are made by identifying (re)weighting factors for every household in the survey. Various methods for determining these factors have been suggested. One proposal has been to infer the weights using partial information that may exist on non-respondents including, for instance, the number of attempts required to obtain a response (Politz and Simmons, 1949). An alternative method infers these weights from the distribution of non-respondents across certain identifiable subgroups of the sample, called “adjustment cells” (Thomsen, 1973). External sources of data, such as a population census, have also been proposed to determine the ‘true’ number of units in the various subgroups of the population (Hansen et al., 1953). Our contribution in this paper is to clarify how non-response may be related to household size and to show how external sources of data can be used to correct for such non-response. We also show how these external data can be used to assess whether adjustment reduces survey bias in dimensions beyond the one (household size) for which the correction is carried out.

We address these issues using the 1992 Uganda Integrated Household Survey (IHS) and the 1991 population census. We show how survey non-response varies with household size using the “correct” census distribution of household size (Section 2) and adjust survey weights accordingly (Section 3). We then show that using the new weights, the comparability between census and survey improves considerably. Section 4 investigates the bias that household size related non-response may create for commonly used welfare indicators, explores the importance of income re-

lated non-response, and shows that the problem is not confined to Uganda. Concluding remarks, finally, present a case to experiment with different visiting strategies to altogether avoid household size related non-response.

2. Survey Non-response and Household Size

Survey non-response is likely to be highly correlated with household size as the probability of finding someone at home is likely to be higher for large than for small households. With a fixed number of attempts to visit a household small households are disproportionately likely to not be included in a survey. Formally, let the probability of a respondent i being at home at the time an enumerator calls be represented by p_i and assume that the presence of household members at home is independent. The probability of non-response p_s^n for a household with s number of respondents can then be expressed as:

$$p_s^n = (1 - p_i)^5$$

Expression (1) is non-linear and the probability of non-response decreases exponentially with the number of respondents. Consider a probability of a respondent being at home of, say, 0.7. This probability is probably an over-estimation, considering that it reflects the probability that an adult household member who is able to respond is at home at a time when enumerators carry out their assignment. At this relatively high probability of being at home, the probability of non-response drops to below 3 percent for households of size 3 and larger. But for small households (of course depending on the probability of finding a respondent at home) the probability of non-response is still non-negligible. In our illustration but with only one attempt to visit a household the probability of non-response for a one-respondent household is 30 percent and for a household with two respondents 9 percent. If the number of visits is increased these fractions will diminish, but a sample survey that does not have a visiting strategy which increases the likelihood of finding a respondent at home to close to unity and whose replacement procedure is independent of household size is likely to under-sample small households and over-sample larger households. Without knowing the probability p_i, the final dataset cannot be corrected for this bias.

Another reason leading to the incorrect representation of household size in surveys is that household members may not be recorded; this may be due to memory lapse or enumerator error. Let the probability of non-recording be r_i for each household member. Given that an enumerator is interviewing one household member, the prob-

ability that at least one household member is not recorded in a household of size *s* can be expressed as one minus the probability that all remaining household members are recorded:

$$p_s^r = (1 - r_i)^{s-1}$$

This probability increases with household size at a decelerating pace and approaches 1 for very large households. Even a low probability of non-recording of say 0.01 will lead to an almost 5 percent probability that household size is underestimated by at least one member for a household whose (true) size is 6 and of over 10 percent for a household whose (true) size is 12.

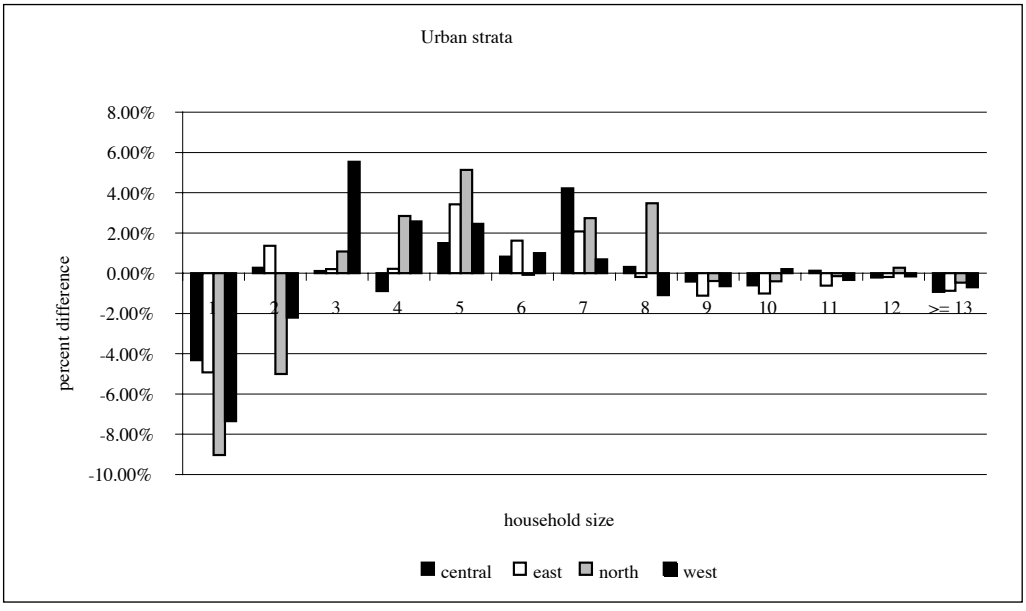
The probability of non-response and the probability of under-recording in a sample survey lead to the tails of the distribution of household size being thinner than they should be. Since both types of measurement problems have potentially offsetting consequences for the mean or the median, the thinning of the tails in the sample survey could easily go undetected without considering the distribution of households over size classes. By its very nature a census focuses on the correct administration of household size. Hence much attention is devoted to training of enumerators to ensure that this aspect of the census is 'right'. Also, census enumeration takes little time so that inclusion of an extra individual has small marginal cost for the enumerator and respondent. In contrast, household survey questionnaires typically take much more time and are less focused on obtaining correct household size. Provided, then, that a census is not subject to non-response problems the 'thinning' of the tails of the household size distribution in survey data can be identified by comparing both distributions.

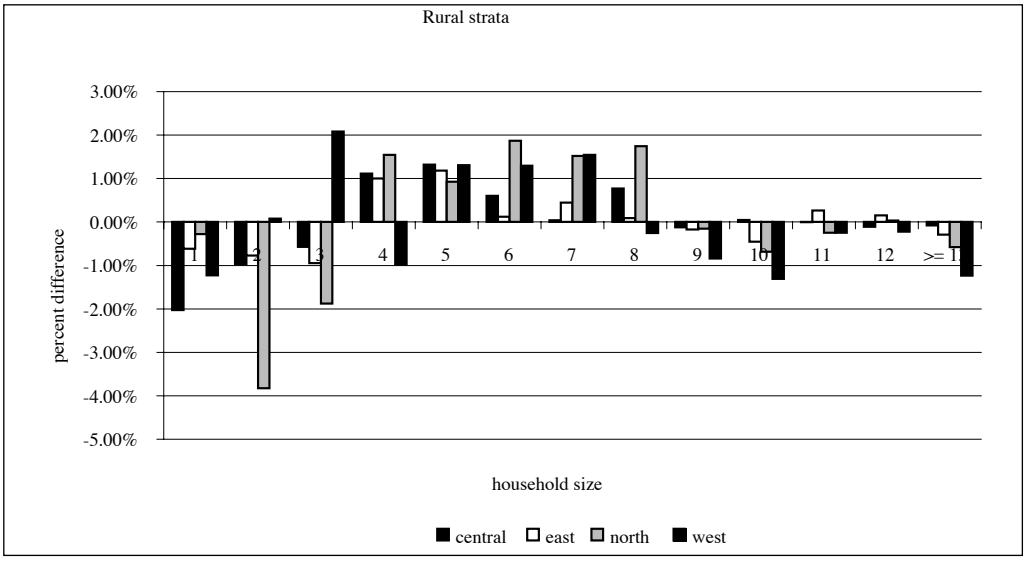
How important are these effects in practice? In the second week of January 1991 a population census was held in Uganda, whereas from January to December 1992 the Integrated Household Survey was implemented. Both the survey and the census recorded household size by collecting information on each member of the household. The definition for a household is identical between the survey and the census.[4] The survey was representative in 8 strata (rural and urban areas in respectively central, east, north and west Uganda), so that measures of household size should be (statistically) identical between the census and the survey.

1: A household is defined as a group of persons who live, cook and eat together or a single person who lives alone and eats independently.

Figure 1 explores the existence of a bias in greater detail. It plots for each stratum and for different household size-categories the difference in the proportion of households reported by the census and the survey. In total 13 categories are distinguished: households of size 1 up to 12 (reflecting 98 percent of the total population) and a rest category comprising all households with 13 or more members. The figure shows a clear pattern. Households comprising 1 or 2 members (in rural areas 1, 2 or 3 members) are under-represented in the survey; households of size 4-8 are over-represented. Large households comprising 9 or more members are under-represented again. Not only is the pattern striking, the differences can be sizeable especially in urban areas where, depending on the stratum, the fraction of one-person households in the survey is four to more than eight percentage points less than that in the census. In rural areas the differences are less pronounced, yet the pattern is comparable to that in urban areas.

Figure 1: Difference in proportion of households of size x reported in census and sample survey (a negative number indicates under-representation in the survey).





3. Adjusting for Survey Non-response

Having established the existence of the problem, the question is "can we control for it?" In the presence of a census, this can be achieved through a re-weighting procedure which ensures that the frequency distribution among mutually exclusive and exhaustive categories in the survey correspond precisely to the frequency distribution among those same categories in the census. Formally, let the total population in the census be N and let there be N_s people in the census living in households of size s. Corresponding variables for the survey are indicated by a lower case. If the inflation factor for a household i of size s is w_{is} then, according to a survey of h households, the total population living in households of size s is given by $n = \sum^{h} s_i w_i$ and the total population is: $n_s = s\sum_{i\text{-}1}^{h_s} w_{is}$. Denote the fraction of the population living in a size s household as f_s, which corresponds to the fraction F_s that can be obtained from the census as N_s / N. If, as is the case in Uganda, the total population according to the survey (n) corresponds to the total population according to the census (N), while the distribution of the population over different household size-categories does not correspond to that of the census, then this can be corrected by multiplying w_{is} by N_s / n_s.

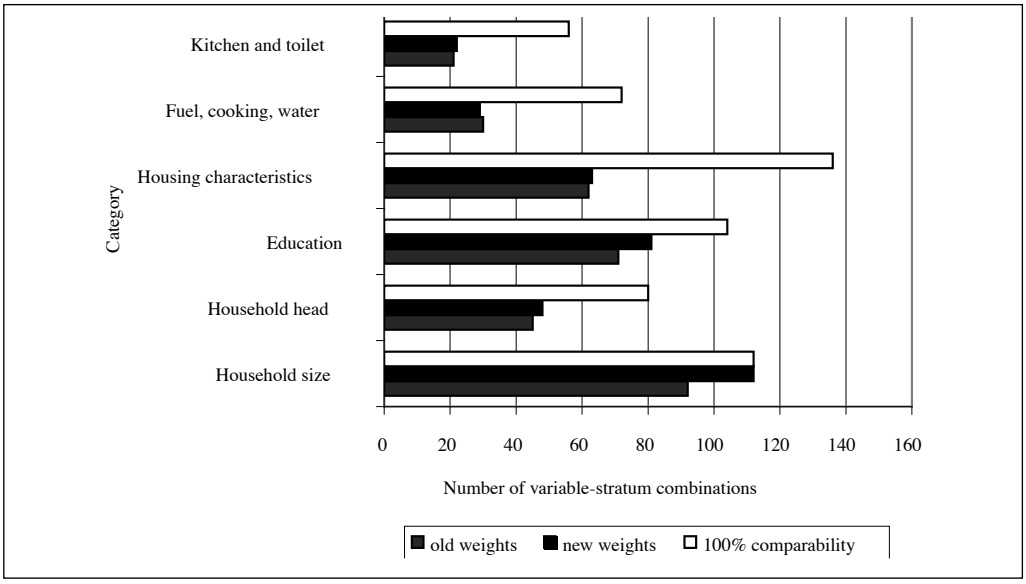
Though the procedure to correct for a bias due to household size related non-response is straightforward, in practice there may be reluctance to rely on re-weight-

ing because adjusting a survey in one dimension may make it less comparable to a census in others. However, if the adjustment is to deal with size bias due to non-response, then adjusting the survey weights should not only lead to improvements in comparability in household size, but in other dimensions as well. By considering whether re-weighting results in improvements – or at least not in deterioration – in the comparability with other variables, we have a test for the validity of the exercise. Clearly this test can only be applied if the census and survey reflect the same time period (which is the case in Uganda), and if variables are identically defined.

After comparing enumerator instructions and the way questions were phrased and coded between the census and the survey, 70 potentially identical variables were identified.[5] As the survey is representative at the stratum level (there are 8 strata), there are 560 possible variable-stratum combinations for which the census mean should lie within the 95% confidence interval of the survey. Having sorted the 70 variables into six variable categories, we present the number of 'matching' variable-stratum combinations by variable category in Figure 2. The figure also indicates the maximum number of matches attainable ('100% comparability'). By definition, after re-weighting all variables relating to household size pass the comparison test and comparability increases from 321 variable-stratum combinations to 355 (+10.6%). Comparability for the non-household size related variables increases from 229 to 243 variable-stratum combinations (+ 6.1%).[6] We therefore conclude that household size related non-response did affect the IHS and that the re-weighting procedure resulted in an improved set of household weights.

5: Even when variable definitions, enumerator instructions and coding are identical, different responses may arise, due to different enumerator training procedures, or because of minor differences in response codes.
6: Only with respect to two variables, marital status (head of household never married) and fuel use (household cooks using paraffin) does comparability deteriorate by one variable-stratum combination.

Figure 2: Number of variable-stratum combinations for which census mean falls within survey mean's 95% confidence interval



4. Discussion

Our analysis has shown how household size related non-response may lead to an unrepresentative sample and, consequently, biased estimates. Table 1 provides insight into the importance of re-weighting for poverty incidence, per capita consumption and the Gini coefficient.

Table 1: Comparison of consequences of reweighting on various indicators of household welfare

Domain	Poverty Incidence		Per capita consumption		Gini Coefficient	
	IHS, official	IHS, re-weighted	IHS, official	IHS, re-weighted	IHS, official	IHS, re-weighted
Urban	27.8 (2.4)	27.8 (2.4)	33158 (2063)	32534 (1699)	0.395 (0.03)	0.383 (0.03)
Central rural	54.3 (2.2)	54.1 (2.2)	18046 (638)	18131 (629)	0.329 (0.01)	0.330 (0.01)
East rural	60.6 (2.3)	60.6 (2.3)	15427 (480)	15460 (486)	0.321 (0.01)	0.322 (0.01)
North rural	73.0 (2.9)	72.3 (2.9)	13663 (632)	13899 (636)	0.330 (0.02)	0.331 (0.01)
West rural	54.3 (2.4)	55.0 (2.6)	16368 (500)	16256 (537)	0.309 (0.01)	0.311 (0.01)

Notes: The columns IHS official presents welfare estimates as released by the Uganda Bureau of Statistics. IHS re-weighted adjusts the IHS sampling weights for household size related non-response. Standard errors are in parentheses and are corrected for survey design.

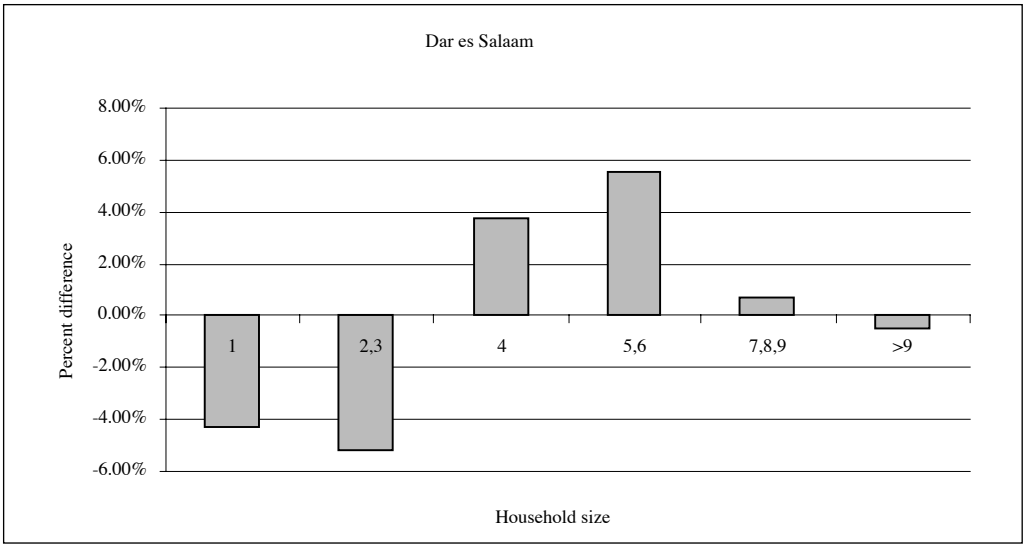
The table shows how re-weighting the IHS to adjust for household non-response has little effect on the various welfare estimates. This is encouraging in that it implies that official poverty estimates do not need to be revised. This absence of an impact of re-weighting on the welfare indicators can be traced to two aspects: (i) the fraction of poor one and two person households is small; and (ii) even after re-weighting, members from small households make up only between 8% and 9% of the total population. It should not be inferred from this lack of impact that re-weighting is superfluous. Whether this is the case depends on the research question at hand. For instance, if the interest were in the fraction of non-poor living in small households then re-weighting makes a significant difference (at the 95% level of confidence) as it increases the fraction from 39.3% to 45.1%.

Non-response is not related to household size alone. A distinct concern may be that survey participation varies with household income. Very poor and very rich households, for instance, may be less inclined to answer because of the high opportunity costs of their time. Household size related non-response may also be correlated with income. This could occur, for instance, when non-responding households are more likely to be non-poor because they are employed (and hence not at home when

the enumerator visits). In the latter case, correcting for non-response with the procedure outlined previously will not correct this bias. Mistiaen and Ravallion (2003) show how, in the presence of information about non-response rates, income related non-response may be corrected for. They present an application for the USA, showing a considerable under-representation of wealthy households. Applying their approach to the Ugandan survey we obtain an inverted-U shaped compliance-expenditure pattern with people in middle quintile groups more likely to comply than either the richest or the poorest. The difference in compliance rates is only marginal. After correcting for wealth related non-compliance the largest divergence we find for the poorest quintile an estimated true population proportion of 0.2097 (rather than 0.20); for the wealthiest quintile it is 0.1986. On the basis of this information, we conclude that household size related non-response is a more pressing issue than income related non-response.

So far we illustrated the importance of household size related non-response with data from Uganda. However, the issue is not unique to Uganda. In Tanzania, for instance, a pattern comparable to that presented in Figures 1 and 2 was detected when comparing the distribution of household sizes in the 2000/01 Household Budget Survey with that of the 2002 Population and Housing Census. There is an under-representation of small and very large households and an over-representation of medium sized households. Figure 3 illustrates the Tanzania pattern for its capital city, Dar es Salaam. We therefore conclude that many more surveys are likely to be affected by this easily remedied problem.

Figure 3: Difference in proportion of households of size x reported in the Tanzanian 2002 census and the 2000/01 Household Budget Survey (a negative number indicates under-representation in the survey).



Source: Adjusted from Kilama and Lindeboom (forthcoming)

5. Concluding Remarks

Our analysis has shown how household size related non-response may lead to an unrepresentative sample and, consequently, biased estimates. A straightforward way to deal with the bias has been presented in which the frequency distribution of household size categories in the survey is made to correspond precisely to the frequency distribution among those same categories in the census.

The problem of household size related non-response can largely be avoided by reconsidering the visiting strategy of enumerators. If enumerators make various attempts to enumerate a household and (re)-visit on different days of the week and at different times during the day, the probability of non-response by small households is likely to decline. Little can be said, a priori, which visiting strategy would be optimal. Yet a survey that would experiment with visiting strategies that vary across strata and which checks, ex post, the stratum level frequency distributions of household size in the survey and with those from the census would provide valuable information that would help avoid this kind of bias in the future. Identifying a visiting strate-

gy that would avoid household size related non-response is critical, especially as the ex post procedure sketched in this paper is only feasible when the survey and census are contemporaneous. As censuses are only implemented once every ten years, most surveys will not benefit from the procedure. Improving the household visiting strategy and thus reducing household size related non-response therefore remains an important objective to pursue.

References

Hansen, H.M., Hurwitz, W.N. and Madow, W.G. (1953). Sample Survey Methods and Theory. John Wiley & Sons, New York

Kilama, B. and Lindeboom, W. Where are the Poor in Tanzania (forthcoming).

Kish, L. (1965). Survey sampling. John Wiley & Sons, New York.

Lessler, J. T. and Kalsbeek, W.D. (1992). Nonsampling Error in Surveys. John Wiley & Sons, New York.

Mistiaen, J. and Ravallion, M. (2003). Survey Compliance and the Distribution of Income. World Bank: Policy Research Working Paper no. 2956.

Politz, A.N. and Simmons, W.R. (1949). An Attempt to Get 'not-at-homes' Into the Sample Without Call-backs, Journal of the American Statistical Association, 44, 9-31.

Thomsen, I. (1973). A Note on the Efficiency of Weighting Subclass Means to Reduce Effects of Non-response When Analyzing Survey Data, Statistik Tidskrift, 4, 278-283.

Measuring Progress Towards Global Poverty Goals: Challenges and Lessons From Southern Africa

Sebastian Levine[1]

Summary

This paper draws on the work in Lesotho and Namibia of tracking progress towards cutting poverty in half by 2015, which is the key poverty target of the Millennium Development Goals. The paper serves at least two purposes. Firstly, it outlines the steps and methodological considerations involved in selecting appropriate national indicators and targets for measuring income poverty using household surveys and poverty lines based on observed consumption patterns. Secondly, it highlights some practical lessons and challenges for policy makers in southern Africa when they attempt to access and analyse poverty data under less than ideal circumstances.

Key words

Income poverty, poverty line, household budget survey, Millennium Development Goals

Résumé

L'article se base sur le travail fait au Lesotho et en Namibie sur le suivi du progrès accompli dans la réduction de la pauvreté de moitié d'ici 2015, ce qui est l'objectif–clé de pauvreté des Objectifs du Millénaire pour le Développement. L'article est utile à plus d'un titre. Premièrement, il met en exergue les étapes et les aspects méthodologiques utilisés dans le choix des indicateurs et des objectifs nationaux appropriés pour mesurer la pauvreté monétaire en utilisant des enquêtes auprès des ménages et des seuils de pauvreté basés sur la structure de consommation. Deuxièmement, il met en exergue certaines leçons pratiques et défis dont les décideurs en Afrique australe doivent relever quand ils essaient d'évaluer et d'analyser des données de pauvreté dans des circonstances pour le moins idéales.

1: United Nations Development Programme, Private Bag 13329, Windhoek, Namibia. E-mail: sebastian.levine@undp.org. I am grateful for the insightful comments by Bjørn Wold and to Benjamin Roberts and Julian May for their many valuable contributions in Lesotho and Namibia. The views expressed in the paper are my own and not necessarily those of the United Nations Development Programme. Any errors and omissions are also mine.

Mots clés

Pauvreté monétaire, Ligne de pauvreté, Namibie, Lesotho, enquête budget auprès des ménages, Objectifs du Millénaire pour le Développement.

1. Introduction

Demand is growing for estimates of poverty that are better and more systematically collected and analysed at the sub-national, national and global levels. A key driver of this increased demand has been the advent of Poverty Reduction Strategy Papers or PRSPs. These are comprehensive and country-led strategies that set out a country's macroeconomic, structural and social policies and programs to promote growth and reduce poverty, and they are expected to include appropriate mechanisms for monitoring to ensure that progress can be measured (World Bank, 1999). A second source of increased demand for effective systems to monitor poverty has come from the global and national drive towards the Millennium Development Goals derived from the Millennium Declaration that was agreed to by all UN member states in 2000 (United Nations, 2001).[2] As part of this work a central challenge faced by countries that are reporting on the Millennium Development Goals is to define poverty indicators and tailor global goals and targets to reflect the national context (Vandemoortele, 2005).

This paper looks at the process of tailoring the Millennium Development Goal 1 and specifically the global target to "halve, between 1990 and 2015, the proportion of people whose income is less than one dollar a day." in two countries in southern Africa: Namibia and Lesotho. Both countries have recently completed national poverty reduction strategies as well as first rounds of progress reports on the Millennium Development Goals, and the two countries faced similar challenges in reporting on several of the Goals in particular the first (Government of Lesotho, 2004a; Government of Namibia, 2004).[3] The paper serves at least two purposes. Firstly, it outlines the steps and methodological considerations involved in selecting the appropriate indicators using national poverty lines and monitoring progress towards the first of the Millennium Development Goals. Secondly, it documents some practical lessons and challenges for policy makers in developing countries when they attempt to access and analyse poverty data under less than ideal circumstances.

2: The eight goals are: 1. Eradicate extreme poverty and hunger; 2. Achieve universal primary education; 3 Promote gender equality and empower women; 4. Reduce child mortality; 5. Improve maternal health; 6. Combat HIV/AIDS, malaria and other diseases; 7. Ensure environmental sustainability; 8. Develop a global partnership for development.

3: Lesotho's PRSP was finalised in 2004 (Government of Lesotho, 2004b). As a middle-income country Namibia is not eligible for concessional assistance from the World Bank and does not need to prepare a PRSP. However, poverty reduction remains a national priority within the country's National Development Plans of which the third edition is currently being prepared.

The paper is structured in the following way: Section 2 describes the two household surveys from Lesotho and Namibia and outlines the process of deriving poverty lines based on actual consumption recorded in the surveys. Section 3 presents some main results from the analysis of income poverty and inequality in Lesotho and Namibia. Section 4 looks at some implications for reporting on Millennium Development Goal 1 and the setting of poverty targets. Finally, Section 5 concludes.

2. Old Datasets and New Poverty Lines

In order to measure progress under Millennium Development Goal, 1 United Nations (2003) advises that indicators based on national poverty lines should be used. Specifically countries are advised to use the poverty headcount ratio, which is the proportion of the national population whose incomes are below the official threshold (or poverty line) set by the national government. This is instead of the original standard measure of the proportion of the population that lives on less than USD 1 per day (using purchasing power parties). The latter indicator is primarily useful for international comparisons, while the former allows for monitoring and analyzing poverty using a national standard based on the specific characteristics of the country. The initial challenge for national authorities in Lesotho and Namibia was therefore to establish national poverty lines.[4] However, even before that some data issues had to be examined.

In Lesotho a comprehensive Household Budget Survey (HBS) was conducted in 1994/95 and in Namibia a National Household Income and Expenditure Survey (NHIES) was conducted in 1993/94. In both cases, nation-wide stratified samples of more than 4000 households kept detailed diaries of income, consumption and expenditure over a full 12 month cycle, and were therefore ideal for analysis for welfare and poverty analysis. The two surveys were characterised by high levels of ambition and sophistication but had done little to inform policy making.

This is part of a general and system-wide problem. As one review of the national statistical system in Lesotho found: *"Although Lesotho also has a long track record of research and data gathering current environment is one in which data is neither trusted nor used. As a result, policy development in Lesotho shows little sign of being evidence based. In the absence of good information, choices may result in incorrect targeting, inefficient use of resources, and contradictory or less than optimal outcomes. For Lesotho, poverty monitoring is thus not simply a technical task of identi-*

4: Full detail on the methodologies and results in each country is captured in two separate studies (May, et.al. 2001b and van Rooy, et.al. 2006) funded and coordinated by UNDP in collaboration with the Central Statistics Offices in Lesotho and Namibia, and with technical assistance from the Universities of Kwa-zulu Natal, Port Elizabeth and Namibia, and the Human Sciences Research Council.

fying a bundle of indicators but must extend to rebuilding national confidence in information usage." (May, et.al, 2001a).

The 1994/95 HBS in Lesotho is a case in point, while data gathering and entering had been completed, processing and weighting of data had come to an abrupt end once the external technical assistance ran out. As a consequence the vast dataset was never analysed and accordingly, the results never officially released.[5] In order to investigate whether the dataset could be 'rescued' and used in the preparation of country's Poverty Reduction Strategy Paper and a progress report on the Millennium Development Goals half a decade later, UNDP supported a technical review of the dataset. The review process, subsequently captured in Wollard and Roberts (2001), pieced together the various data files and administrative records but found evidence of large human errors in the data entry process due to pervasive confusion between monthly and annual income among both respondents and data entry staff. Also, data on domestic consumption expenditure from own production was only captured sporadically. As a result of this technical review, the Lesotho Bureau of Statistics decided to re-enter, clean and validate the entire dataset from the original questionnaires. Obviously having to recapture the data almost a decade after the survey was conducted does not represent an ideal situation and the process introduces new elements of potential bias, but under the prevailing less than ideal circumstances there were few other options. Moreover, it was found that the older data could serve as useful baseline once a new survey scheduled for 2002/03 had been finalised.

In Namibia the situation was more straightforward. The 1993/94 survey was completed and published, although only after some delay (National Planning Commission, 1996). However, the survey results were only released in the form of basic tabulations and only a very limited analysis was conducted. Moreover, the poverty analysis used a food-ratio method as a poverty line identifying the "poor" as those whose expenses on food comprise 60 percent or more of their total expenditure. "Severely poor" were identified as those whose food expenses were 80 percent or more. This poverty line originates from "Engel's Law" which states that poor households devote a greater share of their budget to food compared to better-off households. However, using this food ratio approach is problematic. As noted by Ravallion (1992) the relationship between the food share and consumption will generally differ across households because of differences in the relative prices they face, demographic differences, or differences in preferences. Also, the income elasticity of demand for food can

5: The lack of analysis and usage was not only confined to poverty and inequality. Household surveys of this kind also provide critical information for estimating the consumption component of the National Accounts and are essential for updating the Consumer Price Index basket.

be very close to unity for poor households. Furthermore, stakeholders in Namibia expressed concern about the overly arbitrary nature of the existing poverty line.

Most countries that have official poverty lines define these in an absolute sense, interpreting them as a fixed standard of living (Lanjouw, 2001; May, 2001). Using the available data from the surveys on household incomes and consumption focus in Lesotho and Namibia focus turned to setting up an absolute poverty line based on a minimum food basket and allowing for essential non-food consumption.

Setting the poverty line

The first step in specifying the food basket of the absolute poverty line in both countries was to examine the actual food consumption patterns of a particular segment of the population. Looking only at households in the second to fifth deciles (the bottom decile of the distribution being discarded to avoid possible data errors) of the two household surveys, the average total household food expenditure for each food and beverage item was calculated. These were subsequently ranked according to the percentage of households consuming the item and the average amount spent per item. Having done this, the top 30 *purchased* items were selected for inclusion in the basket of goods to be used for calculating the poverty line. Focusing exclusively on the consumption of households in the lower deciles of the expenditure distribution ensures that expensive, luxury food items are not heavily represented in the basket. Moreover, by basing the composition of the basket on existing consumption patterns, the combination of food and beverage items included in the basket is consistent with local tastes and preferences. In addition to these items that were purchased by the household, a bundle of commonly consumed own produce food items were included. The components of the Lesotho basket are shown in table 2.

Table 2: Components of Lesotho's food basket

Thirty Food and Beverage Items from Purchases				
Beans	Chicken	Maize meal	Powdered milk	Sterilised milk
Beef	Cooking fat	Malt	Rice	Sugar
Beers	Cooking oil	Mineral water/soft drinks	Salt	Tea
Bread	Eggs	Mutton	Samp	Tinned fish
Bread flour	Fresh milk	Offal	Sorghum meal	Tomatoes
Cabbages	Home brew	Potatoes	Soups, all types	Wheat meal
Ten Food Items from Own Produce				
Maize	Pumpkin	Cabbage	Wild vegetables	Chicken
Sorghum	Radish	Beans	Wheat	Spinach

Source: May, et.al. (2001b).

Having selected the basket of items, the total expenditure per item for all households for each of the twelve months of the year was calculated. This calculation was done separately for households living in the urban areas (Maseru and Windhoek) and all other households. *Total monthly expenditure* per item was computed by summing up each of the values. This was subsequently converted into *average monthly item expenditure per household* by dividing by the number of households in either the urban sample or the other households sample and *average monthly item expenditure per capita* by dividing average household item expenditure by the mean household sizes. This calculation should ideally accommodate geographically-determined price differentials. Unfortunately, neither of the surveys in the two countries collected community-level price data. This made it difficult to factor in locality-specific price differences in determining our poverty line. As an alternative, in Lesotho consumer price data collected for urban and other areas were used as a best approximation. In Namibia, this option was constrained by the fact that the consumer price data until very recently has been exclusively urban-based. In order to accommodate rural-urban price differentials in the estimation of the Namibian poverty line, use was made of price data captured from three rural regions as part of a 1993 study on household subsistence levels (Multidisciplinary Research and Consultancy Centre, 1994). A ratio of rural to urban prices was calculated for the ten items included in the survey that coincide with the food basket. Then the cost per gram for urban prices was inflated to arrive at the cost per gram for the items in rural areas.[6]

The next step taken was to convert the average item expenditures per capita into *average number of grams per item per capita* (Ravallion, 2004). The conversion is undertaken by dividing the average item expenditure per capita by cost per gram for each of the items, yielding monthly number of grams per item per adult.

$$\text{Grams per month} = \frac{\exp}{1} \div \frac{cost}{gram}$$

With the average number of grams consumed per adult on each of the items in the basket every month calculated, this was converted into a calorific value by determining the calorie content per gram of each of the items. The product of average number of grams consumed per capita and the calorie content per gram of item gives the average number of calories obtained from each source (per capita per month).

6: More work is needed to determine spatial differences in consumption patterns which could lead to setting up regional poverty lines based on different consumption baskets (Bidani and Ravallion, 1994).

The sum of all the daily item calorific values for both urban Maseru households and all other households shows an *average per capita calorie consumption* of 989 kcal, of which 965 kcal comes from the purchased items and 24 kcal from the own produced goods. For Namibia the weighted average of the daily item calorific values for both urban and rural households shows an average per capita calorie consumption of 696 kcal, of which 454 kcal comes from the purchased items and 242 kcal from the fifteen produced goods. When setting the poverty lines, both Lesotho and Namibia settled on a minimum of 2200 kcal with reference to FAO/WHO recommended daily allowance, which is defined as the amount needed to maintain health, growth, and an appropriate level of physical activity (WHO, 1985).

Then, to set the food basket, each of the gram quantities is multiplied by (2200/average per capita calorie consumption): 2.22 (2200/989) in Lesotho and 3.16 (2200/696) in Namibia. Since it cost Maloti 23 per month to purchase the 989 kcal per capita per day in 1994/5, it would have cost Maloti 23 multiplied by (2200/989) or Maloti 51 for a household to purchase 2200 kcal per capita per day in 1994/5 in Lesotho. Similarly for Namibia where it cost Namibian $ 24 per month to purchase the 696 kcal per capita per day in 1993/4, it would have cost Namibian $ 24 multiplied by (2200/696) or Namibian $ 77 for a household to purchase 2200 kcal per capita per day in 1993/4. This way Maloti 51 and Namibian $ 77 constitutes the *food* poverty line for Lesotho and Namibia, respectively in the two years of survey.[7]

Even though having sufficient resources within the household to meet food requirements is critical in terms of determining the threshold below which households are classified as poor, there is a strong argument that states that this alone is not adequate to define the poverty line. Households that can afford to meet the food requirements of all its members but who lack the resources to purchase clothing and shelter, for example, are likely to be considered deprived in a very basic sense (Ravallion, 1994). Recognising this, non-food expenditure was included in poverty lines for Lesotho and Namibia.

As with the food expenditure component, the approach adopted to derive the non-food component of the poverty line was based on observed consumption behaviour in both countries. First the median non-food expenditure per capita was calculated for households with per capita *total* expenditure in a small interval (plus or minus one percent) around the food poverty line. Successively larger intervals were selected, a total of five times so that the largest interval was plus or minus five percent, and a

7: The two national currencies of Lesotho and Namibia, the Maloti and the Namibian $, are pegged at 1:1 to the South African Rand.

simple average was taken of the five observations of median non-food expenditure per capita around the food poverty line. The above process was followed because of possibility that none or very few of the households in the survey sample were likely to have per capita total expenditure exactly equal to the food poverty line. The amount derived from the process was then added to the food poverty line to yield the final poverty line (May, et.al. 2001b and van Rooy, et.al. 2006). In Lesotho the average of the five median values of non-food expenditure per capita came to Maloti 27. Therefore, the final poverty line was (Maloti 51 plus Maloti 27) Maloti 78 per capita per month in 1994/95 prices. Similarly for Namibia the average of the five median values of non-food expenditure per capita came to Namibian $ 30. Therefore, the final poverty line was (Namibian $ 77 plus Namibian $ 30) Namibian $ 107 per capita per month in 1993/94 prices.

It should be noted that because of database limitations the poverty lines are expressed in per capita terms. The implication is that all members of a household are assumed to have equal food and non-food requirements and that there are no economies of scale in larger families. This is clearly unrealistic and future analysis should at least test the robustness of alternative equivalence and economies of scale parameters (May and Roberts, 2005).

3. Key Results from the Household Surveys

This Section presents some of the key results from the two surveys in Lesotho and Namibia using standard measures of poverty and inequality. While focus initially centred on the poverty headcount as the share of the population under the poverty line, following Foster, Greer and Thorbecke (1984) three measures that cover the incidence, depth and severity of poverty were derived. Their values are summarised in Table 3. The first key analytical conclusion is that at the time of the surveys the vast majority of the populations in both countries lived in income poverty. For Lesotho, the *incidence of poverty*, or the poverty head count, was calculated at 58 percent for 1994/95 compared to near 65 percent in Namibia for 1993/94 using the poverty lines developed specifically based on the observed patterns of consumption in the two countries.

The second key finding is that the *depth of poverty* in Namibia in 1993/94 was 0.34 and 0.35 for Lesotho in 1994/95, which can be interpreted to mean that on average the expenditures of the poor are 34 and 35 percent, respectively, below the poverty line. Based on the poverty gap for Namibia, it was estimated that if perfect cash

transfers were possible it would have cost Namibian $ 50 million per month or Namibian $ 606 million per year in 1993/94 prices to lift all the poor individuals out of poverty.[8] Converting this into a December 2003 equivalent, the cost of raising all poor individuals above the poverty line would be Namibian $1.2 billion or about 10 percent of the national budget.

Table 3: Summary of results for income poverty in Lesotho and Namibia

	Lesotho	Namibia
Survey year	1994/95	1993/94
Food poverty line	M 51	N$ 77
Poverty line	M 78	N$ 107
Incidence	58.3 %	64.8 %
Depth	0.35	0.34
Severity	0.26	0.21

Sources: May, et.al. (2001b) and van Rooy, et.al (2006).

The *severity of poverty*, or squared poverty gap, takes inequality among the poor into account by weighting the poverty gaps of the poor so the poorer the individual the greater is their weighting. Accordingly, a transfer from a poor to an even poorer individual would reduce the severity index, whereas a transfer from a very poor to a less poor individual would increase the index. Of Lesotho's ten administrative districts, the incidence of poverty in 1994/95 was found to be highest in predominantly mountainous districts of Mokhotlong (75.4 percent), followed closely by Mohale's Hoek (74.9 percent), Quthing (72.7 percent) and Thaba Tseka (72.3 percent). The same applies to both the depth and severity of poverty. Conversely, the incidence of poverty is considerably below average in the capital district of Maseru, where only 39 percent of households are poor. The incidence, depth, and severity of poverty are also generally below the national average in the mostly lowland/foothill districts of Leribe and Berea. The higher incidence of poverty in these districts is related to the higher incidence in the mountain areas as a whole.

8: A perfect transfer of cash to the poorest to lift them above the poverty threshold is obviously fraught with both technical and political difficulties. In Namibia, however, there is growing pressure to introduce a Basic Income Grant to all citizens irrespective of age. The grant is designed to eliminate extreme poverty and is financed through an increase in the Value Added Tax (Haarmann and Haarmann, 2005).

For Namibia vast spatial differences in the distribution of the poverty is reflected in Table 4. For example, in the northern region of Ohangwena 86 percent of the population live under the poverty line compared to 24 percent in Khomas where the capital Windhoek is located. *The poverty share* shows that out of all the poor people in Namibia almost one in five live in Ohangwena and together with neighbouring regions, Oshana and Omusati, these three regions comprise almost half of the nation's poor. However, in contrast to the relative ease in interpreting the situation of the rural-urban distributions in Namibia, there may be instances where the picture becomes less clear. For example, examining the poverty incidence by region reveals that a number of the distribution functions cross within feasible choices of the poverty line. While it is clear that Ohangwena remains the poorest region and Khomas and Erongo the least poor regions for virtually all possible values of the poverty line, for the other regions it is difficult to make an unambiguous ranking. As such, the ranking of the poverty headcount or incidence by region will vary substantively depending on where the poverty line is set (van Rooy et.al, 2005).

Table 4: Namibia poverty measures disaggregated by region in 1993/94

	N	Incidence (P0)	Depth (P1)	Severity (P2)	Poverty Share
Caprivi	91434	78.79	0.44	0.29	8.00
Erongo	74395	38.35	0.16	0.08	3.17
Hardap	54206	45.39	0.24	0.16	2.73
Karas	54114	44.95	0.23	0.14	2.70
Khomas	161754	23.50	0.10	0.06	4.22
Kunene	59029	76.91	0.41	0.27	5.04
Ohangwena	190858	85.83	0.50	0.34	18.19
Kavango	125033	70.93	0.33	0.19	9.85
Omaheke	47101	68.63	0.33	0.19	3.59
Omusati	153030	75.79	0.39	0.24	12.88
Oshana	161491	75.79	0.39	0.24	13.59
Oshikoto	116134	77.36	0.43	0.28	9.98
Otjozondjupa	100438	54.36	0.28	0.18	6.06
Namibia	1389017	64.83	0.34	0.21	100.00
Urban	401325	33.94	0.15	0.08	15.13
Rural	987691	77.38	0.42	0.27	84.87

Source: van Rooy, et.al (2006).

Inter-temporal changes in poverty

One of the key objectives of poverty monitoring is to track developments over time. In Namibia poverty estimates prior to 1994/95 were based on various partial surveys, proxies and qualitative information that does not permit rigorous poverty analysis (World Bank, 1995). Consequently it is not possible to compare the results from the 1993/94 NHIES to any data in the past. A smaller Levels of Living Survey was conducted later in 1999 (Government of Namibia, 2001) but unfortunately comparability was again not possible because of differences in survey methodologies including seasonal coverage and definitions of household membership.[9] Fortunately such problems were not encountered in Lesotho where there the 1994/95 survey was

9: This survey was never formally approved by the National Planning Commission under whose authority it was carried out and appears to represent yet another example of a survey that was planned and executed but without any real or lasting policy impact.

modelled in terms of objectives and survey design explicitly on a previous survey carried out in 1986/87. As shown in Table 5 the data collected for the 1986/7 survey shows that 58.8 percent of Lesotho's population were living in poverty. The overall incidence of poverty had not changed significantly almost a decade later.

Table 5: Changes in poverty in Lesotho 1986/87-1994/95

	1986/87	1994/95
Incidence	58.8	58.3
Depth	0.33	0.35
Severity	0.23	0.26
Gini index	60	66

Sources: May, et.al. (2001).

Nevertheless, the results also revealed that the depth of poverty increased between 1986/87 and 1994/95. This implies that despite a marginal decline in the incidence of poverty over the period, those that are poor are on average further below the poverty line in 1994/95 than was the case in 1986/87. The results also indicates that the severity of poverty increased between 1986/87 and 1994/95. This, together with an increase in the depth of poverty and in the Gini index (from 60 to 66), reveals how looking at the changing incidence of poverty alone can be misleading as it misses the critical point that the poor have gotten poorer over the period.

It should be noted that the inter-temporal analysis in Lesotho is based on inflating 1986/87 expenditure values to 1994/95 prices using the CPI data and does not deal with the important question of changing consumption patterns between the surveys as a result of changes in preferences, relative prices and availability of goods. To address this future work needs to focus on updating the poverty line using CPI

Measures of inequality

The two countries are not only characterised by high levels of poverty but also by extremely high levels of inequality. Table 6 provides income distribution measures for the population based on monthly expenditure per person for Namibia. It shows that the poorest 20% of the population receives 2.45% of total income, while the top 20% receives 71 percent. An even worse picture emerges from the data in Lesotho. Here the poorest 20% of the population receives only 1.11% of income, while the top 20% receives 78% of incomes (as measured by observed consumption). The Gini indices for Lesotho and Namibia are 66 and 63, respectively. According to UNDP (2004) Scandinavian countries have Gini indices of around 25 and several countries in Latin America and in Africa have indices close to 60 but no other country recorded a higher Gini index than those found in the studies of Namibia and Lesotho indicating that these two countries are among the most unequal in the world. data and rebasing the poverty line once the new survey becomes available (Ravallion, 1998; May and Roberts, 2005).

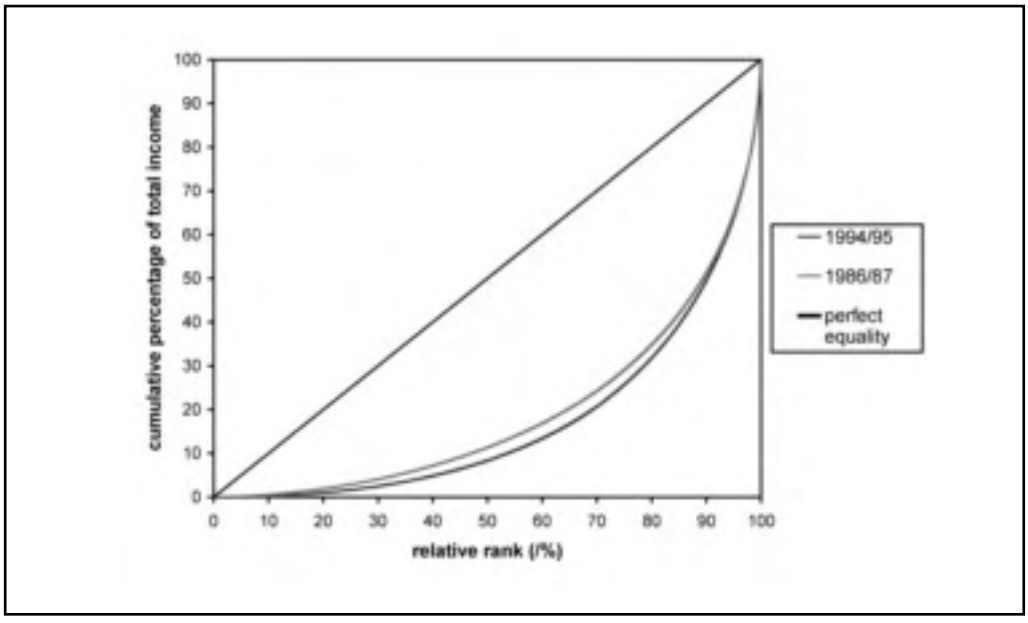
Table 6: Share of national consumption by population deciles in Namibia, 1993/94

Decile	Expenditure share	Cumulative Expenditure Share
Decile 1 (poorest 10 percent)	0.88	0.88
Decile 2	1.56	2.45
Decile 3	2.11	4.56
Decile 4	2.69	7.25
Decile 5	3.42	10.67
Decile 6	4.35	15.02
Decile 7	5.75	20.77
Decile 8	8.28	29.05
Decile 9	13.57	42.62
Decile 10 (top 10 percent)	57.38	100.00
Share of poorest 20% (quintile) in national consumption	2.45	2.45
Gini index	**63**	

Source: van Rooy, et.al (2006).

The Gini index can be seen in relation to the Lorenz curve whereby the proportion of the population ranked from poorest to richest is plotted on the *x*-axis and the percentage of income accruing to the bottom *x* percent of the population is shown on the *y*-axis. Figure 1 shows the Lorenz curve for Lesotho, using the two household surveys. Everyone is ranked according to their income, and then cumulative income is plotted against these ranks. The straight (45° degree) line is the "line of perfect equality". Thus, the higher the Lorenz curve, the lower is inequality.

Figure 1: Lorenz Curves for Lesotho



Source: May, et.al. (2001).

Decomposition of within-group inequality using the Theil measure shows that inequality is pervasive in all of the Lesotho's districts. The data indicates that there is considerable inequality within the different area-types, but also some differentiation between area-types, with 16-22 percent of inequality being accounted for by between-group inequality. Because rural areas contain a larger share of the population than of total income, the Theil-L (which is weighted by population shares) gives greater emphasis to the share of rural areas in overall inequality.

4. Reporting on Millennium Development Goal 1 and Setting Targets for Poverty Reduction

A key purpose for reanalysing the two datasets in Lesotho and Namibia was to fill a key data gap for reporting on the Millennium Development Goals, for both countries in 2004, and in the case of Lesotho, to feed into the Poverty Reduction Strategy Paper. Namibia did finalise its progress report on the Millennium Development Goals in 2004. However, the national authorities in Windhoek decided not to use the data on income poverty. The methodology for deriving the new poverty line remains widely supported from within government structures, indeed the process was led by senior officials, and the results were never questioned. The main problem was of political nature. There was fear that because the poverty head count was significantly higher using the new absolute poverty line compared to the older one based on the food ratio, releasing the poverty figures would embarrass the political leaders in an election year. However, even if the data ended up not being used for the immediate purposes of MDG reporting all is not as discouraging. The new dataset is being used for analysis for instance in the Social Accounting Matrix being developed locally. The data is also referred to in the national Poverty Monitoring Strategy and will serve as a benchmark once the results from the 2004/05 household survey are released. In Lesotho, the data found more immediate use and formed the basis for analysing linkages between economic growth and poverty reduction in the PRSP and for setting national poverty targets.

Setting targets for poverty reduction in Lesotho

The two datasets of household poverty in Lesotho made it possible to estimate the future incidence of poverty assuming certain levels of economic growth, and more specifically assess under which conditions of economic growth and inequality the country will be able to attain the Millennium Development Goals of halving poverty by 2015. Using the incidence of poverty in 1987 and 1995 and data on gross national product (GNP) per capita, a poverty elasticity of –0.12 was calculated for Lesotho (May et al, 2001). This effectively means that, for the period under consideration, a 1% increase in per capita GNP resulted in a corresponding decrease in the incidence of poverty by 0.12%. This observed reduction in poverty can be attributed to two factors, a pure growth effect of –1.37% and a pure inequality effect of 1.25% as shown in Table 7. Therefore, if inequality had not increased in Lesotho between 1987 and 1995, each 1% growth in GNP per capita would have reduced poverty by 1.37%.[10]

10: This figure can be interpreted as the opportunity costs of increased inequality on poverty and once again highlights the central challenge of making economic growth in Lesotho distinctively more pro-poor.

Table 7: Poverty and economic growth in Lesotho 1986/87-1994/95

	Poverty Incidence				Explained by	
	1986/87	1994/95	Annual % Change	**Poverty Elasticity**	Growth	Inequality
Headcount	58.8	58.3	-0.11	-0.12	-1.37	1.25

Source: May, et.al. (2001).

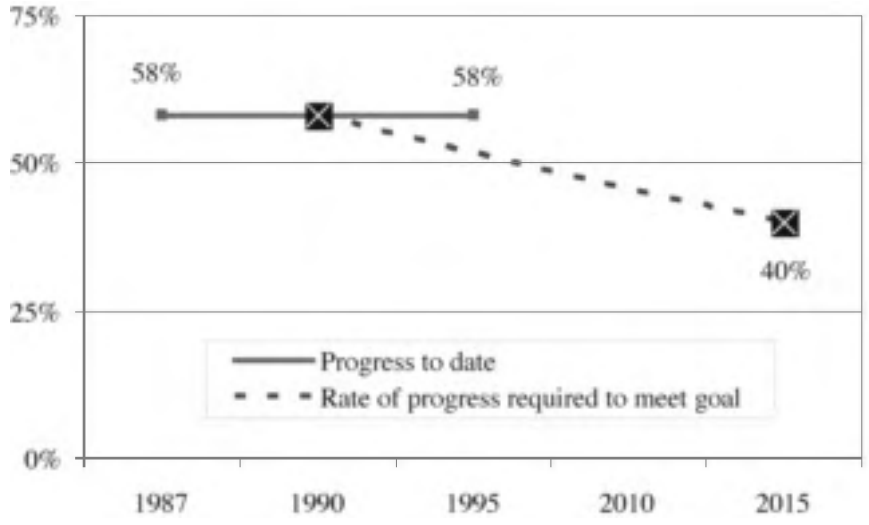
Having obtained an estimate of the poverty elasticity, it was possible to determine the effect of various scenarios on effectively reducing the high level of poverty in the country. For instance with an annual growth rate of 4% per capita GNP from 1995 onwards – a growth that is not wholly inconsistent with what was achieved between the 1980s and early 1990s – poverty would decrease from 58% in 1995 to 53% by 2015. This was still a long way short of the envisaged poverty incidence of 29% required if poverty was to be cut in half. What then would it take the country, in terms of economic growth, to meet the global target? The data revealed that, with a poverty elasticity of –0.12, no realistic growth rates would enable Lesotho to achieve the target of halving the incidence of poverty by 2015 (May et. al, 2001). This type poverty elasticity rests on at least two assumptions. First, that additional growth will be distribution neutral and thus equally benefit everyone in the population, which was not supported by the historical trends. Secondly, that the incidence of poverty in the country has not changed since 1995, which was considered unlikely in Lesotho given slow economic growth, growing levels of unemployment and a high prevalence of HIV.

Despite these assumptions the message for Lesotho's policy makers was clear. Accelerating economic growth may be a necessary but it is highly insufficient when it comes to poverty reduction. The dual challenge remains to make economic growth pro-poor in a strict sense by promoting growth as well as make this growth more biased towards the poorest. The analytical equivalent would be to raise the poverty elasticity, and using estimated levels for other high inequality counties (Hanmer et al 1999; White 1999) Lesotho decided to use a poverty elasticity of –0.5. Under this implied improved responsiveness of poverty to economic growth, an annual per capita GNP growth rate of 4% was found to substantially reduce the incidence of poverty in Lesotho from 58% in 1995 to 39% in 2015. Similarly, with a poverty elasticity to –0.5, an annual per capita GNP growth rate of 6.8% would be required to halve poverty by 2015 (May et. al, 2001). The authorities decided to use the former estimate as its main national target for reporting on the Millennium Development Goals as depicted on Figure 2.

The process of tailoring the global poverty target in Lesotho thus led to a reduction of the target, from cutting poverty in half to reducing poverty by one third. This is a controversial choice and it is symptomatic of a global reality in which Africa is slipping further and further behind the agreed global minimum standards for progress in human development (United Nations, 2005). However, such adjustments should not be viewed as surrendering on the Millennium Declaration and must not lead to complacency by the national governments and their international development partners. Instead it should be seen as part of an essential process whereby the goals and targets agreed by the international community at a global level are taking root at national level.

In Lesotho the global poverty goals now form an integral part the first PRSP (Government of Lesotho, 2005). This and subsequent strategies will need to justify the change in the target and detail measures for making the economic policy making more pro-poor as implied by the adopted poverty elasticity. The poverty data may have been outdated but at least it helped initiate this long overdue discussion.

Figure 2: Lesotho's national poverty target



Source: Government of Lesotho (2004).

The same is needed in the case of Namibia, which has also been caught in a low growth high inequality conundrum. There is only an emerging tradition for approaching economic development and poverty reduction from a quantitative and analytical perspective, and thus policy makers have been constrained in designing and prioritising economic policies according to their impact on the overall development objective of eradicating poverty. This can help explain, though hardly justify, that the extensive NHIES was not used as a basis, or at least to inform, the process of preparing the country's first National Development Plan, which was being drafted at the same time. There is a real risk that this situation is about to repeat itself as the third National Development Plan is currently being prepared in the absence of the final poverty analysis from the new household survey. Synchronising the analysis-to-policy-design cycle must be considered a key priority.

5. Conclusion

This paper has outlined the basic theory on poverty measurement with emphasis on approaches relevant for country level monitoring of income poverty, and described the process of deriving poverty lines under less than ideal circumstances based on actual consumption recorded in two household surveys in Lesotho and Namibia. Such analysis was necessary for reporting on the MDGs, and in Lesotho the data was used to establish national long term targets for the country's PRSP. In Namibia, however, the poverty analysis did not make it into the MDG progress report and has only recently been published as an academic paper (van Rooy, et.al, 2006). The reason not to officially adopt and publish the findings was taken not because authorities questioned the results or the methodology, but because the attention that the publication of poverty data would generate was deemed politically inappropriate in an election year. Once again we are reminded that poverty analysis is not just a technical issue but often also a political one.

However, new household surveys in both Lesotho and Namibia should be able to benefit from the established base lines to reveal new trends in poverty and inequality in the last decade and to guide policy makers in the direction of more pro-poor economic and social policies. Also, through the technical assistance provided to the exercises of revitalizing the data, setting up poverty lines and conducting preliminary analyses capacity strengthening have taken place for statisticians at the two Central Statistics Offices using home grown data sources. Focus in the future should be on putting the methodology to use on the datasets from the recently completed surveys and to build on the analysis to strengthen the robustness of the results especially by

ensuring that the poverty lines are updated to reflect changing consumption patterns over time, household composition and spatial differences.

The experiences in Lesotho and Namibia also point to some more profound lessons. Most importantly it is highlighted how poorly some countries are prepared to report on and set targets for global poverty goals agreed by the international community and the analytical challenges the same countries are faced with when attempting to draw up anti-poverty policies and monitoring systems. There is a general need to strengthen the culture of evidence-based policy making so that in the future policy makers take appropriate measures to ensure that data on poverty, identified as the central policy challenge, is systematically gathered, analysed and put to use.

References

Bidani, B. and M. Ravallion (1994) 'How Robust is a Poverty Profile?' World Bank Economic Review, Vol. 8.

Foster, J.E., Greer, J. and Thorbecke, E. (1984). A Class of Decomposable Poverty Measures. Econometrica, Vol 52:761-6.

Grinspuin, A. (ed) (2001). Choices for the Poor, UNDP, New York.

Government of Lesotho (2005). Poverty Reduction Strategy Paper, Ministry of Finance and Development Planning, Maseru.

Government of Lesotho (2004a). Millennium Development Goals Report for Lesotho, Ministry of Finance and Development Planning, Maseru.

Government of Lesotho (2004b). Poverty Reduction Strategy 2004/05-2006/07, Ministry of Finance and Development Planning, Maseru.

Government of Namibia (2004). Millennium Development Goals Report for Namibia, National Planning Commission, Windhoek.

Government of Namibia (2001). Levels of Living Survey, National Planning Commission, Windhoek.

Government of Namibia (1996). Living Conditions in Namibia, National Planning Commission, Windhoek.

Hanmer, L., De Jong, N., Kurian, R. and Mooij, J. (1999). Are the DAC Targets Achievable? Poverty and Human Development in the Year 2015. Journal of International Development, No.11, pp.547-563.

Haarmann, C., and Haarmann, D. (2005). The Basic Income Grant in Namibia, Desk for Social Development, Windhoek.

Kakwani, N. (2004), New Global Poverty Counts, In Focus, September 2001, UNDP International Poverty Centre, Rio de Janeiro.

Lanjouw, J.O. (2001). Demystifying Poverty Lines,
http://www.undp.org/poverty/publications/ pov_red/Demystifying_Poverty_Lines.pdf

Lipton, M. (1983). The Poor and the Poorest: Some Interim Findings, World Bank Discussion Papers, No 25, World Bank, Washington, DC.

May, J., Kriege, D., Mochebelele, M., Mokitimi, N., Roberts, B. (2001a). Towards a Poverty Monitoring System in Lesotho, Ministry of Development Planning, Maseru.

May, J. and Roberts, B., (2005). Poverty Diagnostics Using Poor Data: Strengthening the Evidence Base for Pro-poor Policy Making in Lesotho. Social Indicators Research (2005) 74: 477-510.

May, J., Roberts, B., Moqasa, G., and Woolard, I. (2001b). Poverty and Inequality in Lesotho. UNDP, Maseru.

Multidisciplinary Research and Consultancy Centre (1994). Household Subsistence Levels: Three Selected Communities. Windhoek: University of Namibia.

Ravallion, M. (1998). Poverty Lines in Theory and Practice. Living Standards Measurement Study Working Paper 133. Washington DC: World Bank.

Ravallion, M. (1994). Poverty Comparisons (Chur: Harwood).

Ravallion, M. (1992). Poverty Comparisons: a Guide to Concepts and Methods. Living Standards Measurement Study Working Paper 88. Washington DC: World Bank.

van Rooy, G., Roberts, B., Schier, C., Swartz, J., and Levine, S. (2006). Poverty and Inequality in Namibia, University of Namibia, Windhoek.

UN Millennium Project (2005). Investing in Development: A Practical Plan to Achieve the Millennium Development Goals, New York.

UNDP (2003). Monitoring Poverty, Essentials, No. 10, May 2003. UNDP Evaluation Office, New York.

UNDP (2004). Human Development Report 2004, Oxford University Press, New York.

United Nations (2005). The Millennium Development Goals Report, 2005, United Nations, New York.

United Nations (2003). Indicators for Monitoring the Millennium Development Goals, United Nations, New York.

United Nations (2001). Road Map Towards the Implementation of the United Nations Millennium Declaration Report of the Secretary-General, September 2001, United Nations, New York.

Vandemoortele, J. (2005). Ambition is Golden: Meeting the MDGs, Development, 2005, 48(1): 5-11.

White, H. (1999). Global Poverty Reduction: Are we Heading in the Right Direction? Journal of International Development, No.11, pp.503-519.

World Health Organization. (1985). Energy and Protein Requirements. Report of a Joint FAO/WHO/UNU Ad-hoc Expert Consultation. WHO Technical Report Series, No. 724, Geneva

World Bank (1995). Namibia: Poverty Alleviation with Sustainable Growth, Washington.

World Bank (1999). Poverty Reduction Strategy Papers – Operational Issues, December 10, 1999, International Monetary Fund and International Development Association, Washington.

Woolard, I. and Roberts, B. (2001). Assessment of the 1994/95 Household Income and Expenditure and Consumption Survey, Ministry of Development Planning, Maseru.

The Role of Household Surveys in Poverty Reduction Efforts: A Case of the Uganda National Household Survey Programme

James Muwonge[1]

Summary

Uganda's Poverty Eradication Action Plan (PEAP) and the Millennium Development Goals (MDGs) have put a lot of focus on the statistical systems that must produce the data to monitor the PEAP and the MDGs indicators. The focus and direction of the data requirements in Uganda has greatly evolved and prompts government and other key stakeholders in the development process to establish a sustainable monitoring system of interventions. The PEAP and the MDGs among other plans require quality and comprehensive data for regular progress monitoring and review. Nationally representative household surveys undertaken by the Uganda Bureau of Statistics (UBOS) have contributed greatly to the policy debate and development in Uganda. The paper provides the UBOS's experience in meeting the various data demands at the national and international levels through this survey programme..

Keywords

Poverty Eradication Action Plan, Household Survey Programme, Qualitative and Quantitative approaches to poverty monitoring, the long term household survey programme.

Résumé

Le Plan d'Action d'Eradication de la Pauvreté (PAEP) de l'Ouganda et les Objectifs du Millénaire pour le Développement (OMD) ont mis une grande priorité sur les systèmes statistiques qui doivent produire des données pour le suivi des indicateurs du PAEP et des OMD. La priorité et l'orientation de la demande des données en Ouganda ont fortement évolué et ont incité le gouvernement et d'autres principales parties prenantes dans le processus de développement à établir un système durable du suivi des interventions. Le PAEP et les OMD parmi tant d'autres plans requièrent des données exhaustives et de bonne qualité pour faire le point et un suivi régulier des progrès accomplis. Des enquêtes auprès des ménages, representatives au plan na-

1: Principal Statistician, Uganda Bureau of Statistics, P.O.Box 7186, Kampala, Uganda
james.muwonge@ubos.org

tional, entreprises par le Bureau Ougandais de Statistique (UBOS) ont contribué largement au débat politique et au développement en Ouganda. L'article présente l'expérience de l'UBOS pour satisfaire diverses demandes de données aux niveaux national et international grâce à ce programme d'enquêtes.

Mots clés

Plan d'Action pour l'Eradication de la Pauvreté, Programme d'Enquêtes auprès des Ménages, Approches quantitatives et qualitatives pour le suivi de la pauvreté, Le programme de long terme d'Enquêtes auprès des Ménages.

1. Background

Discussions of the need for reliable data for policy-making have reached an almost unprecedented level at both national and international levels. Uganda's Poverty Eradication Action Plan (PEAP) - the national development framework, and the Millennium Development Goals (MDGs) have focused on statistical systems that must produce data to monitor the PEAP and the MDGs indicators.

Utilization of data for planning is not new in Uganda but most of the information used in the past was mainly from administrative sources. These included financial and economic data series that were and continue to be produced as part of the routine record keeping. Since 1997, Uganda has undertaken many reforms that have widely affected the economy and the welfare of Ugandans in general. Overtime, the demand for data has also evolved in line with these changes. The demand has shifted from being simple counting exercises to complex programs designed to determine the causes of individual and household behavior and the effect of government policies on the population's choices and welfare. Indeed, the current statistical systems in the world represent the result of centuries of actions, research, and methodological improvements in the world's ability to understand the interaction of various forces, on the socio-economic welfare of individuals, households, and countries. Apparent improvements in data must take into account the existing systems, their users, and the traditional data and emerging needs therein.

The importance of household survey statistics for national development cannot be over emphasized as it is needed to indicate among other things the welfare level to guide policy makers in framing socio-economic developmental plans and initiate interventions for improving people's socio-economic conditions. The discussion

of the need for more and better data for policy making has reached an almost unprecedented level at both national and international levels. Uganda's PEAP and the MDGs have greatly focused on the statistical systems that must produce the data to monitor the PEAP, MDGs and other development indicators. The Uganda Bureau of Statistics (UBOS) has contributed to both the understanding of various phenomena through the household survey programme and specifically to the debate on the effects on welfare levels of programmes on the population.

The availability of information on a regular basis is important in evaluating the dynamics of various economic factors which occur as a result of economic development. Information from the household surveys have provided the foundation and monitoring mechanism for the poverty eradication efforts in Uganda. For instance, Uganda has been able to monitor and measure trends of poverty, and changes that occur among the different population groups using the comprehensive datasets from nationally representative household surveys. According to the information from household surveys, the percentage of the population below the poverty line decreased from 56 percent in 1992 to 44 percent in 1997 and further to 38 percent in 2003. Such information collected through household surveys has influenced the Poverty Eradication Action Plan (PEAP) revision process in Uganda.

2. The National Development Framework

The Poverty Reduction Strategy Papers (PRSPs) have been used by the World Bank (WB) and International Monetary Fund (IMF) as a pre-condition for development assistance to low income countries to qualify for external financing and debt relief. The PEAP is Uganda's development framework and was first drafted in 1997 and revised in 2000 and 2003/04. The PEAP is a dynamic document which is revised every four years in accordance to the changing circumstances and emerging priorities within the national economy. Data and information from household surveys have provided substantial inputs to the revision processes. For instance, the 2002-2004 PEAP revision was influenced by four core challenges, namely: highly unequal growth leading to increased poverty (based on household survey data), less than expected improvement in human development indicators in the 1990s (with the exception of HIV/AIDS), and persistent insecurity which resulted into changes in the regional pattern of poverty. Based on these challenges, the PEAP 2004 has five pillars/components namely: a pillar focusing on Economic management; Produc-

tion, competitiveness and incomes; Security, conflict-resolution and disaster-management; Governance and Human development. In addition, information from household surveys has substantially contributed to the monitoring of MDGs.

3. The Status of Household Surveys Prior to the PEAP

Household surveys were only revived in 1988 after a period of almost 20 years. In 1990, poverty alleviation programmes were started under the Programme for the Alleviation of Poverty and the Social Cost of Adjustment (PAPSCA). Under the Social Dimensions of Adjustment component of PAPSCA, data on a wide range of social and economic statistics was collected with the Integrated Household Survey (IHS) providing the first baseline data on many of the issues. Thereafter, a proposal to develop a clear programme of poverty alleviation was developed and endorsed by the Ministry of Finance, Planning and Economic Development. Because of the many and varied data gaps, the surveys conducted after 1992 were attempts to fill gaps. The modules covered in these household surveys were based on discussions with some of the key users of data and these discussions culminated into the first PEAP in 1997. The uncertainty about availability of funds to undertake regular surveys was the main challenge faced during the post PAPSCA era. An overview of the situation before and after PEAP is summarized in the Table 1 below.

Table 1: Pre and post PEAP data appreciation

Period before PEAP	After PEAP (1997+)
Limited use and availability of household survey data	Increased demand for and use of household survey data and information
Initiated household surveys under Programme for alleviation of Poverty and the Social Cost of Adjustment (PAPSCA)	The PEAP is the building block for the household survey programme
Limited users and producer consultations	Extensive consultation with government, development partners and other stakeholders
Data generated almost exclusively for government and a few development partners	Users of data are many and varied
Virtually no collaboration with research institutions	There is extensive collaboration with research institutions and increased production of policy relevant papers based on household survey data
Very limited analysis of household survey data	Increased access to household survey Data

Access to data limited	Data collected to monitor PEAP and other development frameworks
The main focus was on macroeconomic variables (GDP, Consumer Price Indices etc)	The focus is on both macro and micro indicators
Data generated to fill gaps due to over 20 years without household survey data	Increased level of analysis and appreciation of the quality of data

4. Stakeholders in the Poverty Monitoring Process in Uganda

During the 1990s, poverty reduction became a priority in many countries mainly supported by development partners. The outcome of the currently widespread notion of poverty reduction was the emergence of Poverty Reduction Strategy Papers (PRSP) or the PEAP in case of Uganda, and the need to monitor the performance of poverty related interventions. The PEAP outlines the monitoring and evaluation mechanisms through which the outcomes would be updated. Monitoring and Evaluation is important in keeping on track development programmes and provides the relevant information to decision makers for making informed decisions. It also keeps other stakeholders, namely the legislature, the public, civil society organizations and development partners, informed about the progress being made in implementing the PEAP and other development frameworks.

The monitoring system in Uganda is multi faceted having many institutions and agencies running independent monitoring and evaluation systems. This has resulted into inconsistent reporting, duplication of efforts and wastage of resources. These sometimes uncoordinated systems have undermined the overall goal of informing the poverty eradication programme. An integrated programme has been developed under the PEAP to minimize these constraints and harmonise the reporting mechanism. Poverty monitoring is a responsibility of many stakeholders ranging from UBOS to Civil Society Organizations. Table 2 presents a summary of the stakeholders and their respective roles in poverty monitoring in Uganda.

Table 2: Institutional framework for poverty monitoring in Uganda[1]

Institution	**Responsibility**
Office of the Prime Minister	Collate policy related information and influence national political, socio-economic decisions.
Ministry of Finance, Planning and Economic Development: • Poverty Monitoring and Analysis Unit. • Budget Department • Macroeconomic Dept.	Coordinate poverty data collection particularly the Powerty Participatory Assesments. Analyze the whole range of data. Publish poverty reports that present a holistic view of poverty in Uganda and the implementation of the PEAP. Disseminate findings to Government and civil society, and service budget working groups. Commission poverty research and evaluations. Monitor Public Expenditure Monitor and project macroeconomic indicators.
Uganda Bureau of Statistics	Conduct censuses and surveys (including National Integrated Household Surveys, National Service Delivery Surveys, and Demographic and Health Surveys) and provide basic analysis of these data. Construct National Accounts
Ministry of Public Service	Provide information on public service performance (Results Oriented Management-ROM).
Sectoral ministries, particularly the Management Information Systems in Planning Units.	Design indicators and collect administrative data on service delivery efforts and their immediate outcomes; analyze these data in conjunction with other data sources, and identify policy responses.
Uganda AIDS Commission	Coordinate data and policy response on AIDS.
Police, Judiciary, Prisons, other JLOS Institutions	Compile administrative data on crime and the activities of the police and judiciary.
Office for Co-ordination of Humanitarian Assistance	Co-ordinate and publish data on the number and living conditions of displaced people and refugees.
Inspector General of Government, Auditor-General, Dept. of Ethics and Integrity	Ensuring that inputs are converted into outputs in a transparent manner by monitoring the integrity of public expenditures

1: Adopted from Margaret Kakande: Poverty Monitoring in Uganda 2006: The Practices and Emerging Issues, April 2006
NGO – Non-governmental Organization, CSO – Civil Society Organization

District Authorities	Collect information on relevant outputs, inputs, and outputs at this level for their own planning purposes, and sharing with central government.
Economic Policy Research Centre	Conduct relevant economic policy research that informs the monitoring process, thereby helping in refining it further.
Academic institutions and NGOs • Uganda Human Rights Commission	Conducting research on all aspects of poverty, using official data and collecting their own, providing an independent view on poverty. Providing information on human rights issues
NGOs/CSOs	Undertake value for money evaluation activities
Development Partners	Supporting the monitoring efforts of government and CSOs; and using the findings to positively influence poverty reduction activities.

The Poverty Monitoring Strategy uses three main types of data; quantitative data collected by the Uganda Bureau of Statistics (UBOS) through censuses and surveys, those collected by sector ministries through their Management Information Systems, and those collected through the Participatory Poverty Assessment Programmes (PPAs).

5. The Role of National Household Survey Data in PEAP Monitoring

During the 1970s and the early 1980s, there was great dearth of data on households in general in Uganda. This was a period when data collection at all levels (household, establishments, etc) was in limbo. Since 1988, however, the need for evidence based decision-making emphasised by the country's planning and decision making authorities and the political leadership, culminated into the rebirth of dynamic statistical programmes at UBOS and in Uganda in general. Since then, Uganda has undertaken more household surveys than many Sub-Saharan African countries in the last 15 years. Table 5.1 provides a summary of the Household Surveys undertaken since 1988.

Table 3: Household surveys conducted by UBOS from 1988 to 2004

Survey		Uganda Demographic and Health Survey UDHS	Household Budget Survey HBS	Integrated Household Survey HIS	Uganda National Household Survey 1993-94 MS-1	Uganda National Household Survey 1994-95 MS-2	Uganda Demographic and Health Survey UDHS (+MICS)	Uganda National Household Survey 1995-96 MS-3	Uganda National Household Survey 1997 MS-4	Uganda National Household Survey 1999-00 UNHS I	Uganda Demographic and Health Survey UDHS (+MICS)	Uganda National Household Survey 2002-03 UNHS II	National Service Delivery Survey (NSDS) NSDS
Duration	From	Sep-88	Apr-89	Feb-92	Aug-93	Jul-94	Mar-95	Sep-95	Mar-97	Aug-99	Sep-00	May-02	Feb-04
	To	Feb-89	Mar-90	Mar-93	Feb-94	Mar-95	Aug-95	Jun-96	Nov-97	Jul-00	Mar-01	Apr-03	Mar-04
Sample Size (usable observations)		4,370 women 15-54	4,595	9,925	5,040	4,925	7,070 women and 1,996 men 15-54	5,515	6,655	10,773	7,246 women and 1,962 men 15-54	9,711	18,000
Districts Ommited		9				1	1	1	4	4	4	3	
Questionnaires	Socioeconomic		¦	¦	¦	¦		¦	¦	¦		¦	
	Labour force								¦				
	HH Enterprise			¦	¦							¦	
	Crops					¦		¦		¦			
	Woman	¦					¦				¦		
	Man						¦				¦		
	Establishment			¦	¦		¦				¦		¦
	Community			¦	¦			¦	¦	¦		¦	
Date of Release of publications		Oct-89	Feb-91	Dec-93 Feb-94 Sep-94	Sep-95 Jun-96	Jun-97 Dec-97	Aug-96	Dec-97 Jul-98	Mar-99 May-00	Jan-01 Jan-02	Dec-01	Jan-03 Nov-03	
Months from completion to publications		8	11	9 11 18	19 28	27 33	12	18 25	16 30	6 18	9	-3 7	

From the standpoint of their ability to address the policy issues emerging from the PEAP and other development frameworks, the variety of surveys can be grouped into two major categories namely Consumption Surveys (developed under different names such as HBS, IHS, MS, and, more recently, UNHS,) and Demographic and Health Surveys (DHS and Sero-prevalence Surveys). Below is the summary description of the objectives of each of the consumption surveys;

5.1 Consumption surveys (household surveys with a consumption module)

The Household Budget Survey (HBS), conducted from April 1989 to March 1990 was the first detailed household expenditure-based survey since the 1960s. The main objective of the survey was to provide basic data needed to revise the Consumer Price Index and data for improving estimates of the household final consumption expenditure component of the Gross Domestic Product (GDP) using the expenditure approach. As outlined in Table 3, the first large scale Integrated Household Survey (IHS) conducted in 1992 was a multi-subject inquiry with the main objective of providing a complete dataset needed to understand the mechanisms and effects of the structural adjustment programs, to fill the socio-economic data gaps and to provide base-line data relating to key economic indicators used in development planning.

The IHS was a comprehensive survey covering a 12 months period and provided the first ever baseline poverty estimate for the country.

The subsequent four rounds of Monitoring Surveys (MS1-MS4) were conducted with the main objective of providing time series data to measure economic growth and social development. Although monitoring surveys covered the whole country, they were much smaller in sample size (ranging between 4900-6500 households) and covered a period of between 7-10 months. Both the IHS and the First Monitoring Survey (FMS) yielded useful data on small and household based non-agricultural enterprise activities critical in updating the National Accounts. The Second and Third Monitoring Surveys (TMS) in addition to the socio-economic module included an agricultural survey as the core-subject module while the Fourth Monitoring Survey (National Household Survey 1997) covered a labour force component as the core module as well as socio-demographic data. The subsequent national household expenditure surveys conducted in 1999/2000 and 2002/2003, covered agriculture and labour force modules respectively. The expenditure component of the socio-economic module was maintained in all surveys except the demographic surveys to support the derivation of poverty estimates.

The Bureau is currently implementing the 2005/06 Uganda National Household Survey. The survey focuses on socio-economic characteristics, agriculture and related characteristics among other modules.

5.2 Demographic and Health Survey

The Demographic and Health Survey (DHS) was conducted in 1989, 1995 and 2000/2001 respectively and collected data on fertility, family planning, maternal and child health and other demographic characteristics. In all, DHS information on HIV/AIDS was limited to knowledge about awareness. Although there is widespread knowledge about ways of transmission and protection, HIV/AIDS continues to pose a serious challenge to the health delivery system and poverty eradication efforts in Uganda. As a mechanism of monitoring the incidence and prevalence of the pandemic, a national-wide HIV/AIDS Sero Survey was undertaken by the Ministry of Health in collaboration with UBOS to, among other uses, provide prevalence rates of HIV/AIDS in Uganda. The survey, unlike in many other countries in eastern and southern Africa, was undertaken independent of the DHS.

Evidence clearly shows that UBOS has overtime, evolved into an established organization with proven capability and expertise to conduct a wide variety of nation-wide

household surveys. The Bureau has succeeded in organising and conducting regular surveys and consistently reduced the time gaps between the end of field operation and the release of publications. Most importantly, the organization has developed and maintained expertise and capabilities in the design and implementation of household surveys, which is obviously an essential asset for future household survey system.

5.3 Survey organization and participation of stakeholders in the development of household survey instruments

The Bureau does not have regional or district based offices which makes recruitment of field based interviewers during data collection very difficult. A mobile team approach has been employed for field organization and management of most of the household surveys. The teams constitute centrally recruited and trained field staff responsible for data collection. The field team recruitment criteria is based on the education level, local languages spoken and the clusters to be covered per region. The team comprises of five field staff headed by a supervisor and composed of four interviewers, who travel together from one survey cluster to the next throughout the data collection period.

The alternative to the field team approach is the field based interviewers in which information is collected within or near their area of residence. This model has not been applied in Uganda because the Bureau does not have fully operational district or regional offices. This makes the recruitment, supervision and monitoring of field activities equally costly. Plans are underway to operationalise this model once the regional statistical offices are equipped and become functional. Although experience shows that centrally recruited teams are expensive and unsustainable in the long run due to the limited resources allocated to data production by national governments, in the absence of operational regional offices, the costs of decentralized field activities may exceed those of the centrally recruited teams. A study is planned to establish the costs and benefits of each approach as well as the actual savings realized using the decentralized model of data collection.

6. The Design of the National Household Surveys

6.1 Consultations with stakeholders

Household survey data has increasingly become more of a public good than previ-

ously imagined. Today, users demand for data rather than indicators. For this reason, key stakeholders are involved at the planning stage to ensure the production of demand-driven data from these surveys. The wide involvement of different stakeholders results into:

a) optimal utilization of household survey data for indepth research by analytical and research institutions in Uganda;
b) generation of new ideas and information to enrich and support evidence-based policy formulation and decision-making process;
c) promotion and enhanced multiple ownership of the household surveys data thereby increasing data utilization in the development planning processes; and
d) greater appreciation of the constraints associated with too many demands on the questionnaire content to meet the diverse needs of stakeholders.

Wide consultations have been made involving all the key stakeholders in the production and use of data during the post PEAP period. The PEAP was used as the building block for the survey programme and indicators identified as critical in informing the PEAP were used as benchmarks.

6.2 Sample design

Sample selection and sample size

A similar sample design has been adopted for all household surveys starting with the Integrated Household Survey (IHS) 1992/93. It is typically a stratified two-stage sampling design except in some districts where the sample was selected in three stages due to lack of Enumeration Area (EA) frames. At the first stage, Enumeration Areas (EAs) are drawn with a Probability Proportional to Size (PPS), while at the second stage; households which constitute the Ultimate Sampling Units are drawn with Simple Random Sampling (SRS). Stratification is done in such a way that enables disaggregation of data to rural/urban and regional levels. The demand for district and sub-district level indicators has resulted in increases in sample sizes to derive some district level estimates. The sampling frames used for all the surveys were based on the Population and Housing Census of 1991 and 2002. Household surveys conducted prior to 1991 were based on available administrative information.

The sampling frame for selection of First Stage Units (FSUs) was the list of EAs with the number of households based on cartographic work of the 1991 and 2002 Population

and Housing Censuses. For selection of the second stage units, which are the households, a listing exercise was undertaken in selected EAs in all the household surveys.

In all cases except the UNHS 2005/06, each district was considered a stratum and divided into rural and urban sub-strata. The urban area was further sub-divided into district town and other urban areas. This multifaceted stratification enabled a better spread and representation of the sample, thereby increasing the efficiency of the estimates. As shown in the Table 3, the sample sizes have varied from about 5000 to 18000 households. The allocated sample was selected with probability proportional to number of households.

Level of disaggregation of data from household surveys

Uganda adopted a decentralized planning policy as a way of improving governance and service delivery at local government level. The challenge remains for the national statistical system to provide data and indicators at the various levels within a limited resource envelope. Alternative options have been developed to try and fill the data gaps for example, the poverty mapping initiative of the World Bank which relates household based survey data with the population and housing census information. This option is an alternative way of addressing the demand for small area statistics which is increasingly becoming an important planning input.

Panel data: the quest for monitoring changes over time

Demand for Panel data has become a common request by major household survey data users. It is argued that panel data of living standards addresses a number of critical requirements of the PEAP and MDGs including; appreciation of the functioning of labor markets, child mortality and labor, determinants of and returns on education, among other issues. Nationally representative panel data on living standards can also inform policy makers on strategies for addressing equity issues between households and across geographical areas. For these reasons, panel data in addition to being an important tool for program evaluation, because of its focus on poverty dynamics, substantially complement the ongoing poverty incidence data collection especially from UNHS series. Panel data would also be an important complement to the cross-sectional data of the household expenditure surveys as the latter is most suited to provide information on poverty incidence and other PEAP monitoring indicators. The focus on dynamics that is inherent in panel data would illuminate better the reasons for change in the various indicators (including poverty).

Despite all the enumerated benefits, panel data poses a number of problems to survey statisticians. Tracing of panel households for example, after a long time is the greatest challenge faced by data collectors in Uganda. This is particularly experienced in the urban areas with high levels of mobility. Strong collaborative linkages supported by some incentive mechanisms could be introduced to ensure continued respondents' cooperation. Incentives may take different forms ranging from gifts like t-shirts, and calendars to actual cash. The latter incentive was tried out during the 1993/94 household survey but later abandoned because of the problems associated with its administration. Cash incentives also tend to raise respondents' expectations thereby affecting the outcomes of the survey.

The Bureau collected data from panel households based on the 1992/93 household survey. This involved tracing the same respondents interviewed earlier (in 1992) in 1993/1994, 1994/1995 and 1999/2000. The 1999/2000 panel data was collected on approximately 1300 households. The Government, development partners and research institutions have used the panel data in studying the poverty dynamics between 1992 and 2000. There is a strong argument for the revival of panel data collection in Uganda to enable monitoring of various changes in the economy. Tracing of panel household members focused on household heads and their spouses. It is also difficult to establish the extent of bias introduced by inclusion of a panel. However, the Bureau is about to engage major stakeholders in a discussion to find an amicable solution to panel data question.

7. Data Analysis, Dissemination and Use

7.1 Data analysis

Collaboration in data analysis

Data analysis is one of the fundamental processes in statistical operations. Failure to analyse the collected data renders them useless to policy makers and undermines the real advantages of household survey data in general. Timely dissemination of the household survey results is largely dependent upon the successful finalization of the data processing activity. UBOS considers this process as one of the key steps in any household survey. Since the 1999, UBOS produces the first volume of the household survey reports within six months after field work. It is the policy of the Bureau to ensure that data is shared with all partners and policy makers before undertaking an-

other round of household surveys. To meet the increasing and dynamic demands by policy-makers and other users of household survey data, UBOS has entered into collaborative arrangements with the Economic Policy Research Centre (EPRC) of Makerere University and the Makerere University Institute for Social Research (MISR) to further provide the detailed analytical reports that policy-makers need. This arrangement maximizes the utilisation of the limited capacity at UBOS and other NSOs in Africa in micro level data analysis.

Through such collaboration, the UBOS and EPRC supported by the World Bank, Rockefeller Foundation, International Livestock Research Institute (ILRI) and the Department for International Development (DFID) embarked on a poverty mapping project. This innovation involved linking household level information from both the population and housing censuses and household surveys to provide poverty rates at the lowest levels of administration. This activity is an attempt to meet some of the data demands for small area statistics.

Linking qualitative and quantitative data for poverty monitoring

Poverty is a multi-dimensional phenomenon which is understood and interpreted in various dimensions by different individuals and communities. The conventional approaches for deriving poverty estimates using household income/expenditure involves a detailed series of methodological steps and procedures. There are, however, other approaches to poverty measurements that use other determinants which cannot be quantified. The participatory approaches use a set of methods to assess the community and household welfare by involving the communities in identifying and defining the poor. Each of the approaches has its own merits and demerits. Uganda has embraced both approaches in the poverty monitoring process in order to maximize the benefits from the strengths of each approach.

It is generally agreed in Uganda that quantitative approaches make aggregation possible, provide results whose reliability is measurable and allow simulation of different policy options. On the other hand, qualitative methods comprehensively define poverty, provide more insight into causal processes and produce more accuracy and depth of information on certain questions. It should be noted that each approach has its weaknesses and the poverty monitoring process in Uganda acknowledges them. The interest in Uganda is to build on the strength of each approach to ensure that they play a complementary role[2]. This is achieved at two levels: First, house-

hold surveys are capable of providing the research questions to the qualitative studies and the UPPAP II experience clearly shows that this is possible and achievable. Not only were research questions developed, almost half of the 60 clusters covered under UPPAP II were panel sites for the 1999/2000 household surveys. Secondly, the qualitative approaches are useful in answering questions that may not be easily answered through a household survey. It thus answers the questions on causes of the event, coping mechanisms and social capital for health, education, agriculture, orphanhood, poverty, migration, etc. in more detail than one would expect from a household survey. The quantitative approaches are useful in providing the magnitude of the event and answers to questions like, how many are poor, where are they, etc. A section was thus included in the household questionnaire to capture such issues but also to provide a national perspective of the issues emerging out of the qualitative study.

Data reliability

Improvements in reliability

Improvements in data management have also evolved over time. Range and consistency checks have been included in the data-entry program. The processing team uses MS-ACCESS to carry out more intensive and thorough checks, while various software are used to compute statistical errors. Table 4 showed that while the national level Coefficients of Variations (CV's) for the major indicators are generally below 5 percent, they are below 10 percent at the regional level as per the survey objectives.

Standard errors and coefficient of variations

The household surveys undertaken by the UBOS and her predecessor Department of Statistics provide estimates that are precise at national, regional and at rural- urban levels. Estimates at district level are sometimes provided for fairly large districts. The sample estimates are subject to sampling errors. These arise because the sample of respondents selected is only one of the many samples that could have been selected from the same population using the same design and expected sample sizes. Sampling errors are measured in terms of standard errors and reflect the variability between all possible samples. The coefficient of variation and the standard errors both provide us with precision levels that guide data users in assessing for

themselves the usability of the estimates. Some of the selected key indicators based on the 2002/03 household survey are shown in the Table 4.

Table 4: Standard errors and coefficient of variations (UNHS 2002/3)

Variable	Estimates	Standard Error	Coefficient of Variation (C.V.)	Number of Cases
Average Household Expenditure				
Uganda	136,461	4,476	3.28	9,711
Urban	258,049	22,792	8.83	4,062
Rural	111,412	2,972	2.67	5,649
Average Household Size				
Uganda	5.1	0.046	0.90	9,711
Urban	4.1	0.07	1.94	4,062
Rural	5.3	0.05	0.95	5,649
Enrolment ratio (%)				
Uganda	85.60	0.65	0.76	11.353
Male	85.07	0.82	0.97	5,503
Female	86.09	0.76	0.89	3,984
Literacy Rate (10 Above)%				
Uganda	68.72	0.64	0.93	31,066
Male	75.83	0.64	0.85	14,767
Female	62.23	0.81	1.31	16,299
Literacy Rate (18 Above)%				
Uganda	67.6	0.70	1.0	20,637
Male	78.76	0.69	0.9	9,599
Female	57.98	0.92	1.6	11,038
Percentage of persons Ill or Injured during last 30 days				
Uganda	28.34	0.40	1.4	50,510
Male	27.11	0.47	1.8	24,500
Female	29.50	0.48	1.6	26,008
Percentage of population that usually sleeps under Mosquito Net				
Uganda	10.7	0.51	4.8	50,510
Male	10.1	0.50	4.9	24,500
Female	11.3	0.57	5.0	26,008
Houses by Roofing Type (%)				
Iron Sheets	63.3	1.1	1.8	9,711
Thatched	2.4	0.3	10.6	9,711
Houses by Wall Type (%)				
Bricks	50.7	1.2	2.4	9,711
Mud/Poles	45.8	1.2	2.6	9,711
Houses Floor Type (%)				
Cement	24.0	0.9	3.7	9,711

Earth	73.5	0.9	1.3	9,711
Concrete/Stone	2.0	0.2	10.0	9,711
Households by source of Fuel for Cooking (%)				
Electricity	0.5	0.1	26.0	9,711
Paraffin	1.6	0.2	11.1	9,711
Charcoal	18.0	0.8	4.5	9,711
Wood	78.2	0.9	1.1	9,711
Household Ownership of (%):				
Bicycle	42.7	0.9	2.1	9,711
Television	6.9	0.5	6.8	9,711
Radio	63.3	0.9	1.4	9,711

It can be seen from the table that apart from variables like concrete/stones, electricity and television sets which are not common in all communities in the country, other variables have low C.Vs of much les than 5 %. This shows that these estimates are highly precise or reliable.

7.2 Data dissemination and use

Dissemination of data and information flow to users

Household survey data have provided critical inputs to policy debate in Uganda. UBOS has a free and equal access policy to data from both household surveys and other data sources. UBOS information is released at the same time to both Government and the general public. Household survey findings have been disseminated using various channels. The main findings are disseminated at national level through workshops and the media. Of recent, the dissemination channels have widened to include Press Releases, Publications, UBOS Website, Radio and Television and CD-ROMs.

How have the datasets been used in monitoring the PEAP?

The household surveys that UBOS conducts have been the major source of data and information on poverty trends over a period of 10 years, an experience that is rare in sub Saharan Africa. In addition, trends in mortality, fertility and other indicators have also been generated through a series of demographic surveys. The Bureau has deliberately encouraged researchers to utilise the existing data ets to further inform policy dialogues and this has gone down well with researchers in the

country. The production of the first ever poverty maps for Uganda was a joint effort between UBOS, Makerere University Faculty of Economics and Management (FEMA), International Livestock Research Institute (ILRI), World Bank and the UK Department for International Development. Moreover, collaborative research involving UBOS staff and the research institutions has increased tremendously over the last four years. The consistent increase in the demand for the data is a manifestation of the data utilization phenomenon. Below are a few highlights on the data utilization of UBOS data.

Not only has the data from household surveys enhanced trend analysis, it has also informed several policy concerns including the employment and unemployment rates, the nutritional status, immunization coverage, and literacy levels and enrolments. Explanations to some of the puzzles on the datasets have been provided. Furthermore, Uganda committed itself to monitor the MDGs and this is supported by the availability of regular and reliable data. The household surveys have so far provided information to monitor progress of some of the MDG indicators (poverty, youth employment, maternal mortality, enrolment among others).

8. Concerns About Sustainability and the Long-term Household Survey Programme

8.1 Financing of household surveys in Uganda.

Nation-wide household surveys are expensive undertakings that many governments including that of Uganda are unable to finance out of their own budgets. Household surveys in Uganda have mainly been funded by the World Bank since 1988. Both the 1988 Household Budget Survey (HBS) and the 1992/93 Integrated Household Survey (HIS) cost about US dollars 900,000 and 2 million respectively. A major proportion of the resources for household surveys were initially spent on building capacity and infrastructure development in form of procurement of field vehicles and hiring of consultants. The two surveys (HBS and IHS) were managed with technical assistance from UNDP. Thereafter, the more lighter monitoring surveys covering about 5000 households were undertaken between 1993 through 1997 at a cost of approximately US dollars 500,000. All these surveys were undertaken on a regular basis but without a well documented program of surveys.

2: Weaknesses of the quantitative approach include: (i) sampling and non-sampling errors, (ii) misses what is not easily quantifiable, (iii) fails to capture intra-household issues. For qualitative methods, the weaknesses comprise of (i) difficulty to generalize and, difficulties in verifying information

However in 1998, a program of household surveys and censuses covering the period 1999 through 2006 was developed and funded by both the government of Uganda and development partners especially the World Bank. Since then, government contribution towards household surveys has been substantial. Government contribution towards the UNHS 2005/06 for example was approximately US $ 600,000. This is a substantial contribution and demonstrates government commitment towards monitoring the PEAP and MDGs. This contribution was in form of project support to the Bureau, and did not greatly affect other statistical programmes. Another programme of household surveys and censuses has been developed and once endorsed by the major producers and users of data, government would be requested to commit funding for the programme.

8.2 Concerns about sustainability

The PEAP monitoring and evaluation strategy identifies indicators that are generated from household survey data. The progress made in monitoring poverty and the effects of other interventions has been mainly based on the information from household surveys. However, Uganda, like many developing countries does not have enough resources to finance, on a sustainable basis, regular household surveys and yet, these are invaluable sources of data. The challenge is how to continue generating data amidst the constrained resource envelope. The household surveys undertaken by the Bureau so far have been funded mainly by donors. Government has also stepped up its financing of household surveys but because of the resources involved, it is unlikely that it would fully fund the proposed household survey programme. A detailed costing plan is being prepared to cover the survey programme and once finalized further discussions will be held to enlist support from all the stakeholders including development partners. Government has indicated willingness to support the programme in line with the PEAP objectives.

In addition to trying to convince government and development partners about the need for making more ressources available to continue the survey programme, the Bureau is carrying out a number of innovations to ensure sustainability of the survey programme including production of more innovative statistical products like poverty maps, production of thematic and targeted reports, speeding up data processing and analysis to achieve timelines and is now looking into possibilities of establishing a permanent field organization in form of regional statistical offices.

8.3 Future monitoring of PEAP for policy:–the long term household survey programme

UBOS has developed a ten-year household survey programme that will ensure continuous monitoring of poverty indicators as well as filling existing data gaps. The programme will continue to provide nationally representative information for monitoring the PEAP and other development needs of Uganda. This is mainly driven by the need for information at all levels of decision making. Figure 1 below outlines the proposed programme and the proposed periodicity. The periodicity of the various household surveys has been determined by the pace of the decision-making process. In Uganda, this is given by the PEAP revisions every four years and by the bi-annual Poverty Status Reports (PSRs).

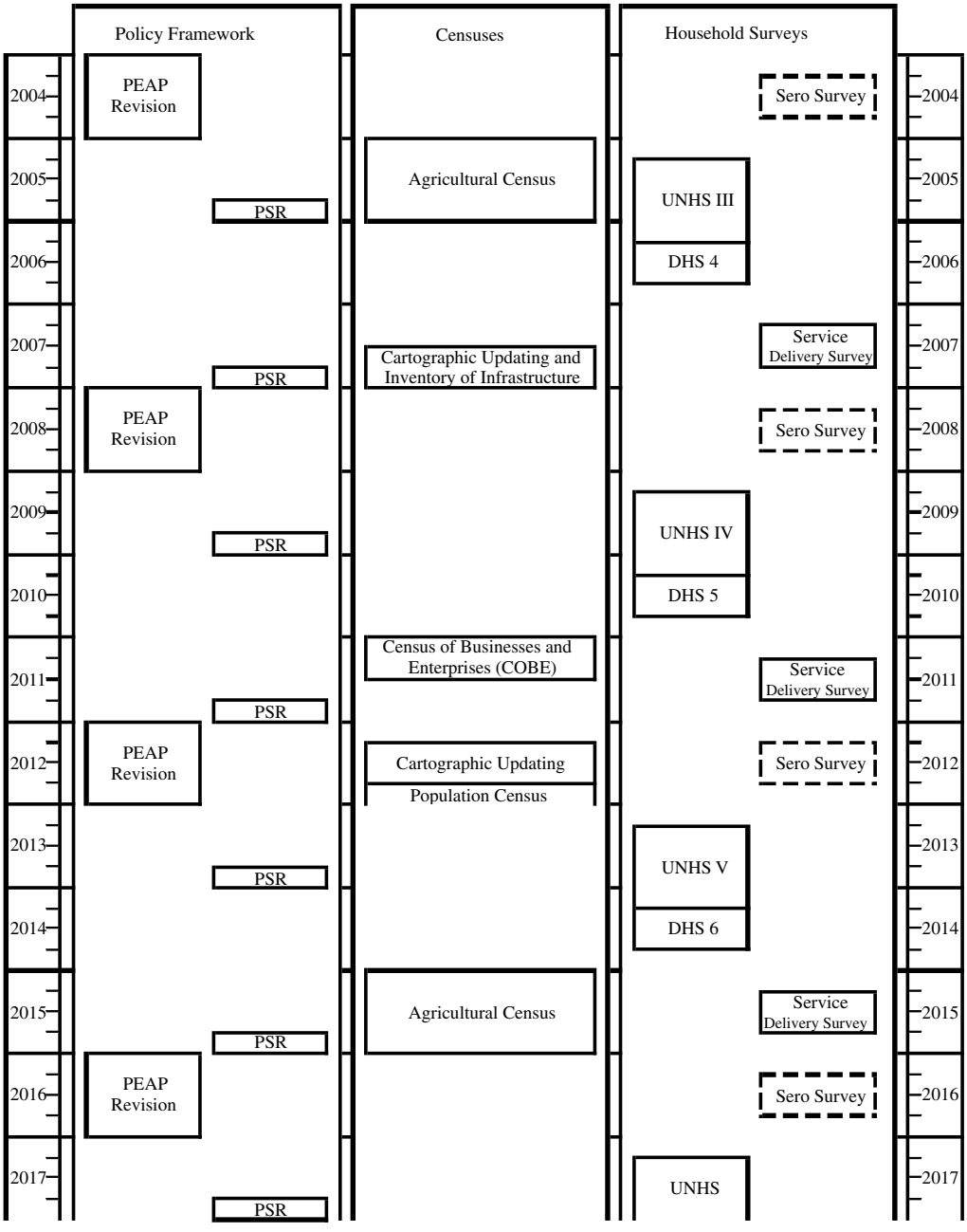
Another factor considered was the dynamism of the indicators. Some indicators – such as the rates of unemployment or immunization coverage–reflect economic conjunctures or short-term government actions that can be expected to change rapidly and may deserve to be observed frequently. In contrast, other indicators, such as access to electricity and housing condition among others, will change slowly and may be measured adequately by the decennial census or household surveys undertaken during the inter-censal period. In between these two extremes, certain aspects of household welfare – such as the amount and composition of household budgets, and derived indicators such as poverty measures – may or may not change quickly.

The level of geographic disaggregation was another aspect of the household survey system that requires careful consideration, since it is related to policy needs; it also has serious managerial and budgetary implications.

According to the proposed household survey programme, surveys with limited objectives use simpler questionnaires and pose fewer exigencies in terms of the selection, training and supervision of the field staff. Their simplicity gives them the potential of using the large samples of households needed to provide sub-national estimates, as well as national figures. On the other hand, Consumption and Demographic surveys are much more delicate operations. They use complex questionnaires that require careful selection, training and supervision of field staff. As a result, they are difficult to implement on very large samples and are unable to provide results for numerous sub-national units.

3: The districts have now increased to 80

Figure 1: Suggested program of UBOS household surveys for the period 2005-2017[1]



4: Adopted from Munoz J: Household Survey Programme for Uganda Bureau of Statistics 2004

Uganda presents a very particular challenge in this sense, because the first level of geographic disaggregation below the country is the district, and the country has 69 of them, some very small. Data users have traditionally complained about UBOS' difficulties in providing district-level figures, pointing out to the increasingly relevant role of districts in the decentralization process.

Important as this policy factor might be with regard to the demand for small area estimates, the constraint imposed by sampling theory are even harder. The fact is that the number of households needed to make a good sample for a district is almost independent of the size of the district, and it needs to be fairly large. However, very large samples, even in the unlikely case they could be financially afforded, are vulnerable to big non-sampling errors, as a result of the difficulties of selecting, training and supervising a tremendous group of interviewers.

In the figure the first column shows the time lines proposed for each programme. The second column shows the national censuses and other nation-wide data collection activities that UBOS needs to conduct in addition to household surveys. Some of these activities are essential inputs into the household survey system, but even those that may not be directly linked to household surveys still need to be taken into account because they are big efforts that may compete with household surveys for UBOS technical resources and infrastructure (vehicles, etc.) These include; Agricultural Censuses and the Population and Housing Censuses, updating the cartographic maps during the inter-censal periods and the Census of Business Establishments (Uganda Business Inquiry).

The last column shows the proposed sequence of household surveys. It consists of the cyclic replication of four elements every four years: one Consumption Survey, one Service Delivery Survey and two kinds of demographic and health surveys – a "standard DHS" (with anthropometrics and no blood testing,) and a "Sero Survey" (with blood testing and no anthropometrics.) As said before, combining the last two elements into a single, integrated instrument seems to be a technically feasible alternative that deservers further exploration.

Consumption surveys need to be fielded over a 12-month period, to take into account the seasonality of the underlying phenomena. The fact that all other surveys have traditionally been fielded by UBOS over shorter periods is reflected in table 3. The proposal is to spread the DHS surveys over the 12 months period as well to capture seasonality.

Besides seasonality, another reason for spreading the data collection period over several months or a year is to capitalize on the heavy investment in human resources usually engaged in surveys. It is cheaper and easier to select and train a small number of good interviewers and supervisors, engage then for many months, than to train a larger number and use them for a shorter period. Longer periods of data collection also permit the implementation of more reliable quality control measures. When the period is too short, problems are often detected after the survey is finished and no corrective actions are possible.

The implementation of the household survey programme, once adopted will further ensure continued provision of policy relevant information over a fairly long period of time

9. Conclusion

The poverty scourge in Uganda represents one of the greatest challenges confronting all stakeholders in the development process. The need for reliable, timely and regular information from household surveys as well as other sources becomes even increasingly more important than ever before. Policy makers need to be informed of the progress (or lack of progress) in the targeted sectors in order to address the identified constraints. It is through the provision of reliable data that appropriate decisions and interventions can be made. This is only possible if there are regular mechanisms of data production and dissemination. The long-term household survey programmes is aimed at sustaining the data production process and hence keeping the development debate and analysis alive.

References

Angus Deaton 1997: The Analysis of Household Surveys: A Microeconomic Approach to Development Policy

Juan Munoz and Kinnon Scott 2004: Household Surveys and the Millennium Development Goals

Kakande M.: A Paper on Poverty Monitoring: The Practices and Emerging Issues Lecture Notes April 2006

Ministry of Finance, Planning and Economic Development 2004: Poverty Eradication Action Plan 2004/05-2007/08

Mijumbi P. and Okidi J. Analysis of Poor and Vulnerable Groups in Uganda, Occasional Paper No 16

Ravi Kanbur, Cornell University: Qualitative and Quantitative Poverty Appraisal: The State of Play and Some Questions January, 2001

Uganda Bureau of Statistics 2003: Uganda National Household Survey Report: Socio-economic Report

Uganda Bureau of Statistics 2001: Uganda National Household Survey Report: Socio-economic Report

Uganda Bureau of Statistics 2004: The draft Uganda Household Survey System 2005-2015

Welfare and Environment in Rural Uganda: Results from a Small-Area Estimation Approach

Paul Okiira Okwi[1], Johannes G. Hoogeveen[2], Thomas Emwanu[3], Vincent Linderhof[4] and John Begumana[5]

Summary

This study combines census, survey and bio-physical data to generate spatially disaggregated poverty/biomass information for rural Uganda. It makes a methodological contribution to small area welfare estimation by exploring how the inclusion of bio-physical information improves small area welfare estimates. By combining the generated poverty estimates with national biophysical data, this study explores the correlation between poverty (welfare) and natural resource degradation at a level of geographic detail that has not been feasible previously. The precision of the resulting estimates of poverty have improved by the inclusion of bio-physical information and the poverty estimates appear to be more robust, as the standard errors show a decline in some cases by up to 40 percent. The coefficient of variation, that is, the ratio of the standard error and the point estimate decline in general as well. Overall, we conclude that the estimates of the poverty measures are more robust when biophysical information is taken into account. Part of the output from this study are maps showing poverty and biomass overlays for Uganda. These maps can be used as a planning tool and for targeting purposes.

Key words

Poverty, land use, rural, geo-referenced information, poverty maps, welfare measurement.

Résumé

Cette étude combine des données du recensement, d'enquêtes et données biophysiques pour produire des informations spatiales de pauvreté/biomasse désagrégées pour le milieu rural en Ouganda. Elle est une contribution méthodologique à l'estimation du bien-être dans les petites unités géographiques en explorant la façon dont l'inclusion des informations biophysiques améliore les estimations du bien-être dans

1: Paul Okwi, International Livestock Research Institute, ILRI, Nairobi, Kenya,
2: Johannes Hoogeveen, World Bank
3: Thomas Emwanu, Uganda Bureau of Statistics (UBOS),
4: Vincent Linderhof, Institute for Environmental Studies IVM, Free University, Amsterdam
5: John Begumana, National Forest Authority (NFA), Uganda

les petites unités géographiques. En combinant les estimations de pauvreté avec les données nationales biophysiques, cette étude examine la corrélation entre la pauvreté (bien-être) et la dégradation des ressources naturelles à un niveau de détail géographique qui n'était pas possible auparavant. On note une amélioration de la précision des estimations de pauvreté obtenues par l'inclusion des informations biophysiques. Les estimations de pauvreté semblent être plus robustes car les erreurs types ont diminué dans certains cas de 40%. Le coefficient de variation c'est à dire le rapport de l'erreur type par l'estimation ponctuelle a en général aussi diminué. En conclusion, les estimations de pauvreté, sont dans l'ensemble plus robustes quand les informations biophysiques sont prises en compte. Une partie des résultats de cette étude est constituée de cartes présentant des superpositions de pauvreté et de biomasse en Ouganda. Ces cartes peuvent être utilisées comme des outils de planification et de ciblage d'objectifs.

Mots clés

Pauvreté, utilisation des terres, information géo-référencée, cartes de pauvreté, mesure de bien-être

1. Introduction

Attaining sustainable use of bio-physical resources and sustainable growth in agriculture are important for Uganda because the economy is agriculture-based and nearly 90 percent of its 25 million people live in rural areas. Ugandan policymakers, having few resources at their disposal, must make critical decisions concerning the type of land use patterns that will prevail in future and also at the same time alleviate poverty. Unfortunately, information about poverty and land use is often incompatible. For instance, spatially disaggregated biophysical information is available but disaggregated poverty information is not. As a result, decisions are often made in an information vacuum and there is limited understanding of the dynamic processes linking poverty and land use patterns.

For both researchers and policy makers alike, various questions need to be answered. Where are the poor located? What is the state of the natural environment? What is the relation between the location of the poor and the state of the natural environment? What role do initial environmental conditions play in poverty reduction and what is the relevant level of policy intervention: regional, district, county or sub-county level? To answer these, rather basic, questions, high resolution comparable wel-

fare and biophysical data are required. To date such information has not been available and none of the research questions formulated above could be addressed.

Recent research on poverty and the environment is either based on case study approaches or on cross-country studies. The former is unrepresentative, the latter is clouded with data incomparability problems [see Atkinson and Brandolini (1999) on the problems associated with use of the Deininger and Squire data set]. Other numerous studies have only looked at the theoretical link between poverty and environmental degradation (Ambler 1999, Barbier, 2000 and Roe, 1998). These studies show that the relative strength of links between poverty and environment may be very context-specific (Chomitz, 1999, Ekbom and Bojo, 1999). By providing comparable welfare and biophysical information for many data points, the proposed data base solves these problems. However, though some studies such as Fujii (2005) have incorporated environmental data into there estimation procedure using small area estimation techniques, biophysical information has not been linked to welfare information in Uganda as yet.

For poverty, data below the regional level are often not available. However, Hentschel, Lanjouw, Lanjouw and Poggi (1998) developed an approach to examine the geographic distribution of poverty by combining sample survey information with census data. This approach is elaborated in Elbers, Lanjouw and Lanjouw – ELL, (2003). Their approach generates welfare estimates at low levels of spatial disaggregation, and additionally, it estimates standard errors for the poverty estimates. For Uganda this approach was taken up and the results show comparable welfare estimates are feasible for rural counties for both 1991 and 1999 (Okwi et al. 2003; Hoogeveen et al. 2003). These estimates only rely on census and household survey data and do not use the available biophysical information. The ELL approach leads to high precision maps and is more robust than the more conventional approaches.

This paper builds on an existing effort to generate small area welfare estimates and combines spatially disaggregated poverty and biophysical data for 1991. We use the detailed information provided in the 1992/93 Integrated Household Survey (IHS) and combine it with the 1991 Population and Housing Census and the 1990-93 biomass data to analyse the links between poverty and bio-physical information at a more disaggregated level. This study has a spatial dimension because environmental problems are inherently geographical. These estimates are based on household per capita expenditure as a measure of welfare. In the first step, we use data from the survey of 1992/93 to estimate the relationship between poverty (measured) by household expenditure and other indicators of welfare (including household eco-

nomic and demographic characteristics, district and regional dummies) and biomass data. In the second step, we use the values from the first stage regression for each stratum to get poverty estimates at lower levels including district, county and sub-county. In the third step, we develop poverty–biomass maps (overlays) to show the relationship between poverty and biophysical data.

Such a combination of information is valuable to policy makers who continue to struggle with the twin objectives of alleviating poverty in the short run and preserving the natural resource base in the long run[6]. This information is also valuable for research analysts who want to better understand the environmental-poverty nexus. From the analysis conducted in this study, we have been able to produce sets of maps (overlays) locating the poor in Uganda using an integrated database that combines census, survey and biomass information. This paper also refines the methodology of small area estimation by including biomass variables and other spatial/environmental information in the first stage regressions for poverty mapping and considers how this improves the accuracy of the poverty/biomass maps for Uganda. Indeed, the first stage regressions results (R-square) improved on average by 2 percentage points over all the rural strata after including biomass data and the point estimates (standard errors) also improved at all levels. The small area estimates are then used to explore several dimensions of poverty and natural resource relationship in rural Uganda.

Following this introduction, this paper is divided into five sections. Section 2 provides an overview of the Ugandan country setting, providing a discussion of the patterns of poverty, natural resource use and the current policy framework. Section 3 describes the data and methods that form the basis for the research reported in this paper. It also provides an overview of the three stage empirical model that underpins the analysis of the data, drawing exclusively on the existing literature on small area estimation techniques. The results are presented and discussed in section 4, while the last section concludes and discusses the broader implications of the research.

2. Poverty and Natural Resources in Uganda

For many years, the Government of Uganda has been committed to poverty reduction and environmental protection. Government strategies are summarized in the

6: Personal communication Muhumuza, F., Economic Policy Research Centre; and members of the National Biomass Study, Ministry of Water Lands and environment; and National Environment Management Authority.

Poverty Eradication Action Plan (PEAP) and implemented by the Poverty Monitoring and Analysis Unit of the Ministry of Finance, Planning and Economic Development, and the National Environment Management Authority (NEMA). With respect to poverty reduction, the Government has been quite successful although Uganda remains among the poorest countries in the world. For instance, during the 1990s poverty in Uganda almost halved from 56% in 1992 to 35% in 1999/2000. At the same time Uganda has faced a significant change in its landscape. Reliable figures are hard to come by but the Forest Department (2002) shows that forest cover in Uganda is shrinking at a rate of 55,000 ha per year. This has raised concern about the future supply of fuel wood, other forest products and environmental services. Many of these changes are believed to be linked to conversion of woodlands to agricultural land.

2.1 Poverty in Uganda

The results from different studies on poverty and inequality (Appleton 1999, 2001, Okwi et al. 2000, UPPAP 2000) in Uganda have wide ranging conclusions and are not easy to compare because either the poverty lines used were not always constant or due to other methodological differences. However, there is little correspondence of results across the studies. The studies based on survey data collected by Uganda Bureau of Statistics (UBOS) show some similarity while the other studies have some contrasting findings. Estimates of the prevalence of poverty range from 66 percent to 44 percent in 1997. Recent results from Ssewanyana and Appleton (2003) show that poverty has risen to 39 percent and inequality has remained more or less the same at Gini of 0.38 in 2002/03. All the studies clearly show that rural areas suffer from higher prevalence of poverty and inequality than do the urban areas. This situation holds even after adjusting for the cost of living differentials. This is not a surprising finding given that in many other developing countries like Kenya and Tanzania, the situation is the same. However, there may be some bias, in favour of overestimating rural relative to urban poverty in all the studies. The reason is that income and expenditure are more accurately measured in urban areas and systematic under

measurement of these variables in rural areas is possible (UBOS, 2002). Without a concerted effort to measure all income and expenditure accurately, the degree of overestimation of rural inequality and poverty cannot be accurately known. Despite this bias, the studies universally conclude that the prevalence, depth and severity of poverty are greater in rural Uganda. Table 2.1 presents poverty estimates in Uganda between 1992 and 1999.

Table 1: Poverty estimates for Uganda, 1992-1999

Domain	Poverty incidence FGT(0)			Poverty gap FGT(1)			FGT(2)		
	1992	1999	2002	1992	1999	2002	1992	1999	2002
Central rural	54.3 (2.2)	25.7 (1.4)	27.6	18.7 (1.2)	5.9 (0.4)	6.9	8.8 (0.7)	2.0 (0.2)	2.49
East rural	60.6 (2.3)	38.4 (1.6)	48.3	23.0 (1.3)	10.5 (0.6)	14.9	11.4 (0.8)	4.2 (0.3)	6.28
North rural	73.0 (2.9)	67.7 (3.8)	34.3	29.0 (2.0)	26.4 (2.9)	4.8	14.8 (1.3)	13.3 (2.0)	3.39
West rural	54.3 (2.4)	29.5 (1.9)	65.0	19.2 (1.3)	7.0 (0.6)	24.3	9.3 (0.8)	2.4 (0.2)	11.88

Notes: the 1992 estimates are derived from the Integrated Household Survey (IHS). The 1999 estimates are from the Uganda National Household Survey (UNHS) and 2002/03 from UNHS. Standard errors are in parentheses.

On average, between 1998-2002 (Table 2) Uganda registered a GDP growth rate of 6.1 percent (UBOS, 2003). Previously, the country had experienced GDP growth rates of about 7.2 percent (between 1991-1997) but the slack in GDP growth start-

ed in the fiscal year 1999/2000 due to a fall in world coffee prices, droughts, civil wars and the war in the Democratic Republic of Congo (DRC), increases in pests and diseases and a rise in world prices of oil (UBOS, 2001). These shocks affected the expansion of the productive sectors and the economy's position with the rest of the world. Infant mortality stood at 88 per 1000 live births while maternal mortality was 504 per 100,000 live births in 2001.

Table 2 Uganda: Key economic and social indicators

Indicator	Year or period	Index
Surface area ('000 of Km squared)	2002	**241.0**
Population (millions)	2002	**24.7**
Population (Annual growth rate)	1991-2002	**3.4%**
GNP per capita (US $)	2002	**320**
GDP annual growth rate	1998-2002	**6.1%**
Agriculture (percent share in GDP)	2002	**44.0%**
Agriculture (percent annual growth rate)	1998-2002	**3.7%**
Deforestation (percentage of total area)	1990-1995	**0.9%**
Labour force (millions)	1999	**11.0**
Average annual growth of labour force (percent)	1990-1999	**2.6%**
Infant mortality (per 1,000 live births)	2001	**88**
Maternal mortality ratio (per 100,000 live births)	2001	**504**
Life expectancy (number of years)	2002	
Male		**48.1**
Female		**45.7**
Total fertility rate	2001	**6.7%**
HIV/AIDS prevalence	2001	**6-7%**
Nutrition (stunting)	2001	**39%**

Source: World Bank, World Development Report (2002) and UBOS (2001, 2003).

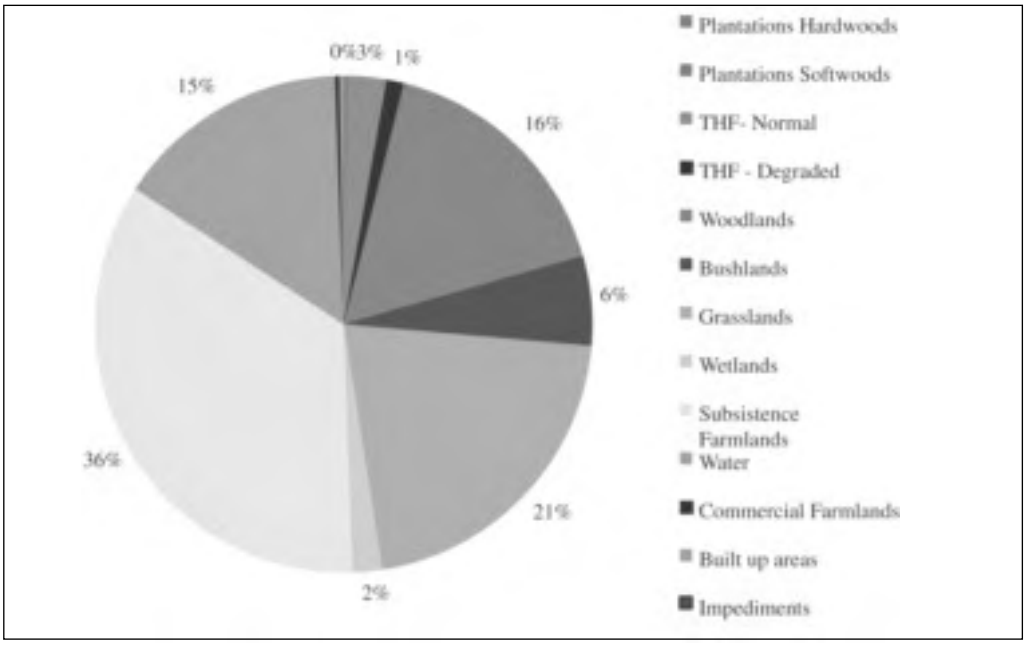
2.2 The state of the natural environment in Uganda

Uganda occupies an area of 241,038 square kilometers of which 43,941 square kilometers is open water and swamps and the rest is land area. The population of Uganda was estimated at 24.7 million in 2002, with an annual growth rate of 3.4 percent during 1991-2002 and a population density of 126 people per square kilo-

meter (UBOS, 2002). The settlement patterns in the rural areas vary, depending on a number of factors: areas with consistently good rains, good soils, and free from disease agents, have high and rising population densities. Areas with less rain, less fertile soils, and which are not free from disease agents, have low population densities. Security is another major factor, which determines settlement pattern in Uganda: for instance, the serious security problems in the northern region since the 1980's are one reason for its low population density.

Besides the other uses like pasture, farmland constitutes the biggest proportion of land use (35%) in Uganda, see Figure 1. The average landholding size in Uganda ranges from 0.4 to 3 hectares per typical household of seven persons. This landholding size has been declining over the years due to population pressure (UBOS, 2002). The climate of Uganda is more of the "equatorial" type. The area has two wet seasons, with intervening short dry seasons of one to three months. The vegetation is typically savannah, though there are some forests on the mountain ranges, and riparian vegetation in river valleys. There is a wide range of savannah woodland. This savannah is usually interspersed by perennial grasses (Department of Forestry, 2002).

Figure 1: Relative Land Cover Distribution



Source: National Biomass Study, Uganda.

Figure A1 in Appendix A shows how the land use covers is divided over Uganda.

2.3 The institutional and policy framework

Since 1987, the government of Uganda has been implementing an economic reform program supported by a large number of donors like the World Bank, International Monetary Fund (IMF) and United States Agency for International Development (USAID). This reform program aims to promote fiscal and economic management, develop the human capital through investment in education, health and other social services, reform the regulatory framework and improve incentives to the private sector. The result of this program has been realized through macroeconomic stability now seen, and the continued growth of GDP, on average about 5 percent per annum since 1987. Some studies have found that policy reforms included in the current economic and liberalization and adjustment efforts may increase the pressure on forests (Jones and O'Neill 1995, Angelsen and Kaimowitz, 1999). For instance, Kant and Redantz, (1997), show a positive correlation between external indebtedness and deforestation. However, some of these empirical studies are based on poor quality data; the analytical models make very simplistic assumptions about government objectives and policy formulation that limit their relevance.

There are two major players in the use and conservation of natural resources, the individuals/households and the government institutions. In Uganda, the government has more power in the conservation and use of natural resources. Government plays two main roles in the management of natural resources. They often own them and influence their allocation by setting the legal framework through policies that affect incentives to which other resource users respond. The natural environment is managed by the Department of Forestry in conjunction with the National Environment Management Authority (NEMA). Tropical forests are almost invariably publicly owned, and the infrastructure of water resources, as well, is often developed and owned by the public sector. It is important to note that the property rights are often unclear to the communities around them. The reason for government management of the natural resources is that the government is best placed to pursue multiple objectives - environmental protection, economic growth, regional development and support of indigenous people and cultural heritage. But the government ownership and management in the pursuit of such public objectives need to be effective if they are to overcome the incentives for private gain.

In Uganda, government stewardship of resources has shown a mixed record of successes and failures (NEMA, 2002). The failures are basically bureaucratic. The institutions are often inefficient and overstaffed with unqualified personnel. The other related problem is that under-priced natural resources put additional pressure on resource management agencies. By creating opportunities for corruption and personal gain, under-pricing makes the agencies vulnerable to influence from the politically powerful. For instance, forestry departments come under pressure to provide low-cost materials to industries and allow encroachment on gazetted areas so as to serve politically important areas and people. Meanwhile, essential tasks with little political appeal, such as maintenance and regeneration are overlooked.

3. Data and Methods

3.1 Data

The central element in this study is the availability of survey, census and biomass information. For purposes of this project, we used census data for 1991 and sample survey data from 1992 (Integrated Household Survey) to derive welfare estimates and maps. The surveys are multi-purpose household and community surveys, in the same vein as the World Bank's Living Standard Measurement Study (LSMS) surveys and were designed and implemented by the Uganda Bureau of Statistics (UBOS). The Integrated Household Survey (IHS) used a stratified sample of 10,000 households in both rural and urban areas. The survey questionnaire collected information on household and demographic characteristics, education, assets, employment, income and expenditure (Uganda Bureau of Statistics, 1992/93). The sample was designed to be nationally representative, as well as representative of the four regions divided into rural and urban strata. In this study, we only use the 4 rural strata as for these strata we can include bio-physical information in the update of welfare estimates (using a sample of households present in the IHS).

The second data source is the 1991 census which was conducted by the same institution (UBOS) and was meant to cover the entire population in both rural and urban areas. Two forms of questionnaires were used, a short and long form. The short form of the questionnaire covered mainly information on household members and education and was administered to all households in the country. The long form of the questionnaire covered housing characteristics and access to basic utilities and was administered to only 10 percent of rural areas (UBOS, 1991). The 10 percent is

representative at district level. Although the census did not collect information on income and expenditure, it provides information on a number of characteristics likely to be correlates of poverty. The census and survey data have several common household variables such as household size, composition, education, housing characteristics, access to utilities and location of residences. In this method, it is important that the survey and census are almost covering the same period. The main assumption of the method is that the parameters estimated from the survey data are almost equally applicable to the period covered by the census.

To capture the environment aspects, we use geo-referenced information from the National Biomass Study of the Ministry of Water, Lands and the Environment. The project developed its own classification system based on a combination of land cover and land uses. This information covers changes in land cover such as broad-leaved tree plantation or woodlots, coniferous plantations, tropical high forests (normal and depleted/encroached), woodlands, grasslands, wetlands, water resources and land use such as subsistence and commercial farmland, and changes in landscape among other aspects. The biomass indicators vary at the cluster level. To capture some of these variables, the proportion of the parish (lowest administrative unit in the country) under each land use type was used. For example, to capture wetlands the proportion of the parish that is covered by wetlands was used. Similarly, for subsistence farmland the proportion of the parish under subsistence farms was used. This criterion was used for all the other land use types.

In the National Biomass Study (NBS) project, the country was split into 9,000 plots with 3 sample plots at each intersection. However, due to influences of population density and agro-ecological zones on land cover and tree growth, some adjustments were made on the overall total sample plots. Topographic maps, land cover maps (1:50,000) and Global Positioning System (GPS) were used to locate the field plots on the ground. There were four categories of data capture and processing i.e. mapping (spatial and its attributes), biomass survey (filed plot measurements), monitoring of biomass and land cover change. This information details the woody biomass stock for each plot and it can be used to assess the relationship between tree cover and poverty. The data is extremely rich in bio-physical factors and also includes the distribution of infrastructure like markets, roads, schools and others. Besides, the GIS format of the data allows us to explore the possibilities of merging the data sets using GIS variables.

In addition to the land uses derived from the GIS information, we have also distance to road indicator as well. For each geographical area (districts, counties, sub-counties or even parishes), we have three different types of roads (main road, tarmac road

and track) calculated the total amount of area within the range or buffer of this particular type of road. We calculated five buffer zones ranging from 5 kilometres down to 1 kilometre. Figures A2 to A4 in Appendix A present the buffer zones for main roads, tarmac roads and tracks respectively. Note that as the buffer zones narrowed down from 5 to 1 kilometre, the percentages of total land area decrease. In particular, in North Uganda particularly, the areas are less close to any type of road than in other parts of Uganda.

3.2 Overview of the analysis

In Uganda, the availability of high-resolution datasets is a strong foundation for us to produce and use poverty-biomass maps. Although several approaches have been developed to design poverty maps, there has been less effort to develop poverty/biomass maps. The Ugandan situation is unique because two decades ago, the country was faced with deteriorating economic, social and environmental conditions. However today, the social and economic trends have been greatly reversed, but it is not clear what the implications of these changes are for the natural resource base. The approach we use to link these problems uses statistical estimation techniques (small area estimation) to overcome the typical limitations in the geographic coverage of household welfare that surveys provide and the lack of welfare indicators in the census data, and includes biomass information to assess these changes.

Our approach to the analysis of the links between poverty and biomass using maps begins with the construction of a poverty map. We adopt the approach developed by Elbers, Lanjouw & Lanjouw (2003). First, we select variables based on comparable variables found in the survey and census data sets. The variables are derived from the comparable questions in the questionnaires. This is done because the empirical modelling of household consumption is limited by the set of variables that is common in the two data sets. A test is done to compare the means for the survey and census variables and the variables that pass the significance test are considered for the regression analysis. Close examination of the data shows that several variables that appear to be the same in the two data sets were really quite different. Reasons for these differences could be attributed to the fact that probably the two exercises measured distinctively different things for these variables or that the survey was simply not representative of population for these variables.

A logical next step is to make the connection between welfare and biophysical information. However, obtaining information on biomass use for administrative units is

not straightforward, because of confidentiality, different data formats, the intricacies of geo-analysis and because environmental conditions do not follow administrative boundaries. There have been attempts to link poverty to other socio-economic factors that do not follow administrative boundaries (e.g. ILRI 2002), suggesting that combining poverty with other information (in this case on livestock) is key for a convincing integrated framework to address poverty issues for pastoralist populations.

Once the census and biophysical datasets are integrated, ELL welfare estimates can be improved (see for instance. Mistiaen et al. 2002 for Madagascar). The preliminary poverty estimates for rural Uganda control for spatial autocorrelation solely by relying on Population Sampling Unit (PSU) means calculated from the census. By controlling for biophysical characteristics of the estimation procedure, the efficiency of the derived poverty estimates may be improved, leading to more precise estimates and enhancing the level of spatial disaggregation that is attainable.

In the regression analysis, we use household survey data to estimate per capita expenditure as a function of a variety of household characteristics. This estimation takes the form:

$$\ln y_{ch} = \chi_{ch} \beta + \eta_c (Z) + \in_{ch}$$

Where y_{ch} is the log of per capita consumption expenditure of household h residing in cluster c, X_{ch} are the household characteristics that are observable in both the survey and census data sets, and β is a coefficient vector. In our household survey, the clustering is done at regional (disaggregated into rural and urban) areas. The error term is composed of two parts. On the one hand, η_c (Z) applies to all households within a given cluster (location effect), which depends on the biomass conditions Z. On the other hand, $\in_{ch}$ is household specific component of the error term (heteroscedasticity). These two error components are uncorrelated with one another and independent of the regressors. This specification of the error term allows for heteroscedasticity of the household specific error component. It also allows for the possibility of spatial autocorrelation. That is, location specific effects that are common to all households within a cluster.

To reduce the magnitude of the unexplained location specific component, we estimate a separate model to explain the cluster specific error terms. As regressors, cluster means of the household specific variables are obtained from the census at enumeration area level and merged into the survey data set. This is a common procedure in poverty mapping. It amounts to explaining spatial autocorrelation between

factors common to a household in a given PSU. To the extent that households attend the same school, make use of the same source of fuel wood, or water and have similar access to markets, this procedure is likely to go a long way in explaining spatial autocorrelation. Yet, various rather obvious determinants of spatial autocorrelation cannot be obtained from the census. Population and tree density, soil type and quality, access to infrastructure are examples of such information. By building an integrated dataset with census and biomass information, we are able to include such bio-physical information in explaining spatial autocorrelation. We estimate equation 1 taking into consideration the location and heteroscedasticity component of the disturbance term. Survey weights are included in some of the regressions depending on the Hausman test (see Deaton, 1997) results for whether the regressions should be weighted or unweighted.

Separate regressions were estimated for 1991 for each of the 4 rural strata of the survey data set. For 1999 only one model was estimated. We consider the set of variables that passed the test (zero stage) selection process and the final selection of variables are determined by a stepwise procedure.

The next step (second stage) is to predict out of sample and apply the survey equation to the census data. Since we are using household level census data, the combination produces estimates of per capita expenditure for each household. We simulate the level of consumption for each household based on Elbers *et al.* (2003).

4. Empirical Implementation

4.1 Zero Stage: Selection of Variables

The first step is known as the "zero stage". In this stage, we compare variables from the survey and census, and we select potential ones, which are later used in the regression models described in the methods above. Principally, the idea is to obtain variables from the household survey, which are comparable to those in the census. The initial step is to look at the question in both the survey and census. This should provide a clue as to whether the responses should provide similar information. However, it is not usually obvious that identical questions will yield similar responses for several reasons. For instance, the way the question was asked, the local language translation of the question, the ordering of the questions or even variations in interpretation of questions may provide major differences in the responses. To verify

that the questions yielded similar answers, we conduct an assessment to determine whether the variables are statistically similarly distributed over the households in the survey and census. This assessment is done for each of the four strata and the comparison is done at regional level (four regions focusing only on rural strata).

After a comparison of wording, coding and instructions in the enumerator manual, we constructed a more disaggregated total of 167 potentially identical variables, which sometimes involved interactions among some variables.[7] Then, using statistical criteria, we compare the stratum level means of the variables to assess the level of similarity. We do this by testing whether the survey mean for a particular variable lies within the 95 percent confidence interval around the census mean for the same variable. A third and final step is to do a comparison of the variables across the two categories of strata (rural and urban) to assess the level of uniformity in comparability. The selection of variables used in the first stage was based on criteria, which picked all continuous variables found to be comparable. For the dummy variables, we tested whether the census and survey means were identical[8].

4.2 Re-weighting

Despite being identified as potentially identical, household size did not pass the distribution comparison test. It differed consistently between the census and the survey in that small households are underrepresented in the survey. For instance, in Central rural the census mean for one-person households is 18.4 percent but the corresponding figure in the survey is 16.3 percent. As household size is crucial when deriving per capita welfare estimates, it was less of an option to drop it from the common set of variables. And fed by the suspicion that small households are underrepresented because of non-response and improper replacement (Hoogeveen, 2003) we decided to reweigh the survey.

The re-weighting strategy followed is known as post-stratification adjustment (Lessler and Kalsbeek, 1992). It ensures that the weighted relative frequency distribution among mutually exclusive and exhaustive categories in the survey corresponds precisely to the relative distribution among those same categories in the census. In total 13 different household size categories were distinguished, reflecting

7: More detailed information on the variables and the zero-stage comparison can be obtained from Okwi *et al.* (2005), which is the supplementary report of this study. The definitions of variables are listed in Chapter 2, while Chapter 3 presents the results of the zero-stage comparison.
8: For a full list of zero-stage comparisons, we refer to Chapter 3 of Okwi et al. (2005).

households of size 1-12 with category 13 reflecting households of size 13 and over, and re-weighting was done at the stratum level. A danger of re-weighting along one dimension, household size in this case, is that survey variables that were representative using the 'old' weights become non-representative once the weights have been adjusted to control for unrepresentativeness in other dimensions. On the other hand, if the adjustment corrects for a genuine sampling error, the comparability between the survey and the census should improve in all dimensions. As a check on the appropriateness of re-weighting, we compared the set of variables that were considered identical on the basis of wording, coding and enumerator instructions and how many passed the survey-census means comparison test before and after re-weighting. Re-weighting increased the number of variables that passed this test in all rural strata considerably, while improving the fit for household size related variables.

4.3 First Stage

The first stage estimation is conducted using the household survey data, census and biomass data. Since we are analysing only rural data, the household survey is stratified into four sub-regions, and we estimate four different models. At this stage, we construct more interaction terms from the selected census, survey and biomass variables, then use a stepwise regression approach in SAS to select the variables which provide the best explanatory power to the log per capita expenditure. As is the case with other similar studies, we use a significance level criterion with no ceiling on the number of variables to be selected. The significance level used for selecting variables was 5 percent.

To develop an accurate model of household consumption, we consider the model specified in equation 1. In this model, the error component is attributable to location and household specific effects. The presence of these errors makes our welfare estimates less precise. Since unexplained location effects reduce the precision of our poverty estimates, the first goal is to explain the variation in consumption due to location as far as possible with the choice and construction of explanatory variables. We attempt to reduce the magnitude of the location effect in four ways.

i. We include in our specification district dummies and their interaction terms with key household level variables (household size, level of education, age of head of household). All districts in Uganda are represented in the survey.

ii. We calculate means at the enumeration area (EA) level in the census of

household characteristics such as household size and composition, and the gender, age and average level of education of household heads. We then merge these EA means into the household survey and consider their interactions with household characteristics obtained from the survey for inclusion in the household regression specification.

iii. For the information collected from the long form questionnaire, (for 10% of the rural households and representative at the district level) on housing characteristics, use of fuel, access to water sources, etc. we calculate district means and interact these with household characteristics.

iv. Finally, we include in our specification biomass variables and their interaction terms with key household level variables. The biomass variables include information on distance to roads, proportion of land under grassland, woodland, water, farmland and forests.

So far in the household model, cluster level means and biomass data that interacted with household characteristics are included. To further select location variables we determine the common component in the household specific error terms and regress this on enumeration area and district means. We then select a limited number (5 at most) of variables that best explain the variation in the cluster fixed effects estimates. The number of explanatory variables is limited so as to avoid over-fitting. The selected location variables are included in the household regression model after which a combined model is estimated comprising of household specific and location variables. A Hausman test described in Deaton (1997) is used to determine whether to estimate our final regression models for each stratum with household weights. We re-estimate the regressions in equation 1, but after adding weights to the selected explanatory variables.

We model the idiosyncratic part of the disturbance by choosing variables from the set of potential variables selected from the census and survey, their squares and interactions. To select a subset of these variables, we use ϵ^2_{ch} as the dependent variable in the stepwise regression and choose not more than 10 variables that best explain the variation in the household specific part of the residual.

Finally, we determine the distribution of η_c and ϵ_{ch} using the cluster residuals η_c and standardised household residuals: $e^*_{ch} = \frac{e_{ch}}{\sigma_{E,ch}} - \left[\frac{1}{H}\sum_{ch}\frac{e_{ch}}{\sigma_{E,ch}}\right]$, respectively,

where h is the number of households in the survey. We use normal distributions for each of the error components. The consumption model is then re-estimated using the Generalised Least Squares (GLS) method using the variance-covariance matrix resulting from the above equation.[9]

Table 3 summarizes the results of the first-stage regression, and it shows that the adjusted R^2s of the models for 1991 vary from 0.35 to 0.46[10], (see also Tables B1 to B4 in Appendix B for examples of regressions results). According to Table 3, inclusion of biomass information helped to raise the R^2s by an average 2 percentage points compared to the models without them. The relatively low R^2s in the rural areas may be attributed to at least two reasons. First, the number of variables in the census short forms is limited to mostly household composition, education and ethnic origin[11]. Secondly, household composition and education only change slowly over time. The returns to agriculture are variables much dependent on rainfall, illness of family labourers, incidence of pests and diseases and prices. Again some of this variation may be captured, for instance the age of the head of household and proneness to disease are correlated, but much of the cross sectional variation attributable to any of these sources will remain unexplained and gets subsumed in the error term.

Table 3: Summary Statistics of First Stage Regression Models (Rural Strata)

Number of observations	IHS			
	Central	East	North	West
Number of observations used in regressions	1660	1640	1368	1637
Number of clusters[1]	163	165	144	163
Hausman test for weights	1.29	1.04	1.71	1.84
Regression weighted?	Yes	Yes	Yes	Yes
Adjusted R^2 without location means	0.27	0.32	0.39	0.31
Adjusted R^2 with location means no biomass	0.31	0.34	0.44	0.32
Adjusted R^2 with location means including biomass data	0.35	0.36	0.46	0.34

Note: In the IHS the cluster is defined by the census enumeration area. The models without location means and with location means and no biomass are derived from Okwi *et al.* (2003).

Despite not being high, the explanatory levels are comparable to those attained else-

9: For a description of different approaches to simulation see Elbers *et al.* (2001 and 2003)
10: Note that the regressions are simply association models, and therefore the parameter estimates should not be interpreted as causal effects.
11: Inclusion of all the variables from the short form and biomass data raised the R^2 but not to the urban strata levels implying we still needed to use more information such as housing characteristics to improve them.

where in Africa. For instance, in rural Madagascar the adjusted R^2 range from 0.239 to 0.460 (Mistiaen *et al.* 2002) and in Malawi it ranges from 0.248 to 0.448 (Machinjili and Benson, 2002). Considering that for Uganda, the long form of the questionnaire was available for only 10% of the rural households, the Ugandan R-squares seem to do relatively well.

4.4 The link between poverty and the environment in Uganda.

There have been attempts to link poverty to other socio-economic factors that do not follow administrative boundaries (e.g. ILRI 2002), suggesting that combining poverty with other information (in this case on livestock) is key for a convincing integrated framework to address poverty issues for pastoralist populations. For Uganda, where most households are involved in agriculture, this finding motivates our attempt to combine poverty and environmental information.

A logical next step is to make the connection between welfare and biophysical information. However, as already noted, the regression analysis presents association and not causal models. There is need, therefore, for careful interpretation of the regression results. But it is important to note that obtaining information on biomass use for administrative units is not straightforward, because of confidentiality, different data formats, the intricacies of geo-analysis and because environmental conditions do not follow administrative boundaries. We consider a number of bio-physical factors, including proximity from parish centre to nearest main, tarmac and track roads separated into 1 to 5 kilometres, proportion of the parish land under woodlots, coniferous forests, tropical high forests, degraded forests, woodlands, grasslands, papyrus (wetland), subsistence and commercial farmland, water and impediments.

The regression results presented in Tables B1 to B4 in the Appendix B suggest some spatial correlation between poverty and some bio-physical variables. The ability of these variables to improve the explanatory power of the models is interesting but different variables were selected for the different strata. A few principal variables stand out to be clear correlates of poverty. Access to roads has much explanatory association to poverty in all the four rural strata. Despite the fact that the types of roads differ between the strata, the regression results indicate a close spatial correlation with poverty. In the rural central stratum, access to main and track roads was an important variable while in north rural, access to both main and tarmac roads was important. Likewise for east rural, access to track and tarmac roads was important

and in the west rural, tarmac and track roads are important. The spatial correlation between poverty and access to roads is evident. Although our evidence is indirect, we conclude that access to various types of roads is potentially an important issue in Uganda. By implication, any policy focused on improving access to roads will yield disproportionate benefits for the poor.

Tables B1-B4 and Tables E1-E2 in Appendices summarize the available evidence of the association between poverty and other bio-physical information. Besides access to roads, the proportion of land under woodland, subsistence and commercial farms turned out to be the most important biomass variables associated with rural poverty in central rural. Meanwhile, in the east rural, proportion of land under commercial farms, woodland and the proportion of degraded forests were important spatial variables correlated with poverty. In the north, the proportion of land under water, subsistence farmland and subsistence farmland in the wetlands were the important spatial variables. The selection of water bodies and wet farmland is probably suggestive of the fact that northern region is generally dry and access to water or wetland could be important factors in explaining poverty, given that most of Uganda's rural population depends on agriculture. For west rural, the proportion of land under woodlots and subsistence farmland has spatial relations with poverty. In addition to the selected variables, biomass variables interacted with household characteristics also proved to be important in explaining the correlation between poverty and biomass. The results from the regression analysis clearly display regional variation in spatial correlation between bio-physical and poverty information. This evidence suggests that there is strong relationship between poverty and biomass variables. We conclude that access to subsistence and commercial farmland, wetland/water, woodlands, roads and grasslands are important spatial factors correlated with poverty in Uganda.

5. Results

Once the census and bio-physical data sets are integrated, ELL welfare estimates can be improved (see for instance Mistiaen *et al.* 2002 for Madagascar). The preliminary poverty estimates for rural Uganda control for spatial autocorrelation solely by relying on PSU means calculated from the census. The second stage analyses sought to use the rural models highlight the importance of bio-physical factors in poverty estimation. First, the results of the second stage analysis are used to examine the extent to which the poverty estimates from the census and bio-physical data[12]

match the sample estimates at the level which the survey is representative (region). Secondly, we ask how far we can disaggregate our census/biophysical-based poverty estimates, when we take the survey based sampling errors to indicate acceptable levels of precision. Lastly, we focus on the ultimate goal of the analysis, namely to produce disaggregated spatial profiles of poverty and biomass. Using poverty/biomass maps, we show how projecting poverty estimates and biomass information produces a quick and appealing way in which to convey a considerable amount of information on the spatial relationship between poverty and the natural environment to users. We use poverty and biomass overlays to show the spatial heterogeneity of poverty and the natural environment.

The results of the welfare indicators measured by the conventional Foster-Greer-Thorbecke measures FGT(α) are reported with α-values of 0, 1 and 2 reflecting respectively poverty incidence, poverty gap and the poverty gap squared. As benchmark the official monthly per capita poverty lines (in 1989 prices) are used, i.e. Uganda shillings 15,947 for rural Central, shillings 15,446 for rural East, shillings 15,610 for rural North and shillings 15,189 for rural West. Table 4 below summarizes the poverty inequality estimates based on the predictions of the combined biomass and census at the regional level and the survey based estimates. The detailed estimates for the district level are presented in the appendices. To reduce clutter, the poverty estimates for the county and sub-county are presented in form of maps.

At the stratum level, the results are reasonably close to those from the survey. Interestingly, most standard errors are lower than when no biomass data was included, in some cases by up to 40 percent. As shown in Table 4, the results show a consistent story with the survey and census-based estimates. Central rural emerges with the least level of poverty even when census/biomass data is used for prediction, while north rural remains the poorest of the four strata. When other measures of welfare such as the poverty gap (P-1) and the poverty gap squared (P-2) are used, the comparison among the rural strata still remains consistent with the survey rankings. The inclusion of the bio-physical data improved the poverty estimates at the stratum level and lowered the census-bio-physical based standard errors consistently. This is even when some parishes in the North and West did not have corresponding biophysical data.

12: Some observations were missing in the census/biomass data therefore the populations represented may not be exactly the same as if it was census based data alone

Table 4: Poverty measures for four rural areas from different data sources, 1992.

Stratum		Central			East			North			West		
Poverty Measure		Estimate	Standard Error	CV#	Estimate	Standard Error	CV#	Estimate	Standard Error	CV#	Estimate	Standard Error	CV#
Poverty incidence FGT(0)	Survey	54.30	2.20	0,041	60.60	2.30	0,038	74.30	2.60	0,035	54.30	2.50	0,046
	Census*	54.10	1.69	0,031	63.80	1.57	0,025	74.50	1.84	0,025	55.50	1.69	0,030
	Census/Biomass	53.42	1.25	0,023	63.40	1.48	0,023	74.80	1.07	0,014	55.40	1.37	0,025
Poverty gap FGT(1)	Survey	18.70	1.20	0,065	23.00	1.30	0,057	29.00	1.90	0,067	19.20	1.40	0,071
	Census*	17.90	0.84	0,047	23.90	0.93	0,039	30.30	1.10	0,036	20.30	1.02	0,050
	Census/Biomass	17.85	0.71	0,040	23.90	0.93	0,039	32.00	0.70	0,022	20.10	0.77	0,038
Poverty gap squared FGT(2)	Survey	8.80	0.70	0,080	11.40	0.80	0,070	14.80	1.30	0,090	9.30	0.90	0,094
	Census*	8.10	0.73	0,090	11.70	0.60	0,051	15.60	0.72	0,046	10.00	0.91	0,091
	Census/Biomass	8.02	0.44	0,055	11.70	0.60	0,051	17.05	0.59	0,035	10.04	0.48	0,048
Mean Per Capita Expenditure	Survey	18131	629	0,035	15460	486	0,031	13899	636	0,046	16256	537	0,033
	Census*	17951	564	0,031	15049	382	0,025	12884	370	0,029	16954	509	0,030
	Census/Biomass	18202	345	0,019	19629	4073	0,207	13755	365	0,027	16210	314	0,019

* The 'Census' poverty measures are derived from Okwi *et al.* (2003). The 'Census' and 'Census/Biomass' estimates are predictions based on the ELL method, while the 'survey' estimates are directly calculated form the IHS survey.

\# CV means coefficient of variation, which is defined as the ratio of the standard error over the point estimate.

The inclusion of the bio-physical information in the small-area estimation procedures has one major effect. The level of the precision of poverty measures (standard errors) can change. Table 4 presents estimates of four poverty measures at the regional level in 1992. Poverty measures from three different sources compared. The survey-based estimates are directly calculated from the IHS database. The 'Census predicted' estimates are based on the ELL method without the use of bio-physical information (see Okwi *et al*., 2003), and finally, the 'Census/Biomass predicted' estimates are from the present study. In this study we focus attention on the comparison of 'Census' and 'Census/Biomass' estimates.

The level of precision of the poverty measure estimates has changed due to the inclusion of bio-physical information. In Central, all the poverty measures slightly declined while for East, all poverty measures hardly changed. Except for the poverty incidence, the level of precision of the other poverty measures improved in the northern region. The poverty estimates for West Uganda hardly changed, while the accompanying standard errors declined again suggesting improved precision of the estimates. The graphs in Appendix F show the new poverty estimates and the standard errors of the present study at different aggregation levels in comparison with the 'old' results of Okwi *et al.* (2003).

In addition, we analyse the extent to which the inclusion of spatial features can allow our poverty estimates to be robust. There are two major ways of determining the level of disaggregation at which the error becomes too big. They both yield similar conclusions in most cases. One way to approach this is to consider the absolute level of the standard error. The other method, which is used in this study, is to calculate the coefficient of variation (CV), which is the ratio of the standard error over the point estimate for each administrative unit and compare this with the survey-based ratios.

The inclusion of biomass variables has improved the standard errors (in some cases by up to 40 percent) of our estimators at the stratum level. Finally, this section offered insights about the inclusion of bio-physical and other spatial features in poverty estimation. It demonstrated that relative improvements can be made in the estimation of welfare – with the inclusion of more explanatory spatial characteristics. That is, by controlling for bio-physical characteristics at the estimation procedure, the efficiency of the derived poverty estimates may be improved, leading to more precise estimates and enhancing the level of spatial disaggregation that is attainable. Awareness of this association, combined with well designed policies are key factors that may support poverty reduction in these areas.

Tables C1 and C2 in Appendix C present the poverty estimates at district level. These poverty estimates show some level of heterogeneity. All the standard errors fall below the stratum level survey based ones with the exception of Kalangala district in central region. The case for Kalangala district is an interesting and expected one. First, this is a small district with a total population of 14, 218 people which is significantly less than the population of any most sub-counties and even parishes in the region. For example, in Central region, the poverty estimates range from 25 percent to 63 percent at the district and 19.6 to 74 percent at the county. In Eastern, the poverty levels range from 39.5 to 82 percent at the district level. At the county level, the observed distribution is more interesting than at the district level. In the North, Arua is the least poor district (64 percent) while Kotido is the poorest with 91 percent poor. Similarly, Western region shows significant variation in poverty levels. Whereas Masindi has about 76 percent headcount ratio, Mbarara is the least poor with only 43 percent. Generally, there is wide variation in the poverty estimates in all the strata and we cannot categorically identify one region as being the poorest as there may be pockets of wealthy areas within the poorest region. The level distributions of poverty at various levels are shown in the graphs in Appendix F.

Further, to explain the link between certain bio-physical characteristics and poverty, we use overlays presented in Appendix D.[13] The overlays are simply meant to provide a visual explanation of the relationship between poverty and land-use. For example, from the overlays, we can identify the poverty hotspots and correlate them with the type of land use in the area. A clear example is that poverty is more pronounced in the Northern parts which are typically wooded and grassland areas and less pronounced in the degraded lands of all the regions. The implication of the later result is that the poor are actually using the ecological resources to improve their welfare but in the process they degrade the natural environment as well. However, a contrasting picture emerges from the grassland areas in Western and Northern regions which portray less and more poverty respectively, see also according correlation coefficients with opposite signs in the Table E1 and E2 in Appendix E. A question that emerges is why the difference? Possible explanations for the difference could be because the pastoral lands in Western Uganda have been modified by the people to produce high yielding varieties thus directly improving their welfare, while the pastoralists in the North are still held with the traditional norms of cattle rearing. The overlays generally have helped us to answer the following questions: Where are the poor? Which poor (rich) areas have similar types of land-use features? Which ar-

13: The county level estimates of the household expenditures and the head count are presented in Chapter 4 of Okwi et al. (2005).

eas provide which type/amount of ecosystem services? How do the land-use types overlap with poverty? How does the location of poverty compare to the distribution of ecosystem services? Which areas have access to better resources and what are the benefits and costs? This information may help policymakers to design effective policies to improve the situation. For detailed maps, see the poverty and biomass maps for all strata in Appendices E.

6. Conclusions and implications for policy

This study combines census, survey and bio-physical data to generate spatially disaggregated poverty/biomass information for rural Uganda. It makes a methodological contribution to small area welfare estimation by exploring the inclusion of bio-physical information. By combining the generated poverty estimates with national bio-physical data, this study explores the correlation between poverty (welfare) and natural resource degradation at a level of geographic detail that has not been feasible previously. In this welfare estimation method, association relationships are used to explain welfare rather than causal relationships. However, the resulting estimates of poverty measures have improved by the inclusion of bio-physical information. In some cases the levels of poverty measures have changed. For North Uganda, the poverty gap and poverty gap squared increased compared to the estimates without bio-physical information.

By providing comparable welfare and bio-physical information for many data points, this study solves many problems faced by many previous studies. For instance, previous studies (see Atkinson and Brandolini, 1999) on poverty and the environment were based on case studies which are unrepresentative. This study presents results of a representative sample and population. Secondly, previous studies have also been cross-sectional thus raising data incomparability problems. By using data from one country and collected by the same institution, with comparable questions in the questionnaires and within a period of time less than 2 years, data incomparability problems are solved. Thirdly, this study has provided a practical analysis of the link between welfare and the environment. Other studies have only looked at the theoretical link between poverty and environmental degradation (Ambler 1999; Barbier, 2000; Roe, 1998; Chomitz, 1999; Ekbom and Bojo, 1999). This study has shown that accounting for spatial differences in welfare is key to high precision maps and explaining poverty environment relationships.

The poverty estimates appear to be more robust, as the standard errors show a decline in some cases by up to 40 percent. Moreover, the coefficient of variation, that is,

the ratio of the standard error and the point estimate decline in general as well. Overall, we conclude that the estimates of the poverty measures are more robust when bio-physical information is taken into account. Part of the output from this study are maps showing poverty and biomass overlays for Uganda. These maps can be used as a planning tool and for targeting purposes.

In terms of policy, by implication, any policy focused on improving access to roads is directly related to the welfare of the poor. Similarly, policies focused on conservation of wetlands and forests, improvement of grasslands (mainly pasture land), and access to water could be important policy issues to consider in understanding the relationship between poverty and the environment. Given that most of Uganda's rural population depends on agriculture and the environment, and considering the spatial relationship between subsistence farming, degraded lands and poverty, the results suggest that focusing on improving production in the subsistence sector may prove important in reducing poverty and improving the biomass conditions. The results from the regression analysis clearly display regional up to county level variation in spatial correlation between bio-physical and poverty information and therefore imply region specific policy designs. Finally, in future research, with more information, the causal relationship will be analysed in more detail.

References

Ambler, J. (1999). Attacking Poverty While Improving the Environment. Towards Win-Win Policy Options. Background Technical Paper Prepared for the September 1999 Forum of Ministers Meeting, under the UNDP-EC Poverty and Environment Initiative.

Appleton, S., Emwanu, T., Kagugube, J. and Muwonge, J. (1999). Changes in Poverty in Uganda, 1992-1997. Centre for the Study of African Economies, Oxford Working Paper Series, WPS 99.

Appleton, S. (2001). Poverty in Uganda, 1999/2000: Preliminary Estimates from the Uganda National Household Survey. University of Nottingham. Mimeo.

Angelsen, A. and Kaimowtz, D. (1999). Rethinking the Causes of Deforestation: Lessons from Economic Models, The World Bank Research Observer, 14(1), 73-98.

Atkinson, A.B. and Brandolini, A. (1999). Promises and Pitfall in the Use of "Secondary" Datasets: Income Inequality in OECD Countries as Case Study, Journal of Economic Literature 39(3), 771-799.

Barbier, E. (2000). The Economic Linkages Between Rural Poverty and Land Degradation: Some Evidence from Africa, Agriculture, Ecosystems and Environment, 82, 355-370.

Besley, T. and Kanbur, R. (1993). The Principles of Targeting, In M.Lipton & J.vander Gaag, Including the Poor, Proceedings of Symposium Organized by the World Bank and the International Food policy Research Institute (IFPRI). Regional and Sectoral Studies, Washington D.C: the World Bank

Chomitz, K. (1999). Environment Poverty Connections in Tropical Deforestation, Discussion Notes prepared for the WDR workshop on Poverty and Development. Washington DC. July 6-8.

Datt, G. and Ravallion, M. (1998). Why Have Some Indian States Done Better Than Others at Reducing Rural Poverty?, Economica, 65(257), 17-38.

Deaton, A. (1997). The Analysis of Household Surveys: A Microeconometric Approach to Development Policy. Baltimore, MD: Johns Hopkins University Press.

Demombynes, G., Elbers, C., Lanjouw, J.O, Lanjouw, P., Mistiaen, J.A. and Ozler, B. (2002). Producing an Improved Geographic Profile of Uganda: Methodology and Evidence from Three Developing Countries, Discussion Paper 2002/39, WIDER, Helsinki, Finland.

Ekbom, A. and Bojo, J. (1999). Poverty and Environment. Evidence of Links and Integration in the Country Assistance Strategy Process, World Bank. Africa Region Discussion Paper no. 4 World Bank. Washington DC.

Elbers, C., Lanjouw, J.O. and Lanjouw, P. (2002). Welfare in Villages and Towns: Microlevel Estimation of Poverty and Inequality, Policy Research Working Paper, World Bank: Washington D.C.

Elbers, C., Lanjouw, J.O and Lanjouw, P. (2003). Micro-level Estimation of Poverty and Inequality, Econometrica 71(1), 355-364.

Forest Department. (1988). National Biomass Project Document. Kampala, Uganda

Forest Department. (1992). National Biomass Study Phase I Technical Report. Kampala, Uganda.

Forest Department. (1994). National Biomass Study Evaluation Mission Report. Kampala, Uganda.

Forest Department. (1995). National Biomass Review Mission Report. Kampala, Uganda.

Forest Department. (1996). National Biomass Study Phase III Project Document. Kampala, Uganda.

Forest Department. (2002). National Biomass Study Final Report. Kampala, Uganda.

Foster, J. Greer, J. and Thorbecke, E. (1984). A Class of Decomposable Poverty Measures, Econometrica, 52, 761-66.

Fujii, T. (2005). Micro-Level Estimation of Child Malnutrition Indicators and Its Application in Cambodia, World Bank Working Paper, WPS 3662.

Glewwe, P. (1990). Efficient Allocation of Transfers to the Poor. The Problem of Unobserved Household Income, Working Paper, 70, Living Standards Measurement study. Washington D.C: The World Bank.

Glewwe, P. and van der Gaag, J. (1990). Identifying the Poor in Developing Countries: Do Different Definitions Matter? World Development, 18, 6.

Government of Uganda. (1991). Uganda Population and Housing Census, Uganda Bureau of Statistics.

Hentchel J., Lanjouw, J., Lanjouw, P. and Poggi, J. (2000). Combining Census and Survey Data to Trace Spatial Dimensions of Poverty: A Case Study of Ecuador, World Bank Economic Review, 14,1, 147-65. Washington D.C: The World Bank.

Emwanu, T., Hoogeveen, J. G., and Okwi O.P.(2006). Updating Small Area Welfare Indicators in the Absence of a New Census, World Development, Forthcoming.

International Livestock Research Institute (2002). Mapping Poverty and Livestock in East Africa, ILRI publications.

Jones, D.W. and O'Neill, R.V. (1995). Development Policies, Urban Unemployment, and Deforestation: The Role of Infrastructure and Tax Policy in a Two-Sector Model, Journal of Regional Science 35, 135-53.

Kant, S. and Redantz, A. (1997). An Econometric Model of Tropical Deforestation, Journal of Forestry Economics 3, 51-86.

Machinjili, C. and Benson, T. (2002). Malawi: An Atlas of Social Statistics, National Statistics office, Malawi

Minot, N. (2000). Generating Disaggregated Poverty Maps: An Application to Vietnam. World Development, 28, 2.

Mistiaen J. A., Ozler, B., Razafimanantena, T. and Razafindravonona, J. (2002). Putting "Welfare on the Map in Madagascar". The World Bank: African Region Working Paper Series, 34.

Moller, L. (2002). A Practical Guide to Developing Good Poverty Indicators. Based on Uganda's experience, Mimeo.

NEMA (2002). State of the Environment Report, National Environment Management Authority, Uganda.

Okwi, P.O., and Kaija, D. (2000). The Distribution of Welfare in Uganda. Eastern Africa Social Science Research Review, XVI , 2.

Okwi, P.O., Emwanu, T and Hoogeveen, J.G. (2003). Poverty and Inequality in Uganda: Evidence from Small Area Estimation Techniques. Unpublished Okwi, P.O., Hoogeveen, J.G., Emwanu, T. Linderhof, V.G.M. and Begumana, J. (2005). Welfare and the Environment in Rural Uganda: Facts and Figures. PREM report 2005. Institute for Environmental Studies.

Ravallion, M. and Wodon, Q. (1997). Poor areas, or only poor people? Policy Research Working Paper, 1798. Washington D.C: the World Bank.

Roe, E. (1998). Taking Complexity Seriously. Policy Analysis, Triangulation and Sustainable Development, Kluwer Academic Publishers: Boston USA.

Uganda Bureau of Statistics. (2001). Statistical Abstract. Government of Uganda.

Uganda Bureau of Statistics. (2002). Statistical Abstract. Government of Uganda

Uganda Bureau of Statistics. (2003). Statistical Abstract. Government of Uganda

Wodon, Q. (1997). Targeting the poor using ROC curves. World Development, 25, 12.

World Bank (2002), World Development Report. New York: Oxford University Press.

Appendix A: Bio-physical information for Uganda

This appendix summarizes the bio-physical information for Uganda in 1991/1992. We have two types of biomass indicators. Firstly, we have a land use indicator, i.e. total area per land use type divided by the total area. Secondly, we have distance to road indicators, i.e. the total area within a certain distance of a particular road type divided by the total area.

In Figure A1, we use classes based on natural groupings inherent in the data. The break points are identified by picking the class breaks that best group similar values and maximize the differences between classes. The features are divided into classes whose boundaries are set where there are relatively big jumps in the data values.

Figure A1: Land use classifications: Proportion of county area under different land use types

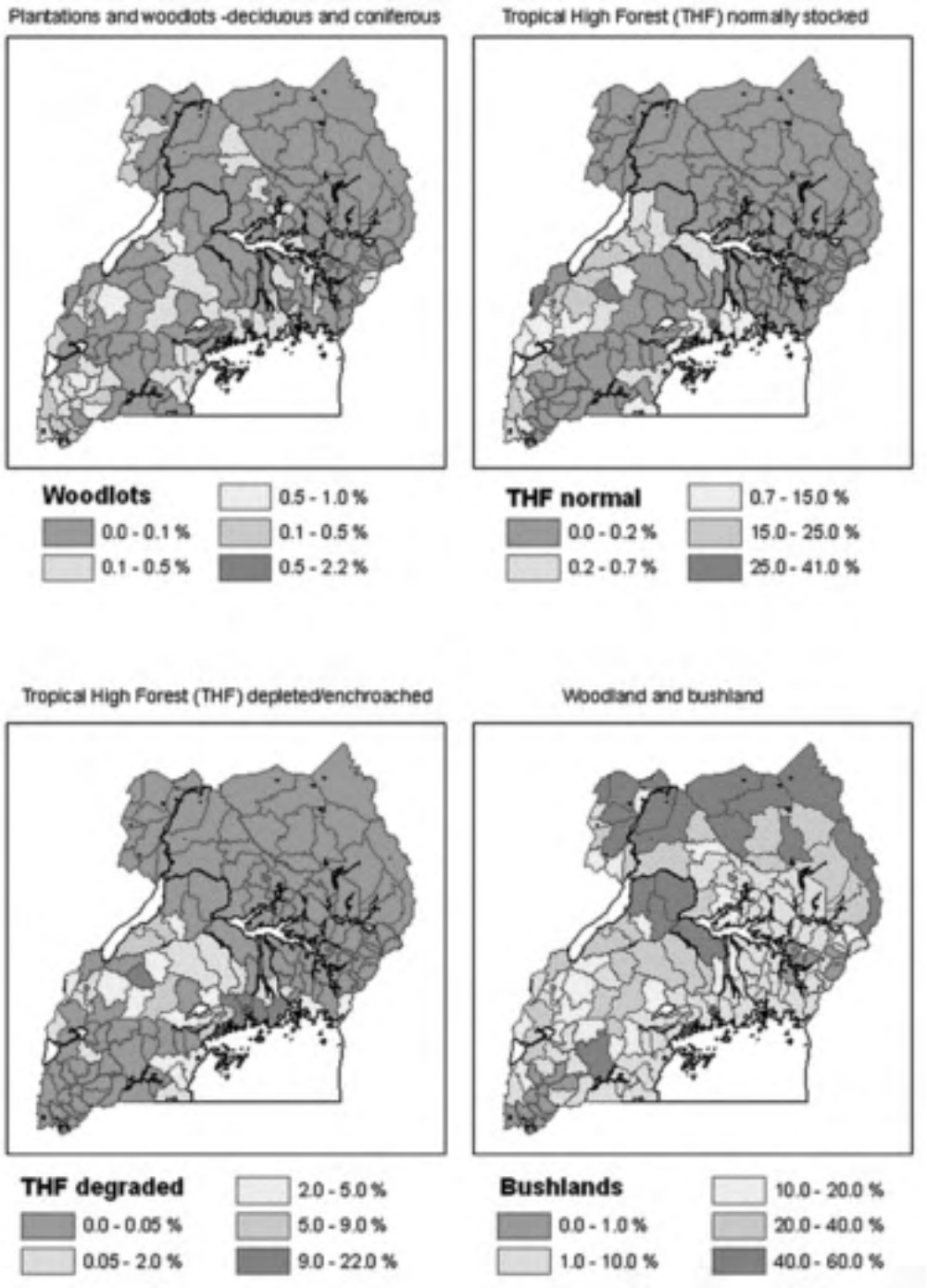


Figure A1: Land use classifications: Proportion of county area under different land use types (continued)

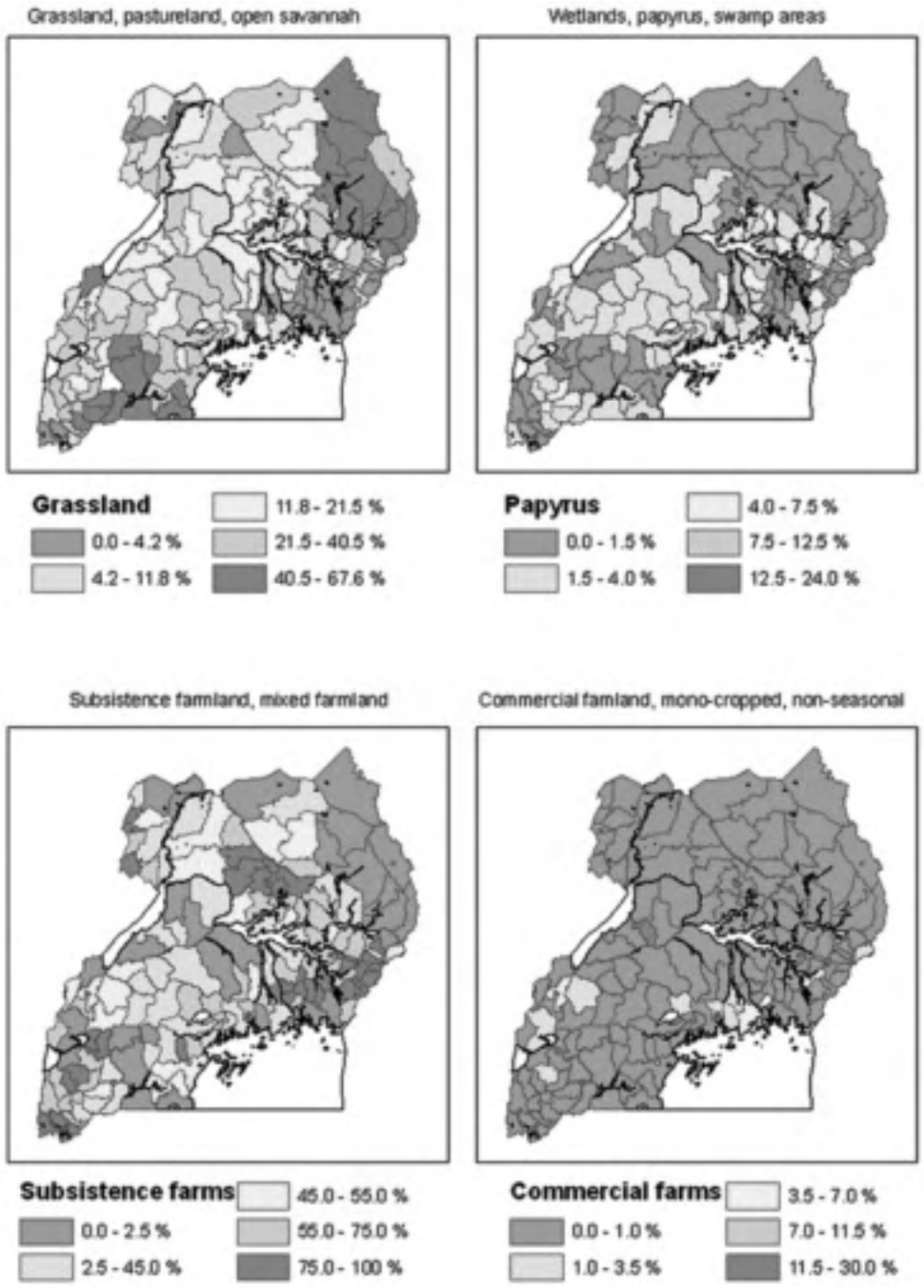


Figure A2: Main road buffers: Proportion of county area within a distance from 5 down to 1 kilometre to main roads

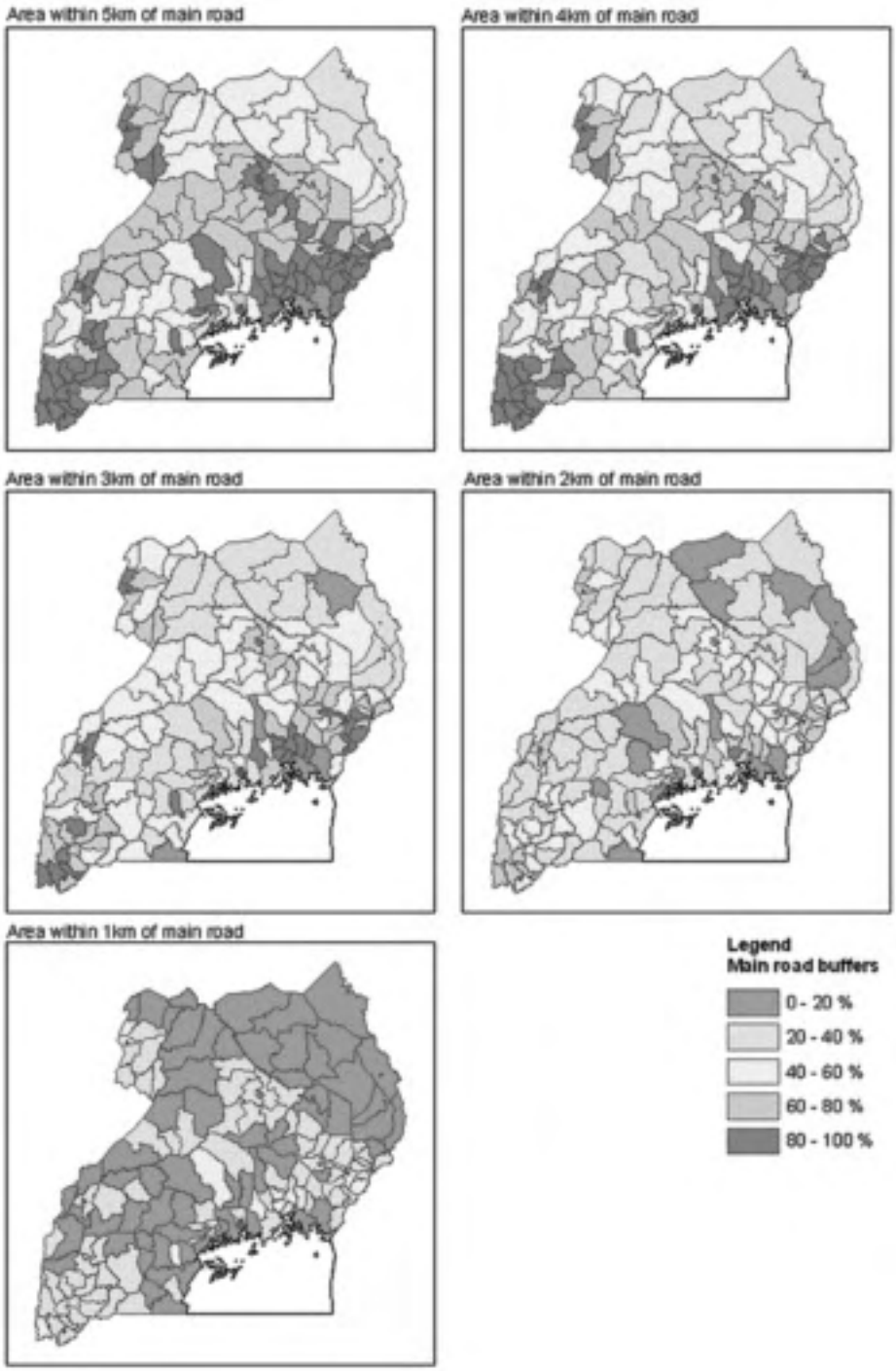


Figure A3: Tarmac road buffers: Proportion of county area within a distance from 5 down to 1 kilometre to tarmac roads

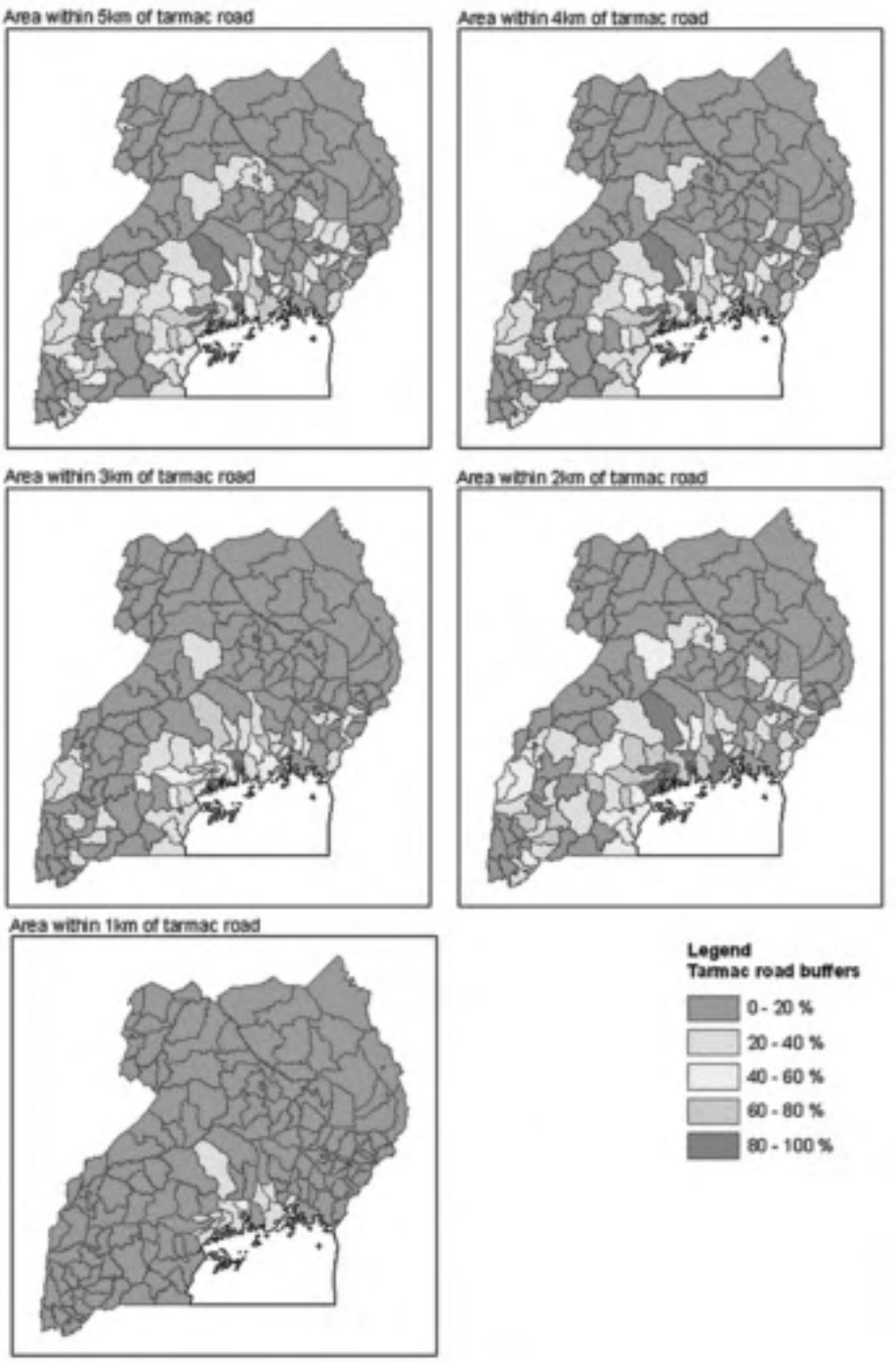


Figure A4: Track road buffers: Proportion of county area within a distance from 5 down to 1 kilometre to tracks

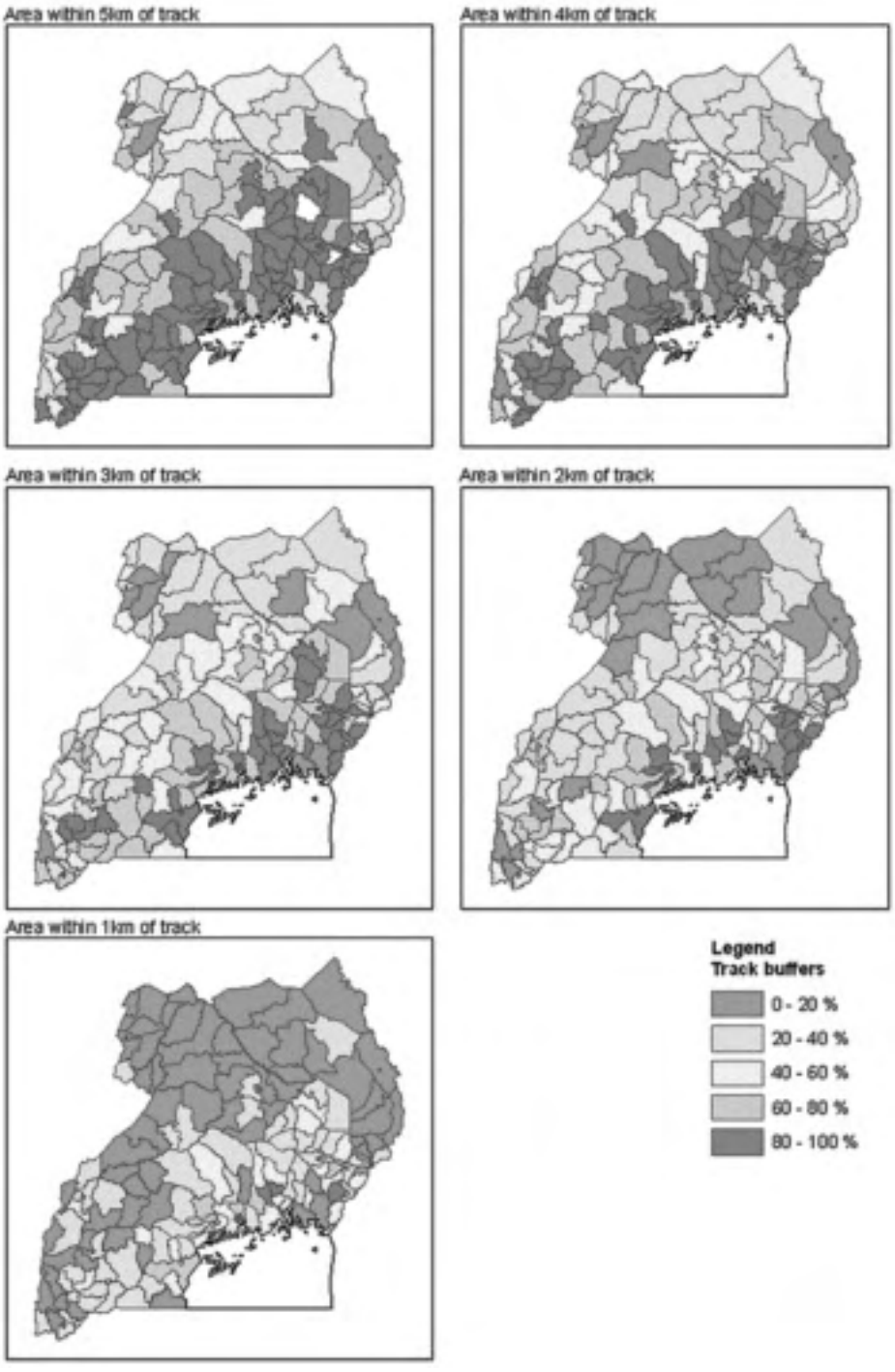
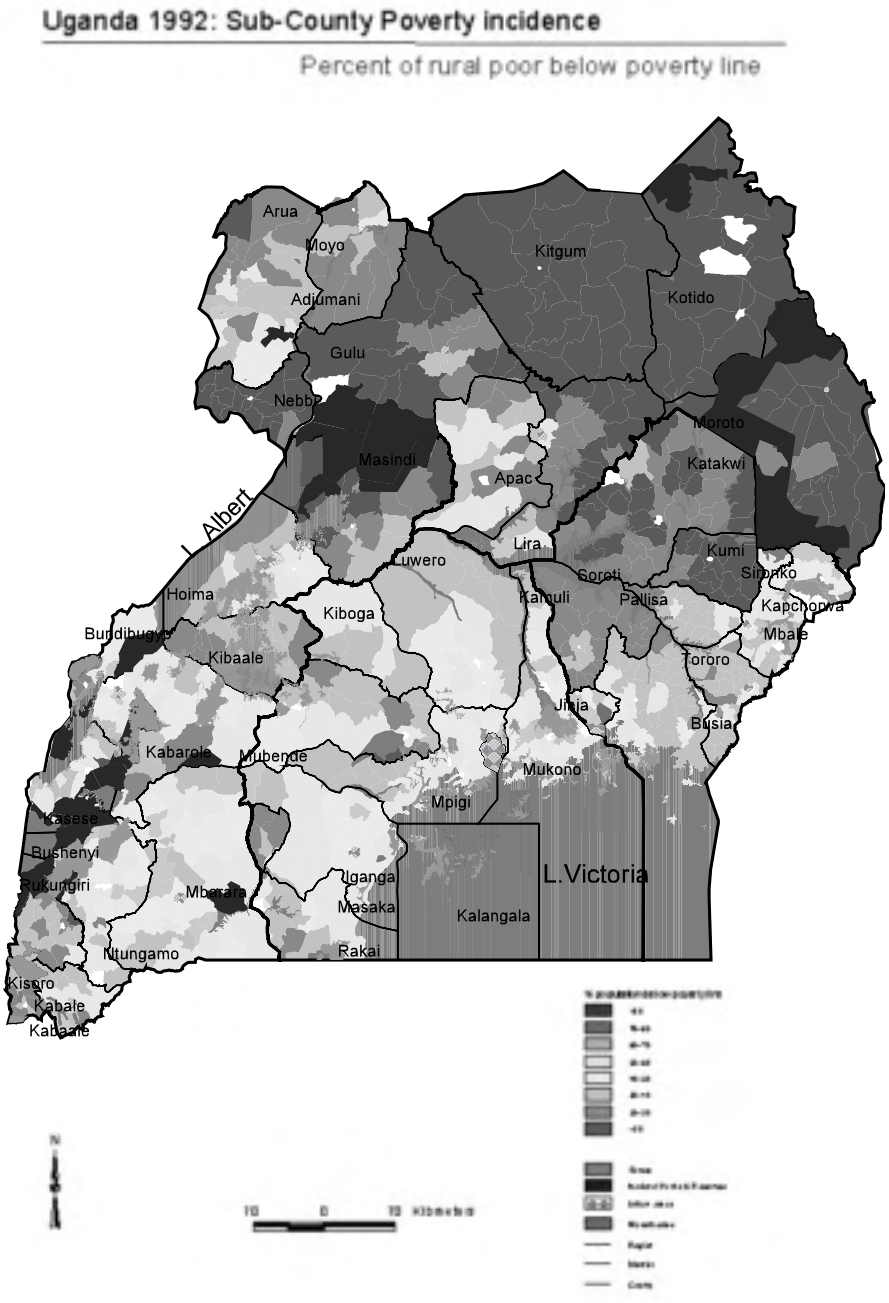


Figure A5: Sub-county poverty rates, 1992



Appendix B: First stage regressions

Table B1. First stage regression results for Central region

Dependent Variable: log of per capita consumption expenditure		
Number of observations:-	1660	
Number of Clusters:	163	
Adjusted R2:	0.35	
Variable	Parameter estimate	Standard error
Intercept	10.326	0.138
Number of females aged 6-14	0.037	0.017
Household size squared	0.001	0.000
Logarithm of household size	-0.382	0.029
Proportion of males with secondary school	0.872	0.150
Proportion of males without education	-0.153	0.046
Proportion of males with education at A 'level	0.426	0.136
Age of household head squared	0.000	0.000
Mean years of education head squared	-0.005	0.001
Number of females aged 45 or older	-0.056	0.025
Buffer zone within 1km of main road	0.341	0.078
Buffer zone within 2km of track road	-0.402	0.116
Buffer zone within 4km of track road	-0.304	0.052
Proportion of woodland (parish)	0.380	0.144
Logarithm of age of household head*Alur tribe	0.929	0.267
Logarithm of age of household head *Toro tribe	2.670	0.423
Logarithm of age of household head *Lugbara tribe	0.422	0.183
Logarithm of age of household head * Males aged 30 or older	-0.213	0.036
Logarithm of age of household head * Males aged 30 or younger	0.081	0.012
Logarithm of age of household head *Kitchen shared	0.703	0.198
Max. numbers of years of education*Ganda tribe	0.022	0.006
Log. of age of household head *Prop. of females aged 0-5 squared	-2.399	0.626
Logarithm of age of household head * Mubende district	-0.062	0.013
Log of age of household head * Prop. of subsistent farming (parish)	0.083	0.020
Log of age of household head * Prop. of commercial farming (parish)	0.183	0.058
Log of age of household head * Prop. of water (parish)	0.057	0.026
Mean number of years of education of adults * Buffer within 5km of tarmac road	-0.026	0.009
Mean number of years of education of adults * Prop. of commercial farming	-0.346	0.103
Proportion of males with A'level education*Kiboga district	-0.300	0.126
Number of males with education at level P5-P7*Prop. of grassland	0.188	0.047
Male hh. head separated or divorced * Number of males aged 30 or younger	-3.089	1.353
Hh. head with education at P5-P7 level * Prop. of town (parish)	2.999	0.766
Hh. head with education at P5-P7 level * Prop. of degraded THF	1.004	0.249
Number of males aged 30 or younger * Mpigi district	-0.050	0.014
Japadhola tribe	-2.278	0.556
Mugwere tribe	5.369	1.371

Table B2. First stage regression results for the Eastern region

Dependent Variable: log of per capita consumption expenditure		
Number of observations:	1640	
Number of Clusters:	165	
Adjusted R^2	0.36	
Variable	Parameter estimate	Standard error
Intercept	9.379	0.142
Household size = 10	-0.152	0.073
Logarithm of adult equivalent size	-0.444	0.024
Prop. of males with no secondary education squared	0.437	0.139
Number of males aged 15-29 years	-0.061	0.018
Age of household head squared	0.000	0.000
Prop. of persons with education under A' level	0.457	0.115
Proportion of males with education years 1 to 4 Squared	0.241	0.061
Buffer zone within 1km tarmac road	-0.255	0.105
Prop. of degraded tropical high forest (parish)	6.927	1.197
Prop. of commercial farm land (parish)	4.100	0.706
Prop. of males with secondary education * Teso tribe	0.229	0.042
Number of males with education between P5-P7* Ganda tribe	2.535	0.521
Maximum years of education * Rwanda tribe	-1.886	0.650
Heads education between P5-P7*Ganda tribe	-2.824	1.261
Log of age of household head *Kamuli district	-0.069	0.016
Log of age of household head * Kapchorwa district	0.093	0.021
Log of age of household head * Kumi district	-0.070	0.015
Log of age of household head *Soroti district	-0.070	0.014
Maximum years of education*pit latrine	-0.070	0.005
Maximum years of education *Kamuli district	-0.070	0.008
Number of males education between P5-P7*Iganga district	0.062	0.021
Number of males education between P5-P7*buffer within 1km track	0.049	0.021
Male hh. head separated, divorced*Kamuli district	-0.348	0.131
Number of males aged 30 or younger* Prop. Of woodlot	1.220	0.303
Number of males aged 30-49 (EA mean)	0.584	0.184
Household size = 1	1.722	0.281
Household size = 8	1.587	0.546
Number of females aged younger than 10 (EA mean)	-0.692	0.266
Number of females aged 6-14 (EA mean)	-1.449	0.235
Number of females aged younger than 15 (EA mean)	1.112	0.288
Number of males with education P1-P4 years (EA mean)	0.444	0.105
Number of males with education P1-P4 years squared (EA mean)	-1.904	0.738

Table B3. First stage regression results for the Northern region

Dependent Variable: log of per capita consumption expenditure		
Number of observations:	1368	
Number of Clusters:	144	
Adjusted R^2	0.46	
Variable	Parameter estimate	Standard error
Intercept	10.225	0.093
Number of males with at least secondary school	0.061	0.029
Household size =5	0.090	0.036
Household size =13	0.366	0.121
Maximum years of education 13 squared	-0.001	0.000
Log of adult equivalent size	-0.681	0.052
Proportion of females aged 30-49 squared	0.350	0.141
Number of males with education years 1 to 4	-0.083	0.019
Number of males with primary education	0.101	0.016
Proportion of males with education O'level and above	0.512	0.179
Number of females aged 30 or older	0.092	0.026
Buffer zone within 1km from main road (parish)	0.682	0.233
Buffer zone within 1km from tarmac road (parish)	6.153	1.623
Buffer zone within 3 km from tarmac road (parish)	-8.865	1.692
Buffer zone within 4 km from tarmac road (parish)	5.732	0.969
Proportion of subsistence farmland (parish)	-0.130	0.054
Proportion of wet subsistence farmland (parish)	-3.714	1.160
Proportion of water (parish)	0.856	0.140
Age of household head age* tribe Lugbar	0.007	0.002
Age of household head age* district Arua	0.008	0.002
Meal hh. head separated or divorced squared	2.866	1.169
Maximum years of education* tribe Madi	0.057	0.008
Number of males aged 30 and above* district Arua	-0.143	0.051
Number of males aged 50 and above * Head male separated divorced	-0.406	0.106
Number of males aged 50 and above*tribe Lugbar	-0.569	0.114
Number of females aged 15 and below* district Apac	0.066	0.012
Age of Household head* Proportion of parish within 1km from main road	-0.020	0.004
Log of adult equivalent size * Distric Gulu	-0.346	0.087
Log of adult equivalent size * Prop. of parish within 1km from main road	0.253	0.115
Log of adult equivalent size * Prop. of parish within 1km from track road	0.105	0.047
Head males separated divorced * district Gulu	0.564	0.233
Head males separated divorced * district Kitgum	-0.445	0.176
Head males separated divorced * district Nebbi	-3.059	1.076

Maximum years of education * district Gulu	0.059	0.011
Maximum years of education* district Lira	0.015	0.005
Maximum years of education* district Moroto	0.106	0.041
Maximum years of education is 13 years* district Gulu	0.025	0.012
Number of males aged 30 and above* district Moyo	-0.177	0.063
Number of males aged 50 or older *Main road buffer zone of 1km	0.609	0.133
Proportion of females aged 0-5 squared (EA mean)	-4.514	1.140
Proportion of females aged 45 plus (EA mean)	-0.599	0.134

Table B4. First stage regression results for the Western region

Dependent Variable: log of per capita consumption expenditure		
Number of observations:	1637	
Number of Clusters:	163	
Adjusted R^2	0.34	
Variable	Parameter estimate	Standard error
Intercept	10.391	0.111
Number of females aged 6-14	0.047	0.017
Number of males with education above O'level	0.079	0.037
Household size squared	0.004	0.001
Household size = 11	-0.343	0.101
Log of household size	-0.246	0.041
Proportion of females aged 0-5 squared	0.934	0.235
Proportion of females aged 30-49 squared	0.451	0.129
Number of males with no education	-0.077	0.013
Number of males with education 1 to 4 years	-0.076	0.016
Age of Household head squared	0.000	0.000
Proportion of parish within 1 km from track road	0.975	0.165
Proportion of parish within 2 km from track road	-0.684	0.145
Proportion of parish within 3 km from tarmac road	0.169	0.049
Proportion of parish within 4 km from track road	0.226	0.066
Proportion of parish under woodlot	-6.715	2.067
Proportion of parish under subsistence farmland	-0.240	0.053
Proportion of parish under wet subsistence farmland	1.096	0.300
Log of household heads age* tribe Kiga	0.034	0.013
Log of household heads age* tribe Konjo	0.206	0.028
Log of household heads age * tribe Nkole	0.107	0.013
Mean education years = 18 * tribe Alur	0.216	0.082
Mean education years =18* tribe Nkole	-0.023	0.011
Mean education years =18* tribe Nyoro	-0.083	0.018
Head no education* tribe Alur	-1.828	0.560
Head male separated divorced* tribe Konjo;	0.574	0.250
Maximum years of education * tribe Alur;	-0.230	0.062
Maximum years of education*tribe Ganda;	0.231	0.077
Log of household heads age*district Hoima;	0.071	0.018
Log of household heads age* district Kasese;	-0.134	0.029
Mean education years 18* Proportion of parish under towns	-1.815	0.881
Head no education * district Kabarole;	-0.157	0.052
Proportion of males with no education*prop.of parish under towns	-11.510	4.072
Head males separated divorced* district Hoima;	0.478	0.198
Household size =6* district Kabarole;	0.348	0.098
Household size =6* district Kabale;	0.367	0.127
Number of males with education above O'level (EA mean)	0.840	0.172
Number of females aged 35 or older (EA mean)	-0.419	0.096

Appendix C: Poverty estimates at district level

Table C1. Rural Strata: District Mean Per capita Expenditure, Poverty and Inequality Estimates

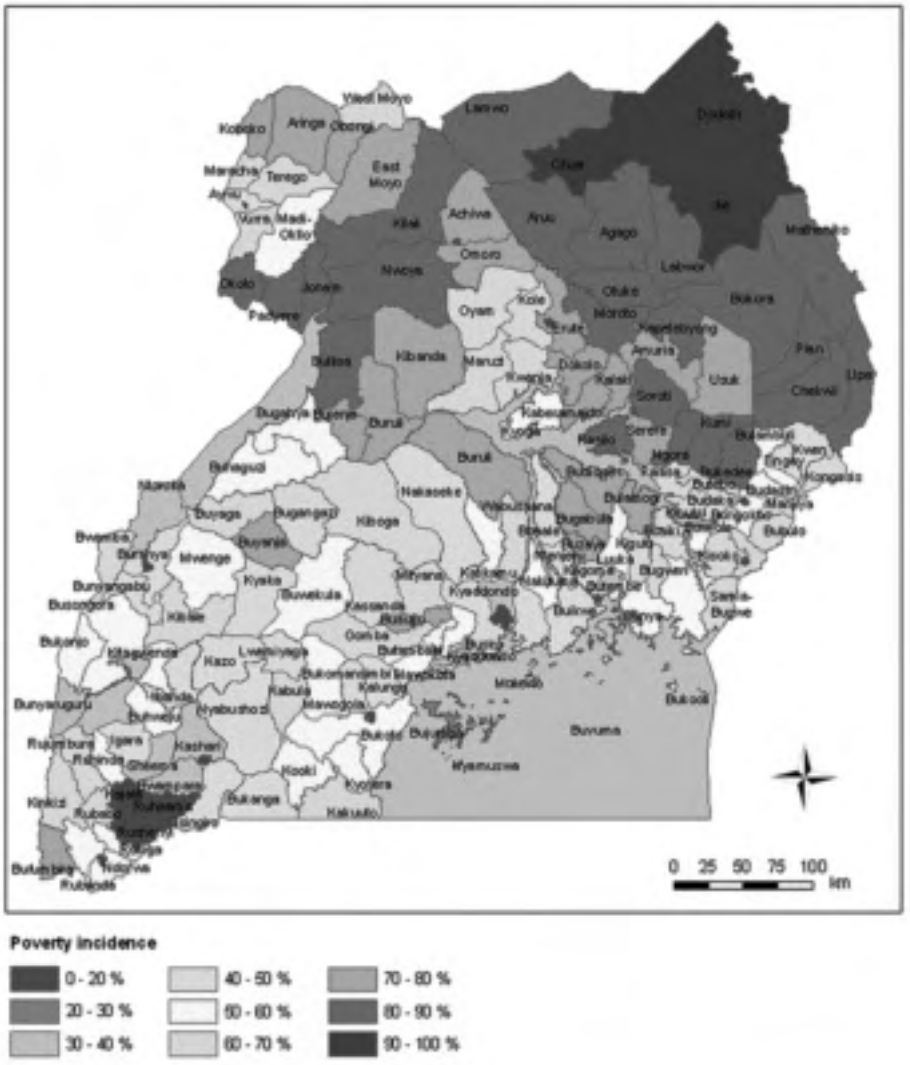
Code	District	Population	Mean Y	FGT0	FGT1	FGT2
Central						
11	Kalangala	14,079	26452.51	25.09	6.21	2.25
			(2198.34)	(0.05)	(0.02)	(0.01)
17	Kiboga	131,445	15858.74	62.11	22.20	10.43
			(756.16)	(0.03)	(0.02)	(0.01)
23	Luwero	403,948	17501.48	55.45	18.41	8.21
			(527.67)	(0.02)	(0.01)	(0.01)
24	Masaka	723,415	18651.63	50.34	15.77	6.74
			(558.90)	(0.02)	(0.01)	(0.00)
30	Mpigi	761,066	19671.96	48.82	15.91	7.05
			(722.53)	(0.03)	(0.01)	(0.01)
31	Mubende	445,077	16176.08	63.00	23.20	11.10
			(888.90)	(0.03)	(0.02)	(0.01)
32	Mukono	705,227	19077.89	49.45	15.94	7.01
			(674.38)	(0.02)	(0.01)	(0.01)
35	Rakai	361,501	16312.77	60.87	21.49	9.99
			(563.14)	(0.02)	(0.01)	(0.01)
East						
7	Iganga	885,398	23364.78	58.38	20.59	9.58
			(8399.97)	(2.24)	(1.24)	(0.75)
8	Jinja	203,021	65272.74	39.53	12.38	5.35
			(54955.40)	(3.37)	(1.36)	(0.69)
13	Kamuli	460,682	12789.89	73.89	30.62	15.93
			(835.61)	(3.68)	(2.96)	(2.07)
14	Kapchorwa	102,019	19059.53	45.81	13.73	5.71
			(1677.03)	(6.31)	(2.70)	(1.37)
21	Kumi	216,150	10945.13	82.40	37.00	20.20
			(776.46)	(3.37)	(3.16)	(2.34)
26	Mbale	640,929	16205.49	58.85	20.51	9.49
			(545.66)	(2.26)	(1.29)	(0.78)
34	Pallisa	347,936	14909.63	63.66	23.01	10.90
			(485.59)	(2.15)	(1.29)	(0.81)
37	Soroti	358,452	11741.12	78.66	34.09	18.22
			(742.83)	(3.33)	(2.76)	(1.96)
38	Tororo	483,104	17926.81	62.84	22.84	10.83
			(1933.28)	(2.00)	(1.28)	(0.82)

Table C2. Rural Strata: District Mean Per capita Expenditure, Poverty and Inequality Estimates

Code	District	Population	Mean Y	FGT0	FGT1	FGT2
North						
1	Apac	440,829	15661.78	64.34	23.61	11.33
			(790.82)	(2.91)	(1.64)	(1.01)
2	Arua	600,141	16778.39	64.01	23.21	11.00
			(862.38)	(2.93)	(1.80)	(1.10)
5	Gulu	277,223	12081.47	79.77	38.53	22.21
			(652.15)	(1.96)	(2.02)	(1.65)
19	Kitgum	327,085	13140.80	88.21	41.92	23.45
			(30480.45)	(1.29)	(1.54)	(1.24)
20	Kotido	111,552	8817.79	90.90	47.29	28.30
			(424.99)	(1.39)	(2.13)	(1.89)
22	Lira	454,193	13526.99	73.46	29.95	15.34
			(577.26)	(2.35)	(1.77)	(1.22)
28	Moroto	123,002	11349.58	83.74	42.62	25.44
			(1609.66)	(2.67)	(2.20)	(1.74)
29	Moyo	132,801	13994.23	70.20	28.16	14.36
			(664.04)	(2.75)	(1.84)	(1.23)
33	Nebbi	286,352	10019.24	87.93	40.72	22.27
			(327.17)	(1.32)	(1.58)	(1.27)
West						
3	Bundibugyo	103,236	16035.53	57.82	23.14	12.22
			(1100.32)	(3.88)	(2.51)	(1.72)
4	Bushenyi	711,713	18688.97	44.60	14.71	6.89
			(753.93)	(2.97)	(1.31)	(0.71)
6	Hoima	188,347	17334.30	52.34	19.01	9.48
			(1452.48)	(5.87)	(2.97)	(1.74)
9	Kabale	382,099	15746.15	55.93	19.74	9.60
			(858.46)	(4.05)	(2.01)	(1.15)
10	Kabarole	693,706	16887.08	51.31	17.82	8.60
			(704.22)	(3.00)	(1.43)	(0.81)
15	Kasese	294,155	15962.43	55.47	19.69	9.62
			(1314.83)	(6.27)	(3.25)	(1.88)
16	Kibaale	212,124	13310.58	68.60	26.68	13.71
			(614.08)	(3.11)	(2.12)	(1.38)
18	Kisoro	176,360	12929.51	70.26	27.33	14.03
			(749.21)	(4.07)	(2.68)	(1.71)
25	Masindi	225,504	11852.71	76.20	33.58	18.74
			(879.23)	(3.91)	(3.37)	(2.51)
27	Mbarara	865,415	19429.69	42.49	13.87	6.45
			(749.62)	(2.53)	(1.06)	(0.56)
36	Rukungiri	371,360	13854.28	65.35	24.11	11.99
			(594.17)	(3.17)	(1.78)	(1.07)

Appendix D: Overlays of poverty and biomass

Figure D1: Map of poverty incidence in Uganda based on the poverty estimates with biomass.



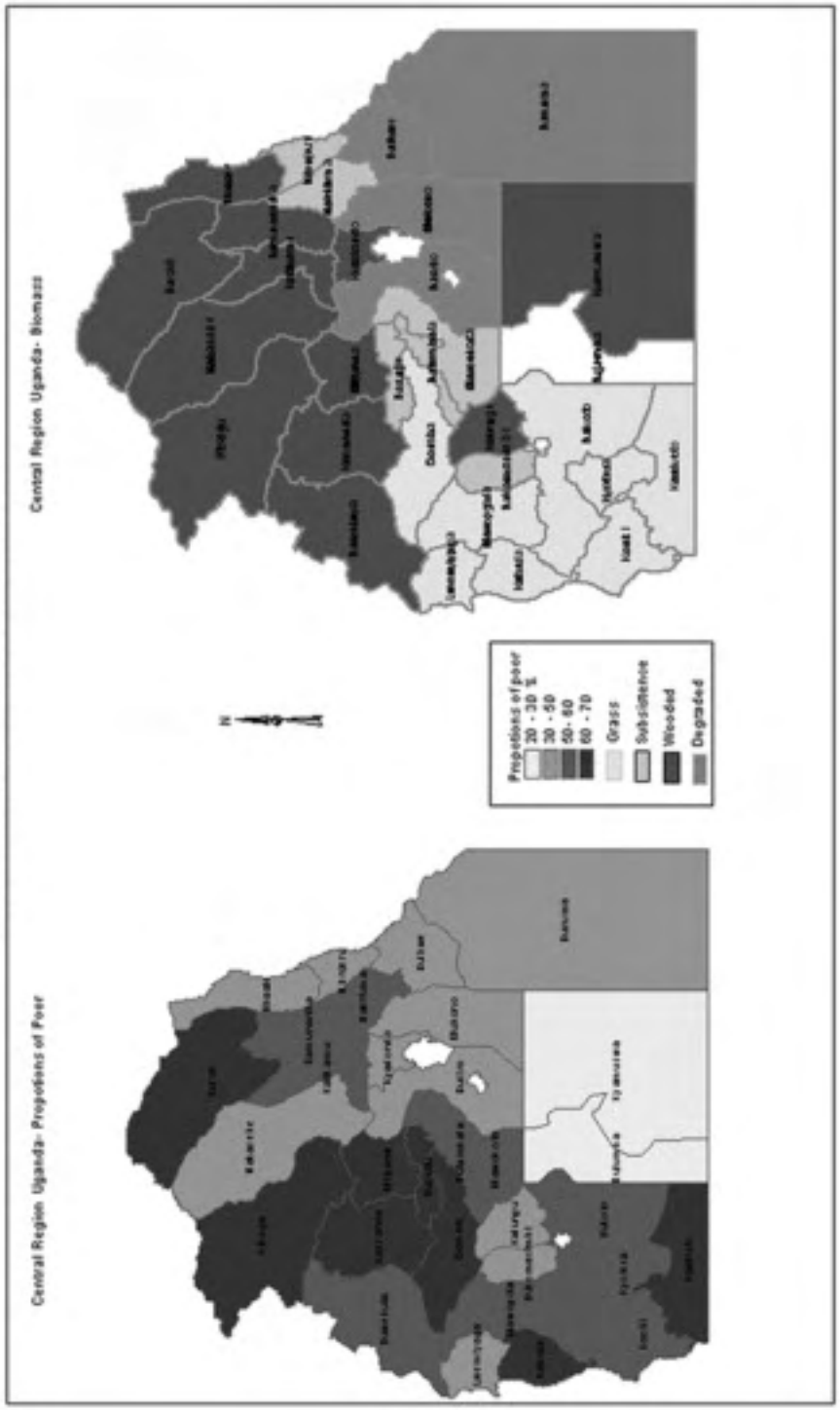


Figure D2: Poverty and biomass in Central region, Uganda, 1992

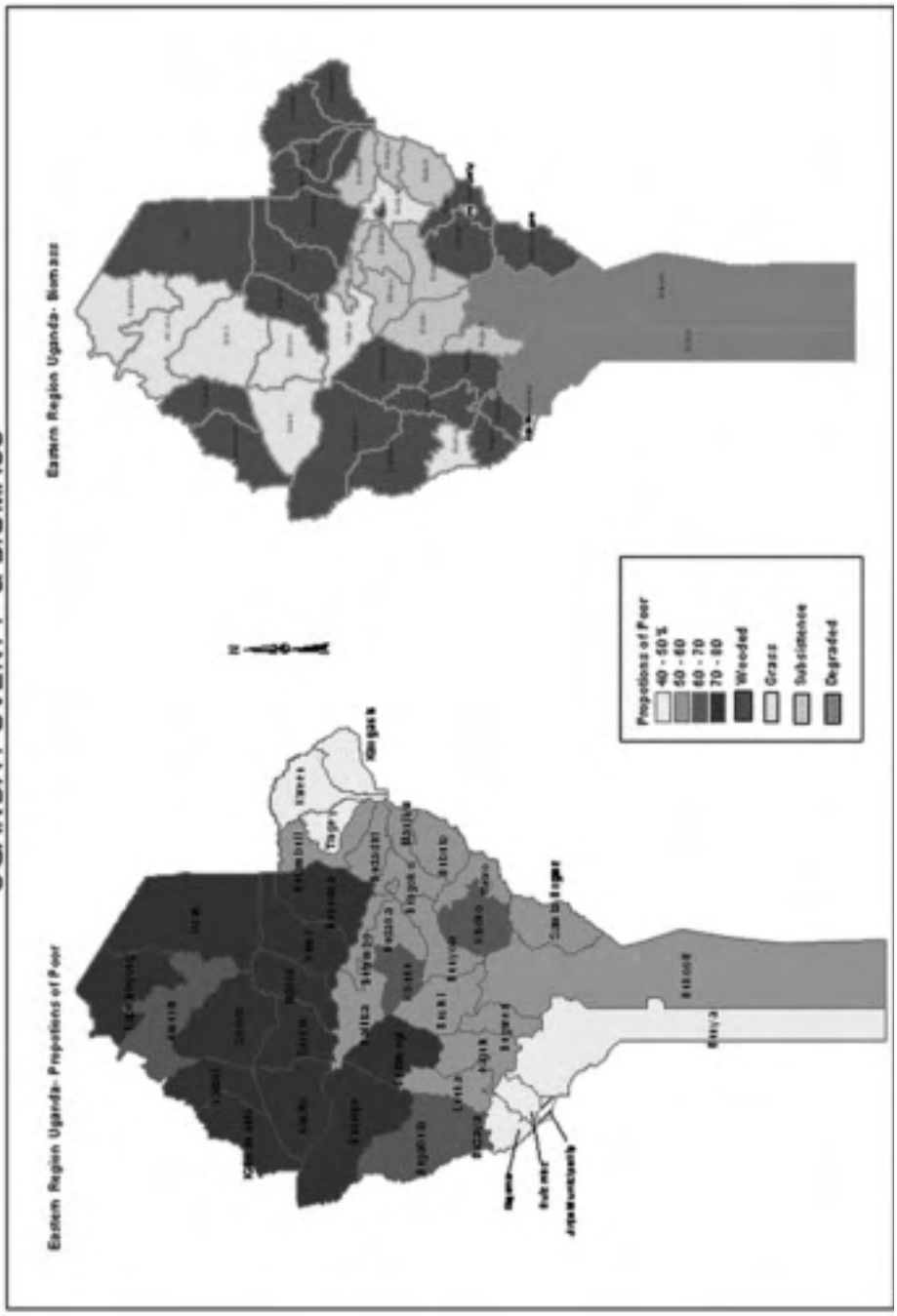


Figure D3: Poverty and biomass in Eastern region, Uganda, 1992

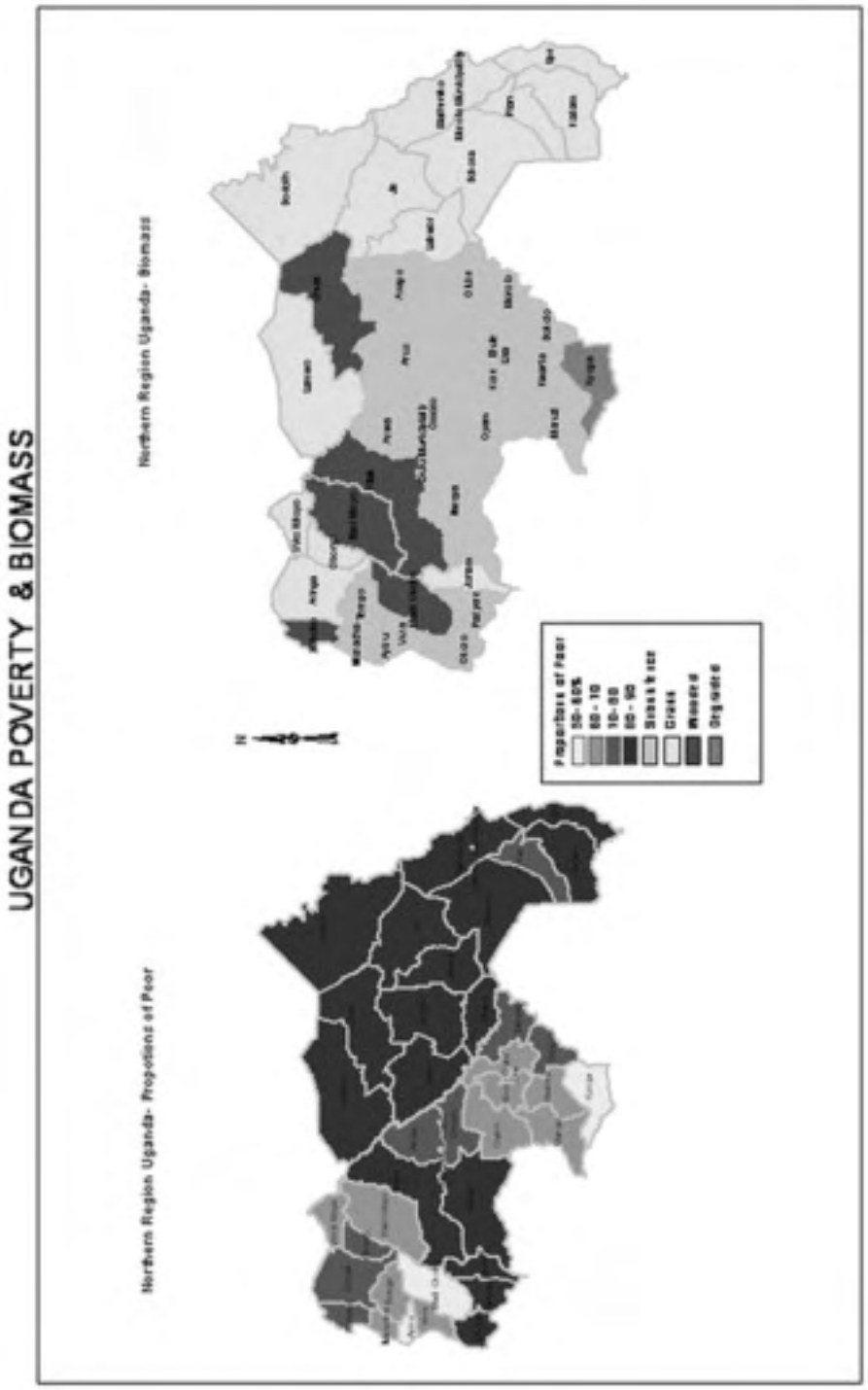


Figure D4: Poverty and biomass in Northern region, Uganda, 1992

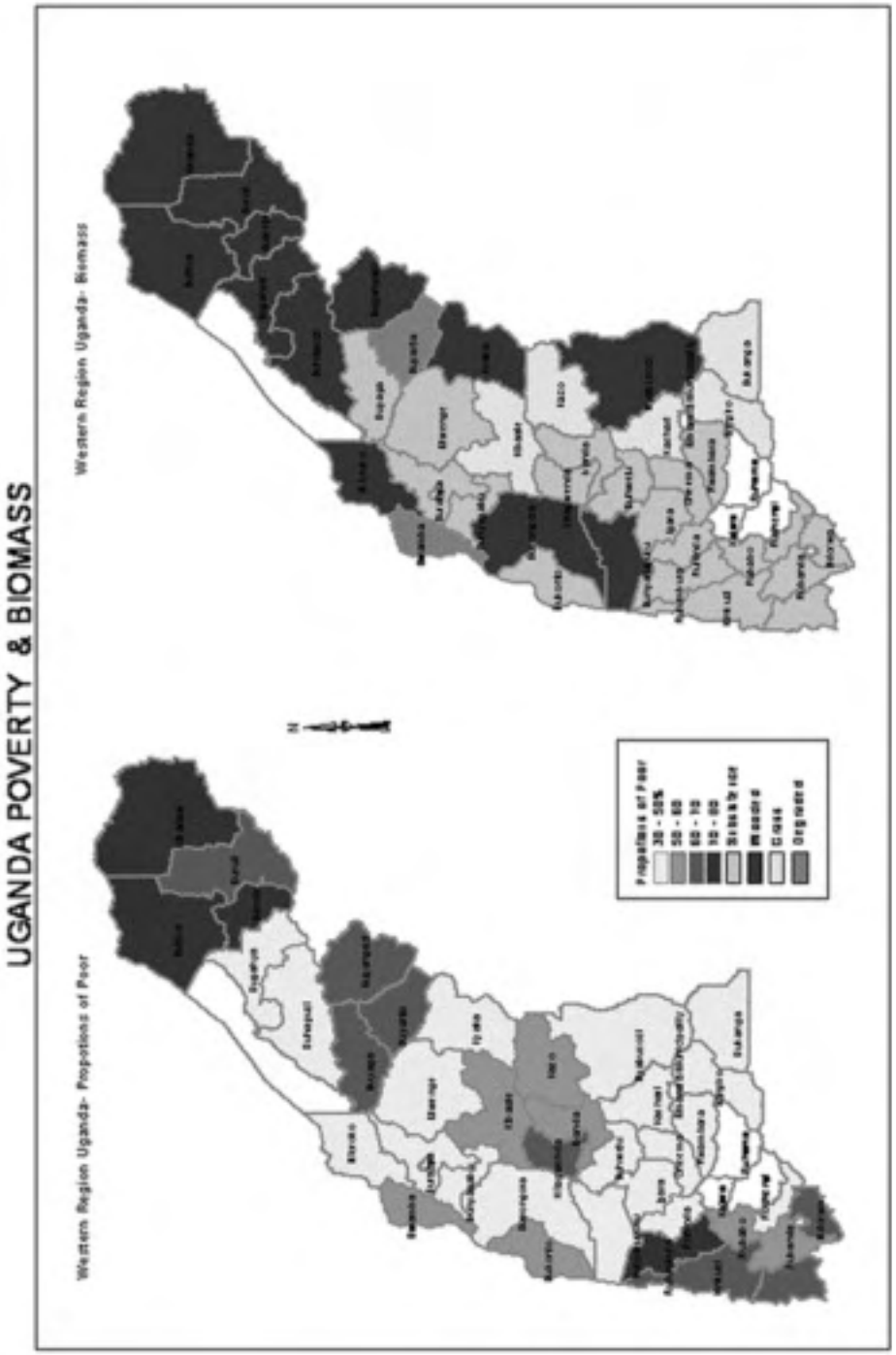


Figure D5: Poverty and biomass in Western region, Uganda, 1992

Appendix E: Correlations between biomass and poverty

Table E1: Correlation coefficients between land use (biomass) and poverty incidence* at county level

Biophysical variable	Uganda	Central	East	North	West
Buffer zones of main road					
1 km	-0.031	0.136	-0.646	0.033	-0.653
2 km	-0.011	0.172	-0.662	0.027	-0.650
3 km	0.012	0.206	-0.664	-0.001	-0.648
4 km	0.046	0.227	-0.643	-0.019	-0.641
5 km	0.085	0.232	-0.604	-0.030	-0.627
Buffer zones of tarmac road					
1 km	-0.347	0.000	-0.511	-0.502	-0.173
2 km	-0.337	0.045	-0.507	-0.506	-0.220
3 km	-0.329	0.081	-0.501	-0.506	-0.267
4 km	-0.321	0.109	-0.494	-0.502	-0.300
5 km	-0.315	0.133	-0.491	-0.498	-0.324
Buffer zones of tracks					
1 km	-0.052	0.125	-0.365	-0.408	-0.337
2 km	-0.045	0.171	-0.401	-0.451	-0.356
3 km	-0.019	0.200	-0.409	-0.474	-0.363
4 km	0.015	0.214	-0.390	-0.462	-0.368
5 km	0.054	0.221	-0.365	-0.434	-0.369
Land use covers					
Hardwoods	-0.238	-0.068	-0.357	-0.109	-0.527
Softwoods	-0.072	-0.006	-0.361	-0.085	0.116
Tropical high forest -normal	-0.294	-0.564	-0.113	0.277	-0.020
Tropical high forest –depleted	-0.208	-0.160	-0.161	0.307	-0.113
Woodlands and bush lands	0.409	0.046	0.587	0.245	0.345
Grasslands	0.095	0.172	0.619	-0.417	0.444
Wetlands	-0.026	0.172	-0.071	-0.154	-0.080
Subsistent farmland	-0.135	0.315	-0.614	0.008	-0.409
Commercial farmland	-0.224	-0.103	-0.428	0.124	-0.424
Subsistent farmland/wetlands#	-0.069	-0.025	-0.429	0.305	-0.064
Built up areas	-0.322	-0.302	-0.673	-0.065	-0.460
Water	-0.239	-0.661	0.000	0.077	-0.146

* The poverty incidences are derived from Okwi *et al.* (2003), and therefore are the poverty estimates without biophysical information.

Subsistent farmland/wetland is part of the Subsistent farmland.

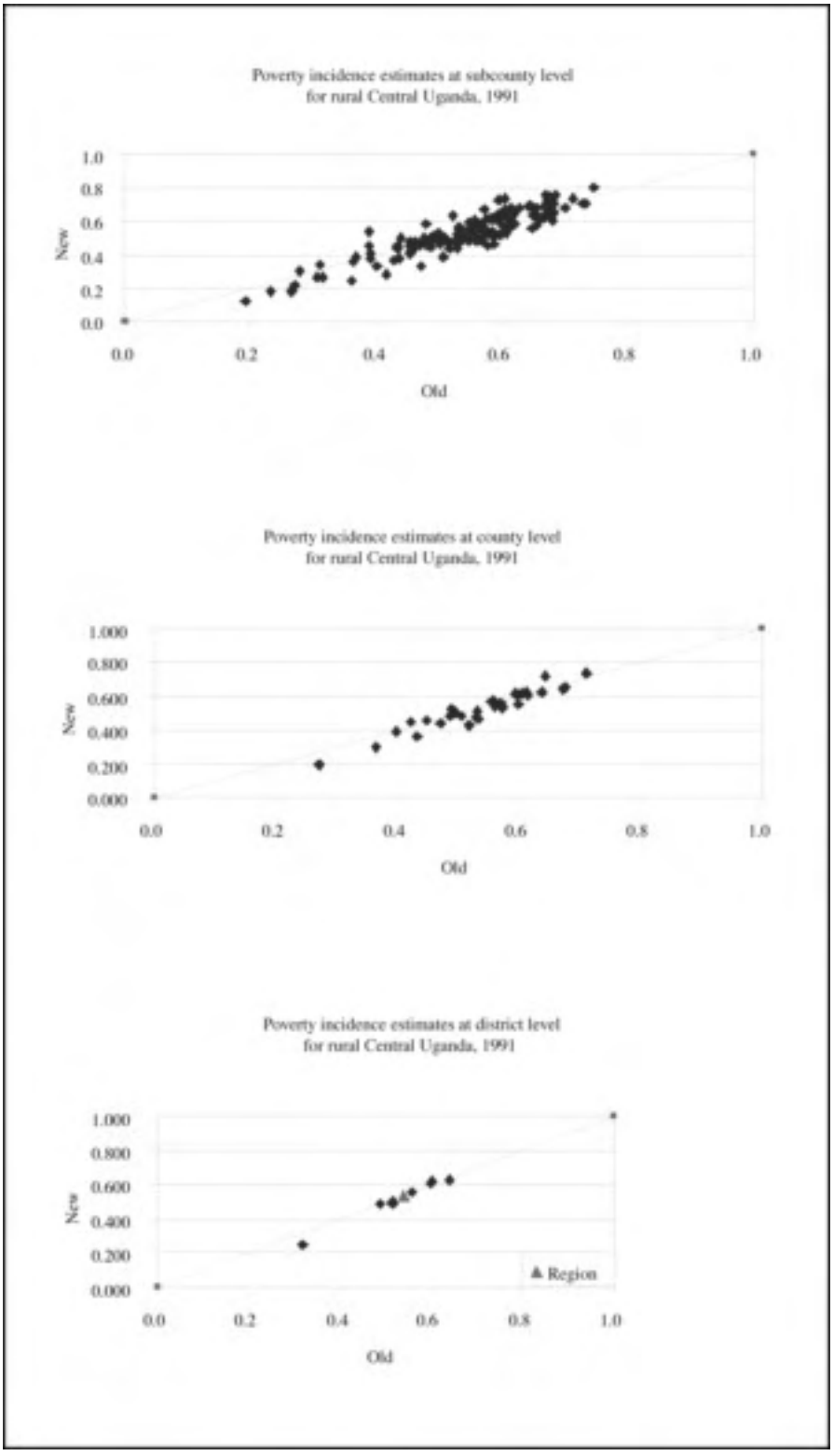
Table E2: Correlation coefficients between land use (biomass) and Number of poor people* at county level

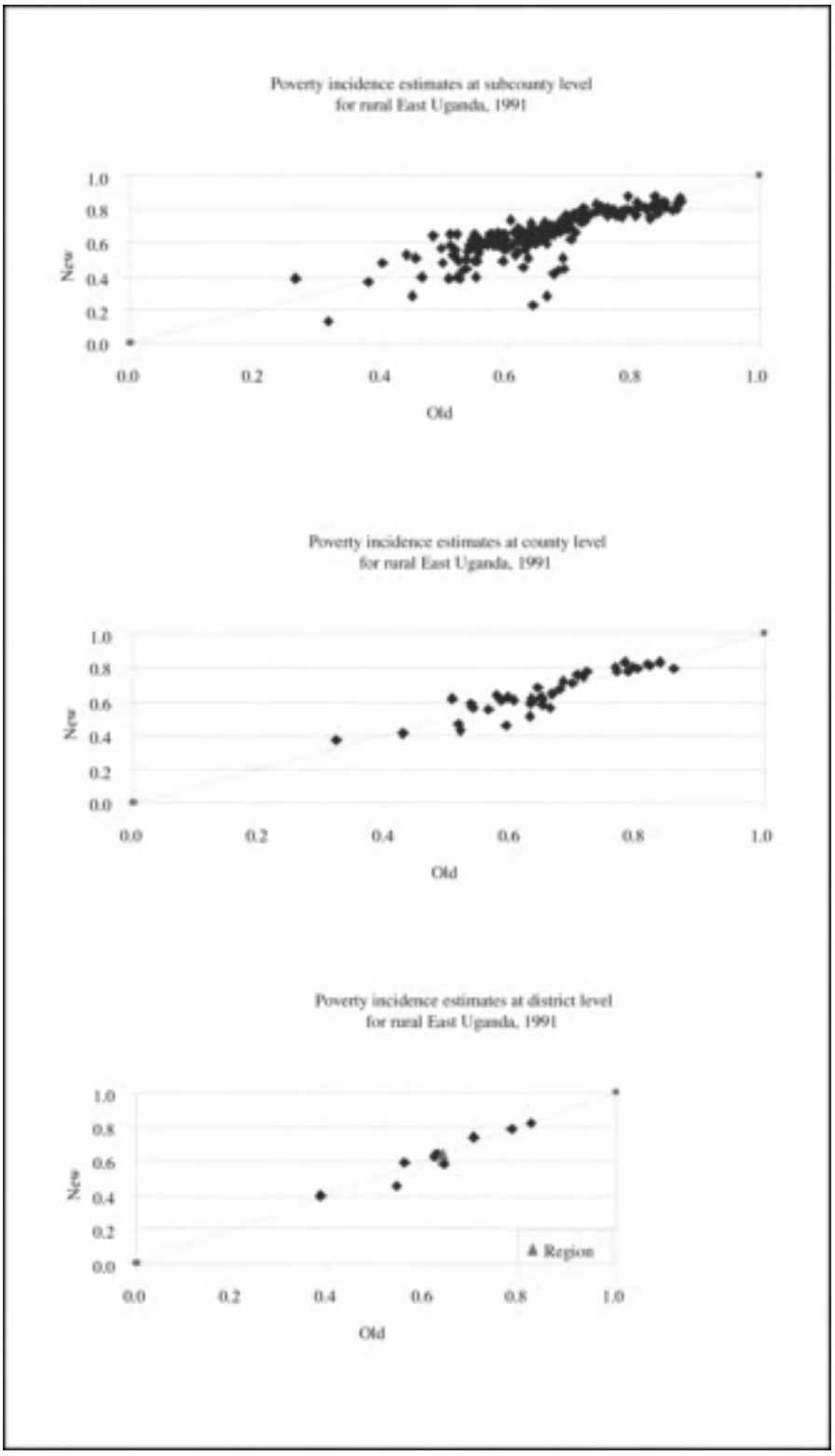
Biophysical variable	Uganda	Central	East	North	West
Buffer zones of main road					
1 km	0.068	0.003	0.124	0.478	0.075
2 km	0.048	-0.003	0.091	0.455	0.109
3 km	0.029	-0.014	0.053	0.442	0.146
4 km	0.010	-0.025	0.018	0.436	0.162
5 km	-0.012	-0.032	-0.020	0.424	0.167
Buffer zones of tarmac road					
1 km	0.115	0.033	0.183	0.093	0.590
2 km	0.117	0.031	0.193	0.088	0.590
3 km	0.124	0.034	0.206	0.090	0.575
4 km	0.129	0.038	0.217	0.093	0.552
5 km	0.132	0.041	0.224	0.098	0.522
Buffer zones of tracks					
1 km	0.071	0.062	0.241	0.101	0.280
2 km	0.040	0.047	0.214	0.068	0.285
3 km	0.012	0.024	0.155	0.066	0.290
4 km	-0.014	0.001	0.088	0.047	0.303
5 km	-0.036	-0.017	0.022	0.012	0.324
Land use covers					
Hardwoods	0.037	0.239	-0.023	0.393	-0.060
Softwoods	0.133	0.157	-0.170	0.274	0.255
Tropical high forest -normal	-0.098	-0.335	0.383	-0.024	-0.050
Tropical high forest –depleted	0.045	0.194	0.372	-0.250	-0.239
Woodlands and bush lands	-0.259	-0.242	-0.381	-0.322	-0.295
Grasslands	-0.386	-0.181	-0.539	-0.419	-0.176
Wetlands	0.027	0.068	0.164	-0.230	-0.414
Subsistent farmland	0.381	0.448	0.279	0.577	0.460
Commercial farmland	-0.065	0.251	-0.139	-0.091	0.292
Subsistent farmland/wetlands#	0.165	0.013	0.318	0.521	0.456
Built up areas	0.144	0.179	0.103	0.245	0.081
Water	0.020	-0.346	0.485	0.167	-0.275

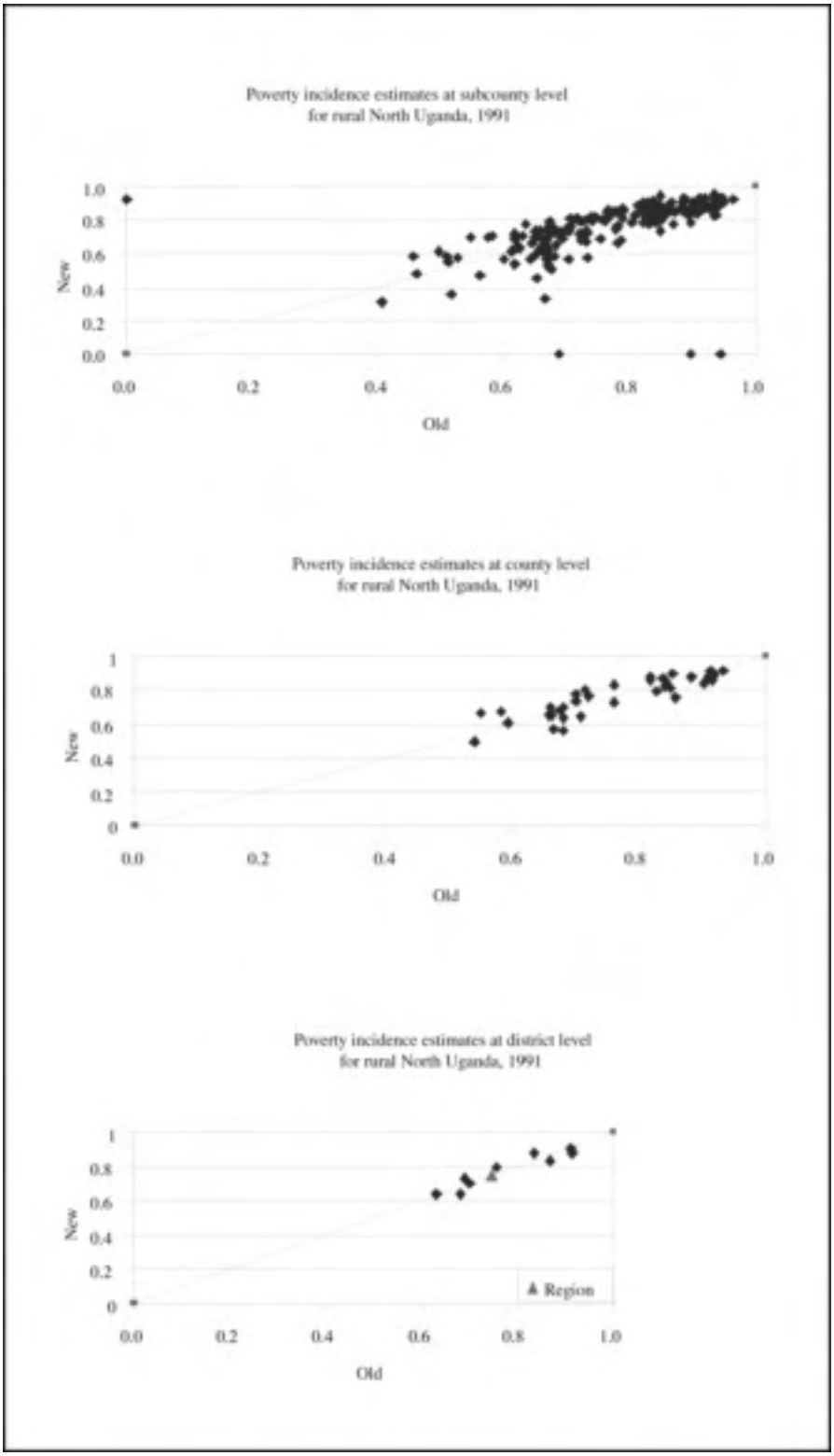
* The poverty incidences are derived from Okwi *et al.* (2003), and therefore are the poverty estimates wit hout biophysical information.

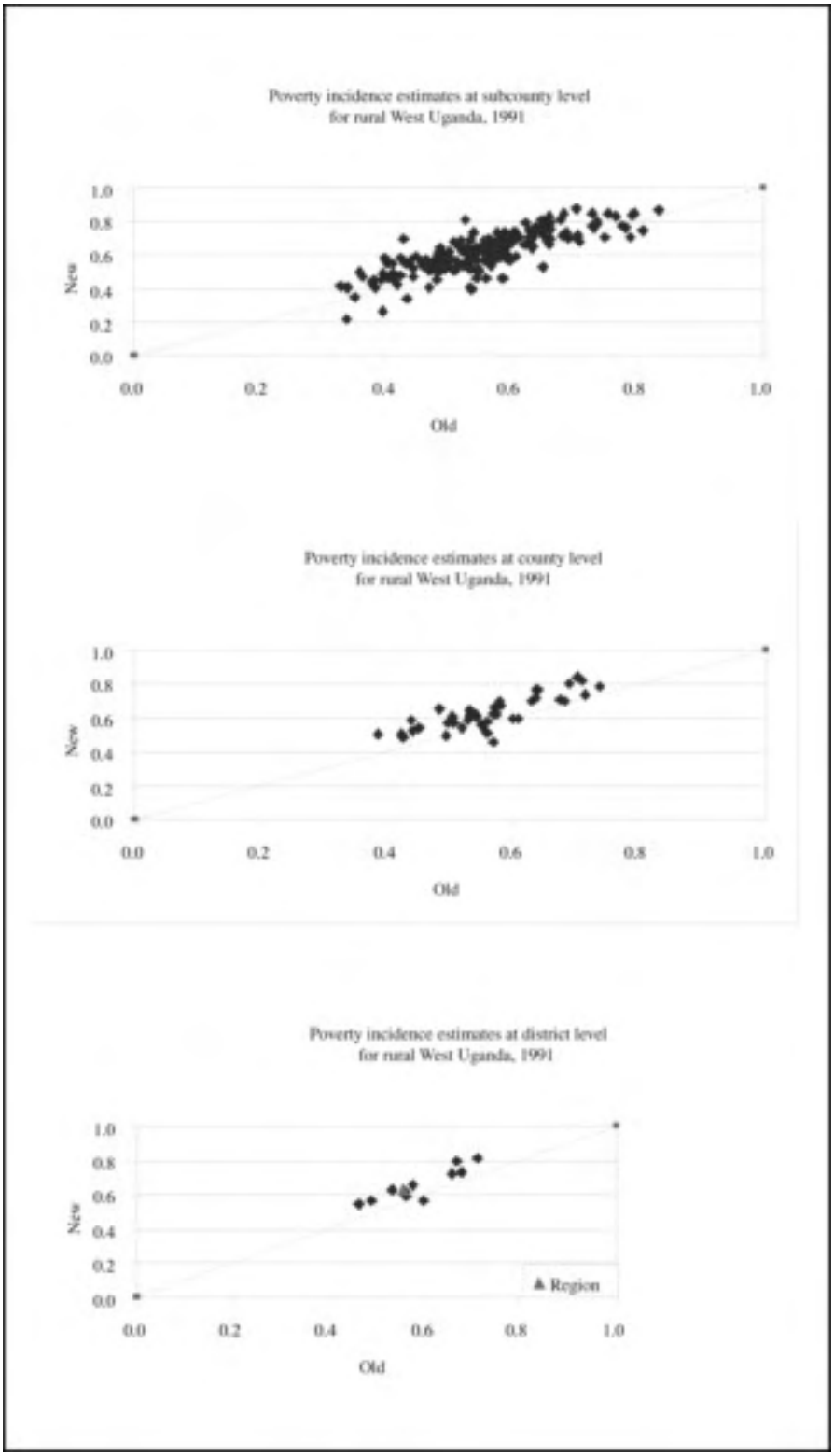
\# Subsistent farmland/wetland is part of the Subsistent farmland.

Appendix F Comparison of old and new poverty estimates




Poverty incidence estimates at subcounty level
for rural East Uganda, 1991
New
Old
0.0
0.2
0.4
0.6
0.8
1.0
Poverty incidence estimates at county level
for rural East Uganda, 1991
New
Old
Poverty incidence estimates at district level
for rural East Uganda, 1991
New
Old
Region


Poverty incidence estimates at subcounty level
for rural North Uganda, 1991
New
Old
0.0
0.2
0.4
0.6
0.8
1.0
Poverty incidence estimates at county level
for rural North Uganda, 1991
New
Old
0
0.2
0.4
0.6
0.8
1
Poverty incidence estimates at district level
for rural North Uganda, 1991
New
Old
Region


Poverty incidence estimates at subcounty level
for rural West Uganda, 1991
New
Old
0.0
0.2
0.4
0.6
0.8
1.0
Poverty incidence estimates at county level
for rural West Uganda, 1991
New
Old
Poverty incidence estimates at district level
for rural West Uganda, 1991
New
Old
Region

Experiences in the Application of the Core Welfare Indicator Questionnaire (CWIQ) Survey Technology in Africa: The Journey So Far

O. O. Ajayi [1]

Summary

The improvement of African statistical processes was a great motivation for the developments of CWIQ Survey Technology. CWIQ was also developed as a package to monitor social and economic programmes and strategies with respect to the state of availability, accessibility and satisfaction of services. CWIQ technology was applied first in Africa.

CWIQ has the advantage that it is simple, cheap and quick to implement. The core questionnaire (fairly inflexible) has been developed with the possibility of adding supplementary modules in structures/formats with responses that can be scanned with pre-developed processing and tabulation programmes. We demonstrate how CWIQ could also be used to measure progress towards the MDG's. Progress made so far in the application of CWIQ in Africa is given as at January 2006.

The CWIQ technology could be adopted for other surveys covering substantive subject-matter areas and one such example is presented. The paper concludes with recent initiatives for capacity building with respect to training, survey techniques, raising the level of expertise of consultants to give technical assistance to countries and data archiving for already concluded CWIQ surveys to enhance further data analysis.

Key Words

CWIQ, data quality, supplementary module, core questionnaire, service delivery, indicators of access, uses and satisfaction, large sample, data disaggregation, agricultural sample census.

1: O. O. Ajayi was Director-General of Nigerian Federal Statistics between 1991 and 1999 and since retirement has been an International Statistical Consultant. P. O. Box 52724, Falomo Post Office, Ikoyi, Lagos. E-mail: ooajayi611@yahoo.com
The author is very grateful to the reviewer of the paper for comments and additions that substantially improved the content and nature of the paper.

Résumé

L'amélioration des processus statistiques Africains a fortement motivé le développement de la méthodologie d'enquête QUIBB. Le système QUIBB a été développé comme un instrument de suivi des programmes et stratégies sociaux et économiques en ce qui concerne la disponibilité, l'accessibilité et l'efficacité des services. Le système QUIBB a d'abord été appliqué en Afrique.

Le système QUIBB a l'avantage d'être simple, bon marché et pouvant être mis en œuvre rapidement. Le questionnaire principal (peu flexible) a été développé avec la possibilité d'ajouter des modules supplémentaires en structures/formats avec des réponses qui peuvent être lues rapidement par des programmes préconçus de traitement et de tabulation. Nous démontrons comment le système QUIBB peut être aussi utilisé pour mesurer le progrès accomplis vers la réalisation des OMD. Le progrès réalisé dans la mise en œuvre du système QUIBB en Afrique est présenté pour une période allant jusqu'en janvier 2006.

La méthodologie QUIBB pourrait être adoptée dans d'autres enquêtes traitant de disciplines spécifiques ainsi que le montre l'un des exemples présentés ici. L'article conclu par la présentation des initiatives récentes de renforcement de capacité relatives à la formation, aux techniques d'enquêtes, à l'amélioration du niveau d'expertise des consultants appelés à donner une assistance technique aux pays et à l'archivage des données d'enquêtes QUIBB déjà réalisées aux fins d'en approfondir l'analyse.

Mots clés

QUIBB, qualité des données, module complémentaire, questionnaire principal, fourniture de services, indicateurs d'accès, utilisation et satisfaction, grands échantillons, désagrégation de données, échantillon de recensement agricole.

1. Introduction

The state of statistical production and usage had been particularly appauling in the last couple of decades in Africa and especially in the 1990's. One of the critical problems of data generation has been the untimeliness of information delivery which was as a result of processing bottlenecks in nearly all the National Statistical Offices (NSO's) in Africa. Also of great concern with the African statistics was the quality of produced data. Errors creep into statistical outputs during data collection and data processing stages. Another issue was meeting the needs of statistics users in Africa

for sub-national estimates for purposes of planning, monitoring, evaluation and critical decision-making which could not be provided as a result of limited sample sizes in surveys. At the same time, African countries have embarked on many social and economic programmes including the adoption of Poverty Reduction Strategies (PRS). All these programmes and strategies require being monitored and being evaluated. Such questions as state of availability, accessibility and satisfaction of services need be answered. The World Bank in collaboration with some other international agencies developed the CWIQ Survey Technology which in its design attempted to remove the problems of untimeliness of data, poor data quality, lack of statistics at the lowest administrative levels and as an effective tool for programme monitoring and evaluation.

However, even though a pessimistic view of African statistics has been expressed implying probably that Africa trails behind the rest of the world, CWIQ is a product of Africa, but one that is now starting to be exported to other countries. It evolved in the late 90s as a response to the growing interest in, and commitment to, results-based development. If one is to focus on results, then one must have the means to measure them and to report on them quickly. Given also that the demand was increasingly focusing on outcomes rather than outputs, household surveys appeared to be the most appropriate tool for providing the data. But traditional surveys were slow, expensive, and difficult to implement. What was needed was something quite different. A survey that was quick, cheap and easy.

The CWIQ was the outcome of a series of brainstorming meetings to come up with a simple solution. The approach that was adopted was to think in terms of a "service delivery" framework. The role of government in development is essentially to provide an appropriate enabling environment for households and individuals to maximize their potential. It does this by providing a range of goods and services that should respond to households most urgent needs. The way to measure whether it is being successful and effective in achieving this goal is to monitor whether the goods and services are indeed reaching the target communities, and whether they are meeting their needs. How is this to be done? By asking the intended beneficiaries directly. For each key service, three indicators are proposed. They are access, use, and satisfaction. Does the respondent have access to public services? Does he/she make use of them? If no, why not? If yes, is he/she satisfied with the quality of service? By including a few additional questions on the socio-economic status of respondent it is possible to measure the extent to which the poorest and most vulnerable households are or are not benefiting from public service programmes. The CWIQ survey thus also has the potential to serve as a very effective instrument for monitoring the implementation of Poverty Reduction Strategies.

At heart, however, it remains primarily a service delivery survey. Because it is relatively simple and cheap, it can be administered to relatively large samples. It can, therefore, also serve as a tool for providing indicators that can be disaggregated to relatively low levels – very useful for supporting the process of decentralization and decentralized decision-making.

Currently, CWIQ addresses monitoring of welfare programmes but could indeed be adopted for regular economic, social and agricultural surveys; it could also handle censuses.

2. General Overview of The CWIQ Survey Technology

2.1 About CWIQ

CWIQ has been developed as an annual national survey which is quick and easy to implement. It is also programmed to be one of the components of an integrated overall national programme of household surveys. This approach requires that a country has a strong statistical system which is a necessary infrastructure needed to support planning, monitoring and evaluation of programmes particularly Poverty Reduction Strategies. CWIQ in this context provides rapid monitoring of key indicators, measurement and evaluation of development outcomes nationally and for different population sub-groups. CWIQ also enables timely assessment of country performance of national programmes including the Millennium Development Goals (MDG's). Monitoring PRSP, assessing the progress of MDG's, assessing and monitoring such issues such as "good governance", "transparency", "accountability" and "results-driven development" need data that are easily accessible, timely and of high quality which CWIQ is able to provide. Indeed all these development frameworks should include statistical capacity building as a key component to strengthen these monitoring tools. See below what CWIQ could do with respect to the measurement of progress of the MDG's:

Table 1: Use of CWIQ to measure progress in MDG's

Goals	Indicators	Available Now	Potentially Available	Not Possible
Extreme Poverty and Hunger	Under weight or U-5	✓		
	Other Indicators			✓
Primary Education	All indicators	✓		

Gender Equality	All indicators except women in Parliament	✓		
Child Mortality	All indicators		✓	
Maternal Mortality	All indicators		✓	
HIV/AIDS/Diseases	All indicators		[✓]	
Environment	All indicators			✓
Global Partnership	All indicators			✓

To ensure rapid implementation, a short questionnaire, applied in a single-shot survey, has been developed with the following sections: Interviewer Information, Household Members, Education, Health, Employment, Household Assets, Household Amenities, Poverty Predictors, and children under 5. However, later expansion has allowed the addition of supplemental modules usually to address some specific issues in countries such as the Gender module (Nigeria), Flood Damage module (Mozambique), MICS module (Mozambique), Consumption module (Rwanda), HIV/AIDS module (Lesotho), Mental Health module (Burundi) and Community CWIQ (Tanzania). The structure of the questionnaire is such it allows for the arrangement and sequencing of multiple choice questions (usually pre-coded) for easy and rapid data collection. The processing strategy uses scanning technology and data entry is done through scanning of questionnaires so as to eliminate data entry bottlenecks. Also prepared ready for use are complete validation specifications and programmes. The standard tabulation plan and programmes have similarly been pre-prepared. One critical benefit and objective of CWIQ realized through this processing procedure was closeness and interaction of survey statisticians and data processing experts. Other important features of CWIQ are the use of large sample (allowing for sub-national estimates) but within the considerations of feasibility and resources availability and yet getting out the results quickly, rigorous control of data quality, quick data entry, validation and processing, simple reporting and of course use of fixed core and flexible modules. With these features, there is no limitation that this technology could be applied to other surveys covering specific subject matters.

2.2. Strengths and weaknesses of the CWIQ approach

As for the strengths of CWIQ, it is an off-the-shelf survey with built-in collection of best practices, producing high quality data and through rapid implementation. CWIQ represents a good tool for training and capacity building. It is also cost-effective with a national survey of 10,000 households costing about $330,000. However, the following weaknesses should be noted, namely, inflexibility of the core questionnaire, use of poverty predictors with no consumption measure, complex training logistics and the need

for technical assistance through the regional training centres. In spite of these and the name "Quick", CWIQ is an effective survey not to be considered as quick and rough!

2.3. Progress of CWIQ programme so far (as at January 2006)

Since 1997, several CWIQ surveys have been carried out in many African countries (Anglophone, Francophone and Lusophone) either as national surveys, sub-national surveys or just pilot surveys. Some countries have, indeed, conducted two or more national CWIQ Surveys and these include:

- Ghana in 1997 and 2003
- Malawi in 2002 and 2005
- Mozambique (2002, 2003 and 2005) with the 2003 combined with Household Budget Survey and 12-month field work
- Rwanda (2001 and 2003)

Those countries with one national CWIQ survey completed are: Benin (2003), Burkina Faso (2003 combined with expenditure/consumption module with single visit), Burundi (2002), Democratic Republic of Congo (2005 with expenditure/ consumption module with a single visit), Gabon (2005 with expenditure/consumption module and a single visit), Guinea (2002 also with expenditure/consumption module and one single visit), Guinea-Bissau (2002), Lesotho (2002), Mali (2001 with Household Budget Survey over 12-month field work), Niger (2005 with expenditure/ consumption module and one field visit) and Senegal (2001 along with Household Budget Survey over a 12-month field work).

The first national surveys is underway in Nigeria while the second national surveys are ongoing in Mali and Burkina Faso in which CWIQ is being combined with a module on expenditure/consumption with only one field visit. Planning is currently going on for first national CWIQ in Angola, Equitorial Guinea, Ethiopia, Kenya, Sao Tome, Sierra Leone, Togo with consumption module and Zambia while the third national CWIQ is being planned for Malawi.

Sub-national CWIQ has also been implemented in some of these countries where, for instance, in Nigeria various states (6 states were covered in 2000) and in Tanzania where various districts have been covered between 2002 and 2004.

As indicated above, additional CWIQ modules covered included for Ghana (on social capital) and for Malawi (on HIV/AIDS). Greater details of the surveys, modules, du-

ration and sample sizes, are given below for some Francophone countries so as to present the complete picture:

Table 2: Details of application of CWIQ survey in some Francophone countries

Country	Date of Survey	Survey Type	Survey Module	Duration	Sample Size
Burundi	Feb. 2002	CWIQ	Mental Health + Alcohol use	3 Months	6,000 HH's
Guinea	Sept. 2002	HBS + CWIQ	Budget + Consumption Community Prices	1 year 3 rounds	7,000 HH's
Mali	Jan. 2001	HBS + CWIQ	Budget + Consumption	1 year 4 rounds	7,000 HH's
Senegal	Jan. 2001	HHS + CWIQ	Budget + Consumption & Perception of poverty	1 year 3 rounds	6,500 HH's
Rwanda	Jan. 2001	CWIQ	Consumption Module	58 days	5,400 HH's
Rwanda	Mar. 2003	CWIQ	Agriculture + Livelihood + AIDS	76 days	5,700 HH's

HBS - Household Budget Survey

3. Application and Methods of Implementation of CWIQ in Africa

Although is survey technology has been applied in Africa, certainly it is useful for data production situations in all developing countries. Use of CWIQ technology has taken place in more than half of Africa (about 26 countries) and in fact has been integrated into the Household Survey Programmes of many of them. As the usefulness of CWIQ technology becomes obvious particularly its quickness in measuring development outcomes, the remaining African countries will join the train. Meanwhile, the countries that have adopted the survey technology covered the entire spectrum of Anglo-, Franco- and Lusóphone language groups demonstrating its adaptability.

Usually the procedure in a country is to carry out a Pilot Survey to test the feasibility of the adoption of the method and see how it integrates into the country's survey programme. The pilot could be on a purposive or probability sample. In situations where a probability sample has been used and a definite administrative unit has been surveyed, the survey results could be used for making decisions. This was the case in Nigeria, Lesotho and several other countries. After successful pilot surveys, a sub-national or national survey usually follows. Both the pilot, sub-national and national surveys have always been supported technically and financially by external donors. A second national survey is also expected to follow and to be supported by donors; thereafter, countries are expected to have internalized the procedure and from then on conduct annual CWIQ with domestic support and build up time series data.

In all the surveys so far (whether pilot or main surveys), use of roving teams has been common and found more effective. Even where resident interviewers are in use, roving teams were still put together who could be assisted by the resident interviewers. The use of roving teams has implications for costs of the surveys, quality of data and rapidity of the results.

The processing facility has always included desk-top computers, laptop computers, printers (dot and laserjet) and scanning machines. The costs of these sets of equipment for the pilot and first national and sub-national surveys have always been substantial proportions of the total costs of the surveys but these tend to go down as surveys are repeated annually.

The processing arrangement leads to a time lag of not more than a week and this allows for immediate feedbacks to the field and, in turn, this enhances the quality of collected data. Other quality assurance strategies used include detailed and intensive training of field staff, supervision arrangement of the field staff at many levels, good and well-written manuals for all categories of field staff, very rigorous validation procedures and detailed and meticulous processing steps. The quality level of data produced so far has usually been very high.

In each country, the activities begin with questionnaire improvement and modifications but these are usually very minor for the short core questionnaire. The limitation in the modification that could be allowed on the core questionnaire has opened an opportunity for countries to develop supplemental modules that address specific issues peculiar to the needs of the countries. For instance, there were the "Flood damage" module attached to the main questionnaire in Mozambique and "Gender" module attached to the pilot survey in one state of Nigeria (Benue State). On the core questionnaire modification, Mozambique improved the child module of the main questionnaire. There are plans in the future to even develop some supplemental modules that will be generally added to the core questionnaire. In this context, supplemental modules that are required and useful to all countries will be those on HIV/AIDS and Malaria, agricultural holding questionnaire which could be canvassed for agricultural census and the general household survey (GHS) as core for Integrated Household Surveys.

For most of the pilot surveys, the turn-around time for releasing the results has been between six (6) and eight (8) weeks which has demonstrated convincingly that the problem of untimeliness in survey data can be eliminated.

Despite CWIQ's limited scope, the analytical content of this monitoring package has been broadened to welfare analysis by enhancing it with a methodology for estimating a welfare function for ranking households across expenditure quintiles for poverty analysis. This methodology was used during the implementation of the Ghana CWIQ in 1997. The different phases of its implementation included, establishing working files through which predictor variables were identified and selection of poverty predictors which are proxy indicators, easy to measure and is reasonably accurate. These are consumption correlates usually derived through rigorous regression analysis using Household Budget Survey and similar other survey data. Predictors vary from country to country and even within a country.

Report writing is expected to be simple and quickly done so as to put out the results on time to users. This has been possible because of the processing strategy which ensures quick validation, use of prepared software packages (Teleform, Access, Excel) to validate and tabulate the data. The planned tables are completed only a couple of days after the conclusion of data collection. Of course, capacity for report writing has to be further built.

User/Producer workshops are expected to conclude the implementation of CWIQ and these have been held in many countries but not in all of them. However, it has to be seen that the interaction between users and producers of data through workshops should be an essential activity in the CWIQ system in order to enhance the usability of the data.

4. Holding Questionnaire for Agricultural Samples Census - Can it Not Adopt The CWIQ Format?

The state of agricultural statistics is generally poor in Africa because of the complexity in the survey implementation requiring objective measurement of areas planted to crops and yield. Apart from this complexity is the slowness of processing the collected data from very large samples particularly for the sample census of agriculture. Isn't there a window of opportunity of using the CWIQ technology to improve upon this situation for agricultural statistics?

The agricultural holding questionnaire used by Nigeria's National Bureau of Statistics [formerly the Federal Office of Statistics (FOS)] for its agricultural sample census of 1993/94 was a supplemental module for the National Integrated Survey of Households (NISH). It was pre-coded and structured nearly like the CWIQ with Questions

column, Response column, Coding column and skip instructions column. With a slight improvement in the design to provide "bubbles" for the multi-choice questions, it could use the CWIQ technology for data collection and data processing. Both qualitative and quantitative questions could be accommodated on this questionnaire. What needs to be done with agricultural surveys in Africa will be to develop interview methods with rigorous trainings to ensure proper questionnaire administration as alternative way of obtaining information previously obtained through objective measurements. Crop area and production could be obtained in local unit measures and be later converted into standard measures. The Holding Questionnaire used by FOS in 1993/94 could be the base to work upon in a CWIQ format. (see the questionnaire in the appendix).

This could work very well, particularly if the service delivery focus is maintained, and the questionnaire focuses on collecting the three access, use and satisfaction indicators for the range of public programs designed to stimulate agricultural production – extension, credit, inputs, land registration, etc. It would obviously be desirable to monitor changes in actual production levels as well. Using traditional methods, as remarked already this would involve the complicated process of taking sample crop cuts and weighing them. This would probably not be advisable for a light survey such as the CWIQ. But the CWIQ way around this would be to simply "ask the farmer" how much he/she produced; much simpler and, surprisingly, just as accurate as the so-called objective measurements under same conditions. However, even without measuring production, the monitoring of adoption and use of the different services by different population groups, would provide enormously useful early indicators of the effectiveness and value of the programs.

5. Recent Initiatives for Capacity Building

The World Bank has gone a step further to build capacity for the implementation of CWIQ in Africa. These various initiatives could speed up development and improvement of the CWIQ survey technology.

Specifically:

(i) Through a grant, the East Africa Statistical Training Centre (EASTC), Dar es Salaam was identified as a centre to train consultants and hire them to guide CWIQ Survey in Africa. A similar training centre exists in Cote d'Ivore for the franco-phone countries to do the same.

(ii) The centres were to expand their curricula to include CWIQ basic principles, procedures and technical components and the students of these centres taken through training in these techniques.
(iii) A network of experts is being established that should be working on various survey modules that could be adapted using CWIQ strategy.
(iv) Encouraging data archiving for already concluded CWIQ surveys so that the data are readily available for further analysis.
(v) Sharing of information through CWIQ newsletter which has been established.

6. Conclusions

CWIQ survey technology presents a strategy for improving data production capability in Africa. But it requires survey experts in Africa to devote time to developing the various survey modules that could adopt the CWIQ format. There should be sustainability of the approach and to achieve this, African governments should be committed to producing and using data for their various development programmes including making adequate funds available to the National Statistical Offices (NSO's). The technical/financial support being given by donors should progressively tapper off. The challenge to adjust to changing technology particularly in the equipment being used should be taken seriously.

References

Bureau of Statistics (2002). CWIQ Survey Report, Mareru, Lesotho

Delaine, G., Personal Communication

Federal Office of Statistics (1999). CWIQ Survey Main Report, Lagos State, Abuja, Nigeria

Federal Office of Statistics (2001). Draft Report of CWIQ Survey in Benue State, Abuja, Nigeria

Menye, Essimi (2003). Presentation, Advanced Training Workshop, EASTC, Dar es Salaam, Tanzania

Marchant Tim (2003). Presentation, Advanced Training Workshop, EASTC, Dar es Salaam, Tanzania

The World Bank - CWIQ Handbook and CD-Rom, Dec. 1999

Appendix

NASC 93/94-Q3 MARCH 1993

FEDERAL OFFICE OF STATISTICS

NATIONAL AGRICULTURE SAMPLE CENSUS 1993/94

PHASE I

HOLDING QUESTIONNAIRE

HOLDING IDENTIFICATION

	NAME	CODE
STATE		
L.G.A		
SECTOR		
E.A.		

Household Master sample No.: .....

No. of Holders in H/H.: .....

Serial No. of this Holding: .....

Name of Holder: .....

Sex of Holder: .....

Age of Holder: .....

Highest level of Educ. Attained: .....

Relationship to Head of HH: .....

Size of Holder's HH: .....

QUES S/NO.	QUESTION AND FILTER	RESPONSE	CODE	SKIP TO
	SECTION I HOLDING CHARACTERISTICS			
1.	Is holding owned by holder (a) As an individual	Yes ..... No	1 2	
	(b) Jointly with another member of the Household	Yes ..... No	1 2	
	(c) Jointly with members of other Households	Yes ..... No	1 2	
2.	What agricultural production activities are carried out in the holding (a) Crop farming	Yes ..... No	1 2	
	(b) Livestock/Poultry (Including Rabbitry)	Yes ..... No	1 2	
	(c) Fish farming	Yes ..... No	1 2	
	(d) Forestry	Yes ..... No	1 2	
3.	Is holding part of an enterprise engaged in other economic activities	Yes ..... No	1 2	
4.	What is the other major economic activity of the enterprise?	(Specify)	1 2	
5.	Where Is holding located? (a) Inside the holder's premises	Yes ..... No	1 2	
	(b) In the field around holder's residence or locality	Yes ..... No	1 2	
	(c) In a different locality, (but holder has no residence there)	Yes ..... No	1 2	
	(d) In a different locality, (but holder has residence there)	Yes ..... No	1 2	

QUES S/NO.	QUESTION AND FILTER	RESPONSE	CODE	SKIP TO
6.	Does the holder own other holdings?	Yes …….. No ……	1 2	
	SECTION II ACCESS TO LAND			
7.	Does any part of your holding (i.e. total land for agricultural production) belong to the following categories of tenure? (a) Rented (b) Squatter (c) Family land (d) Others (specify) ……	 Yes …….. No …… Yes …….. No …… Yes …….. No …… Yes …….. No ……	 1 2 1 2 1 2 1 2	
8.	Which of the following land types do you use for your farm? (CIRCLE ALL APPLICABLE) (a) Upland (Rainfed) (b) Lowland (Swampy) (c) Irrigated	 Yes …….. No …… Yes …….. No …… Yes …….. No ……	 1 2 1 2 1 2	
	SECTION III ACCESS TO CREDIT			
9.	In running this holding has holder taken any credit (or borrowed money) this agricultural season?	Yes …….. No ……	1 2	11
10.	What is the source of fund?	Friends/ Relatives Community/ Peoples Bank Agric. Credit Bank Other Commercial Bank Cooperative Society Local Money Lender Traditional contribution (Esusu, etc.)	1 2 3 4 5 6 7	

<table>
<tr><th>QUES S/NO.</th><th>QUESTION AND FILTER</th><th>RESPONSE</th><th>CODE</th><th>SKIP TO</th></tr>
<tr><td colspan="5">SECTION IV CROP FARMING</td></tr>
<tr><td>11.</td><td>How many crop farm(s) have you already this agricultural season?</td><td colspan="2">Number: ……..</td><td></td></tr>
<tr><td>12.</td><td colspan="4">What crop(s) (including tree crop(s) do you already have on your farm(s) this agricultural season? Please tell me the total area planted to each crop in your local unit. Do not include scattered crops that are not planted in a regular pattern.
AGAINST EACH NAMED CROP(S) ENSURE THAT YOU RECORD THE TOTAL AREA PLANTED (IN LOCAL UNITS) AND THE NAME OF THE LOCAL UNIT
* FOR EACH TYPE OF LOCAL UNIT THE SUPERVISOR SHOULD SUPPLY THE NO. OF LOCAL UNITS MAKING A HECTARE UNDER THE “FACTOR” COLUMN.
<table>
<tr><th>Farm No.</th><th>Name of crop</th><th colspan="2">Area in Local Unit</th></tr>
<tr><td></td><td></td><td>No.</td><td>Name of L/Unit</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
</table>
<table>
<tr><th>Factor</th><th>Crop Code</th><th>Area in Hectare</th></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
</table></td></tr>
<tr><td>13.</td><td>Apart from the farms you are already operating, do you still intend to operate more this agricultural season?</td><td>Yes …….. 1
No …… 2</td><td></td><td>16</td></tr>
<tr><td>14.</td><td>How many farms do you still intend to operate this agricultural season?</td><td colspan="2">Number: ……..</td><td></td></tr>
<tr><td>15.</td><td colspan="4">What crop(s) (including tree crop(s) do you still intend to plant this agricultural season? Please tell me the total area you intend to plant to each crop in your local unit. Do not include scattered crops that are not planted in a regular pattern.
AGAINST EACH NAMED CROP(S) ENSURE THAT YOU RECORD THE TOTAL AREA PLANTED (IN LOCAL UNITS) AND THE NAME OF THE LOCAL UNIT
* FOR EACH TYPE OF LOCAL UNIT THE SUPERVISOR SHOULD SUPPLY THE NO. OF LOCAL UNITS MAKING A HECTARE UNDER THE “FACTOR” COLUMN.
<table>
<tr><th>Farm No.</th><th>Name of crop</th><th colspan="2">Area in Local Unit</th></tr>
<tr><td></td><td></td><td>No.</td><td>Name of L/Unit</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
</table>
<table>
<tr><th>Factor</th><th>Crop Code</th><th>Area in Hectare</th></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
</table></td></tr>
</table>

QUES S/NO.	QUESTION AND FILTER	RESPONSE	CODE	SKIP TO
16.	Do you use the following implement for ploughing?			
	(a) Hoe/Cutlass	Yes ……..	1	
		No ……	2	
	(b) Animal Drawn Plough	Yes ……..	1	
		No ……	2	
	(c) Motorized plough	Yes ……..	1	
		No ……	2	
	(d) Others (specify)	Yes ……..	1	
		No ……	2	
17.	Do you use the following implement for Harvesting?			
	(a) Hoe/Cutlass	Yes ……..	1	
		No ……	2	
	(b) Combined Harvester	Yes ……..	1	
		No ……	2	
	(c) Others (specify)	Yes ……..	1	
		No ……	2	
18.	Which of the following is your main source of ploughing and harvesting implement?	Government	1	
		Open market	2	
		Cooperative	3	
		Others (Specify)		
19.	Do you use the following transportation means for your farm produce?			
	(a) Truck/Pickups	Yes ……..	1	
		No ……	2	
	(b) Motor cycle	Yes ……..	1	
		No ……	2	
	(c) Bicycle	Yes ……..	1	
		No ……	2	
	(d) Boats	Yes ……..	1	
		No ……	2	
	(e) Donkey	Yes ……..	1	
		No ……	2	
	(f) Head Carrier	Yes ……..	1	
		No ……	2	
	(g) Others (specify)	Yes ……..	1	
		No ……	2	
20.	Do you sell any part of your farm produce?	Yes ……..	1	
		No ……	2	
21.	How many people regularly work on your holding(s)?	Holder……	**M**	**F**
		Unpaid HH Member		
		Paid worker ……		
	(RECORD THE NUMBER OF MALE (M) OR FEMALE (F) ACCORDINGLY)			

QUES S/NO.	QUESTION AND FILTER	RESPONSE	CODE	SKIP TO
	SECTION V USE OF FERTILIZER			
22.	Have you used fertilizer on any of your farms this agricultural season?	Yes …….. No ……	1 2	 26
23.	What types of fertilizer have you used? Chemical Fertilizer Farm Manure	 Yes …….. No …… Yes …….. No ……	 1 2 1 2	
24.	What is your main source of supply of chemical fertilizer (PLEASE TICK ONE OPTION)	Min. (Ext. service) River Basin Authority Agro. Serv. Centre Farm Serv. Centre Cooperative Society Local Market Others (Specify)……..	1 2 3 4 5 6	
25.	How far do you usually go to obtain this chemical fertilizer?	Within the locality Outside locality - but les than 10km - more than 10km but less than 100km - more than 100km	1 2 3 4	27 27 27 27
26.	What is your main reason for not using fertilizer? (PLEASE TICK ONE OPTION)	Doubt its effectiveness Too costly to obtain Too far to obtain Don't know where to obtain it Never heard of it Don't need it Others (Specify)	1 2 3 4 5 6	
	SECTION VI USE OF PESTICIDES/INSECTICIDES			
27.	Have you used Pesticides on any farm(s) this agricultural season?	Yes …….. No ……	1 2	 30
28.	What is your main source of supply? (PLEASE TICK ONE OPTION)	Min. (Ext. Services) River Basin Authority Agro. Service Centre Farm Service Centre Cooperative Society Local Market Others (Specify)……..	1 2 3 4 5 6	
29.	How far do you usually go to obtain the Pesticides?	Within the locality Outside locality - but less than 10km - more than 10km but less than 100km - more than 100km	1 2 3 4	31 31 31 31

QUES S/NO.	QUESTION AND FILTER	RESPONSE	CODE	SKIP TO
30.	What is your main reason for not using Pesticides? (PLEASE TICK ONE OPTION)	Doubt its effectiveness Too costly to obtain Too far to obtain Don't know where to obtain it Never heard of it Don't need it Others (Specify)……..	1 2 3 4 5 6	
	SECTION VII USE OF IMPROVED SEEDLING			
31.	Have you used improved seedling on any of your farm(s) this agricultural season?	Yes …….. No ……	1 2	
32.	What is your main source of supply? (PLEASE TICK ONE OPTION)	Min. (Ext. Services) River Basin Authority Agro. Service Centre Farm Service Centre Cooperative Society Local Market Others (Specify)……..	1 2 3 4 5 6	
33.	How far do you usually go to obtain the Improved Seedling?	Within the locality Outside locality - but less than 10km - more than 10km but less than 100km - more than 100km	1 2 3 4	35
34.	If you are using improved seedling, indicate which of these reasons is mainly responsible. (PLEASE TICK ONE OPTION)	Doubt its effectiveness Too costly to obtain Too far to obtain Don't know where to obtain it Never heard of it Don't need it Others (Specify)……..	1 2 3 4 5 6	
	SECTION VIII LIVESTOCK AND POULTRY			
35.	Do you keep any Livestock which are owned by holder's household?	Yes …….. No ……	1 2	
36.	Do you keep any Livestock which belongs to members of other household?	Yes …….. No ……	1 2	

QUES S/NO.	QUESTION AND FILTER	RESPONSE	CODE	SKIP TO
37.	Give the total number of each type of livestock kept in this household			
		Type of Livestock	**Code**	**Quanity**
				M F
38.	Do you vaccinate your Animals?	Yes …….. No ……	1 2	 40
39.	Who does the Vaccination?	Government Private Vet. Others (Specify)…..	1 2 3	
40.	Does your Household keep any poultry?	Yes …….. No ……	1 2	 44
41	If any poultry is kept in this holding, indicate the total number of each kept.			
		Type of Poultry	**Code**	**Quanity**
				M F
42.	Do you vaccinate your Birds?	Yes …….. No ……	1 2	 44
43.	Which of the following is your main source of supply of vaccine?	Government Open Market Cooperative Private Vet. Others (Specify)…….	1 2 3 4	
44.	Do you use veterinary Drugs on your Animals/Birds apart from vaccination?	Yes …….. No ……	1 2	 46
45.	Which of the following is your main source of supply of Veterinary drugs?	Government Open Market Cooperative Private Vet. Others (Specify)…….	1 2 3 4	
46.	Do you feed your livestock/birds with supplementary ration?	Yes …….. No ……	1 2	 48
47.	Which of the following is your main source of supply of Supplementary feed?	Government Open Market Cooperative Private Vet. Others (Specify)…….	1 2 3 4	

QUES S/NO.	QUESTION AND FILTER	RESPONSE	CODE	SKIP TO
	SECTION IX ANCILLARY ACTIVITIES			
48.	Is any fish farming carried out in this holding?	Yes …….. No ……	1 2	52
49.	What type of fish farming does this holding engaged in?	Marine Fishing Fresh Water (i.e. Rivers) Lake Fishing Pond	1 2 3 4	
50.	How many of the following items does the holding possess? (READ OUT THE ALTERNATIVE TO THE INFORMANT AND FILL IN THE QUANTITY SUPPLIED)	Motorized Boat Boats not motorized Nets – Long Nets Round Hooks Others (Specify)		
51.	Does this holding normally operate its boat(s) under the sole control of the holder or in cooperation with other persons?	Sole …… Joint Operation …….	1 2	
52.	Do you have forest trees on your farm(s)?	Yes …….. No ……	1 2	
53.	Indicate the use(s) to which you put the forest product? (CIRCLE ALL APPLICABLE) (a) Fire wood for fuel (b) Foreage (c) Industries (d) Protection (e) Other purposes	 Yes …….. No …… Yes …….. No …… Yes …….. No …… Yes …….. No …… Yes …….. No ……	 1 2 1 2 1 2 1 2 1 2	

QUES S/NO.	QUESTION AND FILTER	RESPONSE	CODE	SKIP TO
54.	Are you engaged in the following forestry industries? (CIRCLE ALL APPLICABLE)			
	(a) Pulp making	Yes …….. No ……	1 2	
	(b) Ply wood	Yes …….. No ……	1 2	
	(c) Lumbering	Yes …….. No ……	1 2	
	(d) Others (Specify)	Yes …….. No ……	1 2	
55.	**STOP INTERVIEW**			

ENUMERATOR	SUPERVISOR
Name: _____	Name: _____
Sign: _____	Sign: _____
Date: _____	Date: _____

Reconstruction des tendances de la mortalité des jeunes enfants en Afrique sub-saharienne de 1950 à 1999 à partir des données d'enquêtes démographiques

Michel Garenne[1] et Enéas Gakusi[2]

Résumé

L'étude propose une reconstruction des tendances de la mortalité des enfants de moins de cinq ans dans trente deux pays sur la période 1950-1999. Elle utilise les données d'enquêtes démographiques de l'enquête mondiale sur la fécondité (EMF) et des enquêtes démographiques et sanitaires (EDS). La méthode utilisée combine les données des différentes enquêtes pour un même pays et permet de recalculer les taux de mortalité par année. Les inversions de tendances sont testées au moyen d'un modèle linéaire logistique. Sept pays ont une évolution monotone ou quasi monotone, huit connaissent une inversion majeure et les autres montrent une ou plusieurs périodes d'inversion mineure de la mortalité. La transition sanitaire est en bonne voie en Afrique mais peu rapide par rapport au reste du monde et sujette à des inversions assez fréquentes, liées soit à des crises épidémiologiques (sida, paludisme) soit le plus souvent à des crises économiques et politiques.

Mots clé

Mortalité des jeunes enfants, tendances, transition sanitaire, Afrique sub-saharienne, crises économiques et politiques, crises épidémiologiques.

Summary

The study proposes a reconstruction of trends in under-five mortality in thirty-two African countries over the 1950-1999 period. It used demographic data from the World Fertility Surveys (WFS) and the Demographic Health Surveys (DHS). Data of different surveys for the same country were combined, and allowed one to estimate yearly mortality rates. Trends reversals were tested using a linear-logistic model. Seven

1: Directeur de Recherche, Institut Pasteur Paris, mgarenne@pasteur.fr
2: Chargé d'Evaluation Principal, Banque Africaine de Développement, a.gakusi@afdb.org
Les auteurs tiennent à remercier tout spécialement le Wellcome Trust qui a financé une partie de cette recherche.

countries had a monotonous or quasi-monotonous evolution, eight experienced major reversals, and the remaining countries showed one or several periods of minor reversals in mortality trends. The health transition is progressing in Africa, though at a slower pace compared to the rest of the world. Frequent reversals were observed, due to either epidemiological crises (HIV/AIDS, malaria) or more often to economic and political crises.

Key words

Under-five mortality, trends, health transition, sub-Saharan Africa, political and economic crises, epidemiological crises.

1. Introduction

La transition sanitaire, définie comme une baisse régulière et importante de la mortalité, a été l'un des phénomènes marquants des dynamiques démographiques du 20ème siècle. Ce mouvement a commencé dès la seconde partie du 19ème siècle en Europe occidentale et dans les pays de nouveau peuplement européen comme les Etats Unis, le Canada, l'Australie et la Nouvelle Zélande. Il s'est répandu assez rapidement dans tous les pays du monde, et dès le milieu du siècle, les données disponibles faisaient état d'une amélioration de l'espérance de vie quasi universelle (Stolnitz, 1955; Stolnitz, 1965). Les données postérieures n'ont fait que confirmer ces premières analyses. Jusqu'à l'avènement du sida, ces tendances semblaient devoir se prolonger. La baisse de la mortalité est concomitante à l'émergence des États modernes et à la mise en place de politiques de santé publique fondées sur les nouvelles connaissances biologiques et médicales. Elle est aussi consécutive à l'amélioration des conditions de vie, en particulier la nutrition et le logement. Les recettes de la santé publique ont été appliquées en quelques décennies dans le monde entier, du Japon à la Terre de Feu, de la Russie aux îles du Pacifique. Elles ont pour premiers objectifs l'eau potable, l'assainissement, l'hygiène, les vaccinations, l'alimentation des enfants. Elles se poursuivent avec la mise au point et la distribution des médicaments modernes ainsi qu'avec l'application de diverses actions de médecine préventives et curatives. Toutes ces actions se sont traduites par une baisse spectaculaire de la mortalité des jeunes enfants et des jeunes adultes, y compris la mortalité maternelle. Cette baisse était surtout due à la diminution de la mortalité par maladies infectieuses et celle des autres causes de décès, dont certaines étaient elles même des conséquences lointaines des infections antérieures (Preston, 1976; Preston, 1980; McKeown, 1976; Schoffield and Reher, 1991).

Le continent qui a le plus tardé à mettre en place ces politiques est sans doute le contient africain, tout particulièrement l'Afrique sub-Saharienne continentale et Madagascar. Les pays de cette région sont ceux qui ont encore les plus forts taux de mortalité des jeunes enfants et les plus faibles espérances de vie. Certes, certains pays d'Asie ne sont guère mieux lotis (Afghanistan, Yémen), mais il s'agit d'exceptions, alors que les pays africains à faible mortalité font plutôt figure d'exception. Cependant, pour juger de l'évolution des conditions sanitaires en Afrique, il convient d'analyser plus en détail les tendances de la mortalité, plutôt que de considérer seulement le niveau actuel en soi. En effet, si la transition sanitaire a démarré plus tard en Afrique, un niveau élevé de mortalité serait plutôt dû au départ tardif qu'à une évolution récente peu favorable. Par contre, montrer que les tendances de la mortalité sont plus faibles qu'ailleurs, voire inversées (remontée de la mortalité) a une toute autre résonance, car cela montre une évolution de nature différente. Cet aspect est particulièrement sensible pour ce qui concerne la mortalité des jeunes enfants, un des principaux indicateurs de la transition sanitaire.

Pour juger des tendances de la mortalité des jeunes enfants en Afrique on ne dispose pas des séries annuelles – comme c'est le cas en Europe, dans les autres pays développés ou en Amérique Latine - de l'état civil provenant de l'enregistrement de routine des naissances et des décès. En effet, aucun pays d'Afrique sub-saharienne continentale ne dispose d'un état civil complet pour la période 1950-1999. Il convient donc de reconstruire les tendances de la mortalité à partir des données disponibles. Ainsi, les annuaires des Nations Unies et différentes publications font état des données d'enquêtes et de diverses estimations directes ou indirectes (Nations Unies, 1982; Brass et al. 1968; Feachem and Jamison, 1991 ; Mandjale, 1985; van de Walle et al., 1988). Une synthèse des données indirectes africaines a été réalisée au début des années 1980 (Hill, 1991 ; Hill and Hill ; 1988). Elle recouvre des périodes très variables selon les pays, démarrant entre 1926 (Angola) et 1966 (Cote d'Ivoire) selon les pays, jusqu'au dernier point disponible en 1985. Cette synthèse souffre cependant de l'imprécision des méthodes indirectes. Une synthèse plus récente a été réalisée sous l'égide de l'OMS (Ahmad et al. 2000) par période quinquennale, et couvre l'ensemble des périodes de 1955-1959 à 1995-1999. Cette synthèse utilise mieux toutes les données disponibles, notamment les données directes provenant des enquêtes par sondage, mais elle ne couvre que les périodes quinquennales, ce qui gomme les changements de tendance de quelques années, et empêche les analyses annuelles plus fines.

Le but de cette étude est de réaliser une analyse fine des tendances de la mortalité des jeunes enfants dans certains pays africains disposant d'enquêtes démographiques. Il s'agit surtout d'identifier le cas échéant les périodes d'inversion de tendance.

Cette étude est un préalable à celle plus détaillée portant sur des situations de remontée de la mortalité ainsi sur leurs déterminants sociopolitiques, économiques et épidémiologiques. Une attention particulière a été portée sur l'impact du sida pédiatrique, devenu une importante cause de décès depuis le milieu des années 1980.

2. Données et Méthodes

Les données utilisées pour cette reconstruction des tendances sont celles recueillies dans les grandes enquêtes démographiques qui font partie des programmes internationaux de l'Enquête Mondiale de Fécondité (EMF, ou WFS selon l'acronyme anglais) et des Enquêtes Démographiques et Sanitaires (EDS, ou DHS selon l'acronyme anglais). Ces enquêtes sont toutes basées sur la même méthodologie: le recueil des histoires des maternités auprès d'un échantillon représentatif de femmes de 15-49 ans. Pour chaque femme enquêtée, dont le nombre moyen est de 6300 par enquête[3], on dispose de la date de naissance de chaque enfant, en mois et année et, le cas échéant, de l'âge au décès, en général en mois, avec le détail en jours pour les nouveau-nés, et parfois seulement l'âge en année pour les enfants plus âgés. Ces données permettent donc de recalculer les taux de mortalité par âge et période. Au total 59 enquêtes ont été utilisées dans cette étude, couvrant 32 pays d'Afrique subsaharienne, soit 70% des pays et 80% de la population d'Afrique sub-saharienne continentale et Madagascar, en excluant les petites îles (Cap-Vert, Sao Tome et Principe, Comores, Seychelles, Mayotte, Réunion, et Sainte-Hélène). À cet ensemble on a ajouté une enquête de type MICS, basée sur des histoires des trois dernières grossesses, réalisée en Angola, et une enquête DHS couplée à d'autres données en Afrique du Sud. Les pays et enquêtes retenus pour l'analyse figurent dans l'annexe.

Les données des histoires des maternités représentent un ensemble imparfaitement représentatif des naissances et décès de la période antérieure. En effet, plus on remonte dans le temps, plus l'âge des mères au moment de la naissance est jeune, puisque les femmes plus âgées ne figurent pas dans l'échantillon. Ceci a tendance à surestimer un peu la mortalité des enfants, car la mortalité des enfants des très jeunes femmes est supérieure à la moyenne. Cependant, les enfants des mères décédées avant l'enquête ne sont pas représentés, ce qui a tendance à sous-estimer la mortalité du fait de la forte corrélation entre mortalité maternelle et mortalité des orphelins. En pratique, on a considéré que ces deux biais se compensaient approximativement, et que les estimations brutes étaient suffisantes. Cette approche se trou-

3: La taille des enquêtes varie entre 3040 (Cote d'Ivoire 1999) et 15367 (Ethiopie 2000).

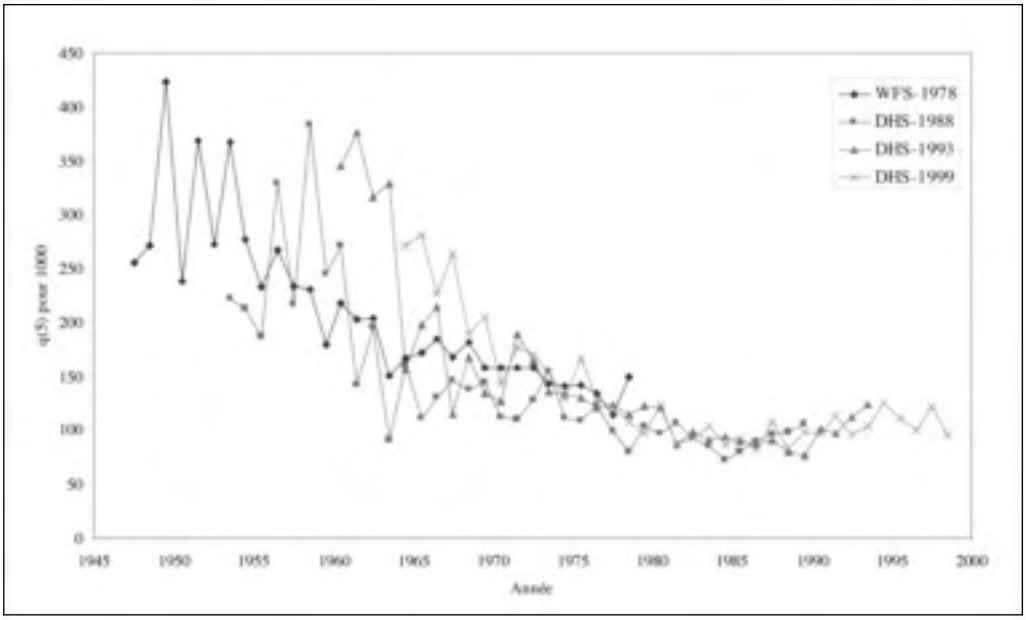
vera justifiée par la compatibilité des tendances entre les enquêtes d'un même pays, quelle que soit la distance à la date d'enquête.

On détaille ci-dessous les différentes étapes des calculs permettant d'établir les tendances avec et sans sida.

2.1 Calcul des taux de mortalité par année de calendrier, par enquête

La première étape consiste à calculer les taux de mortalité par âge et période dans chaque enquête. Dans la première étape, la période a toujours été l'année de calendrier. Dans le cas de l'Ethiopie, le calendrier amharique a été converti en calendrier grégorien. Deux groupes d'âge ont été considérés: 0-1 ans et 1-5 ans, afin de permettre d'affiner les calculs de la mortalité infanto-juvénile. Utiliser toutes les années d'âge entre 0 et 5 ans aurait introduit des biais supplémentaires du fait du petit nombre de décès à 1, 2, 3 et 4 ans. Pour les calculs pratiques, les dates de naissance ont été prises au 1er de chaque mois. Les dates de décès sont calculées soit en prenant le nombre de jours exacts pour les nouveau-nés, soit en prenant le 15 du mois pour les décès postérieurs à un mois. Pour les rares enfants dont l'âge au décès est estimé en années, on a considéré le nombre d'années plus six mois, et pour ceux encore plus rares pour lesquels l'âge au décès est inconnu on a considéré l'âge moyen au décès dans le pays (approximativement 24 mois). Lorsque les dates de naissances et de décès sont recalculées, on peut compter les personne-années vécues et les décès par âge et période, et en déduire les taux de mortalité. Lorsque l'échantillon stratifié était pondéré, on a appliqué les pondérations de la mère à chaque enfant. Ceci a été réalisé par un programme écrit en langage Dbase/FoxPro. À l'issue de la première étape on dispose donc des taux de mortalité infantile (< 1 an) et juvénile (1-4 ans) par année de calendrier, qui sont convertis en quotients en prenant comme âge moyen au décès 0,33 pour les moins d'un an et 1,35 pour les 1-4 ans. Ces valeurs ont été tirées des tables-type de mortalité des Nations Unies, modèle général pour les pays en développement. Ces quotients permettent d'en déduire le quotient de mortalité infanto-juvénile, noté q(5) en appliquant la formule classique du produit des probabilités de survie: $(5) = (1- {}_1q_0)*(1-{}_4q_1)$. La figure 1 donne l'exemple de ces calculs pour le Kenya.

Figure 1 : Evolution des taux de mortalité des jeunes enfants au Kenya selon les enquêtes EMF et EDS



La figure 1 montre d'abord la bonne concordance des enquêtes, en niveau et en tendance, qui permet l'agrégation. Les valeurs pour les périodes les plus anciennes de chaque enquête sont affectées par de fortes fluctuations aléatoires, le nombre de décès étant inférieur à 50 pour les premières années, ainsi que par les biais divers mentionnés ci-dessus. Par contre on ne note pratiquement pas de différence significative entre deux enquêtes successives, et toutes montrent une baisse régulière de la mortalité avant 1984, puis une hausse pour celles qui couvrent la période récente. Ce sont ces inversions de tendance qui sont analysées plus particulièrement par la suite.

2.2 Agrégation des différentes enquêtes

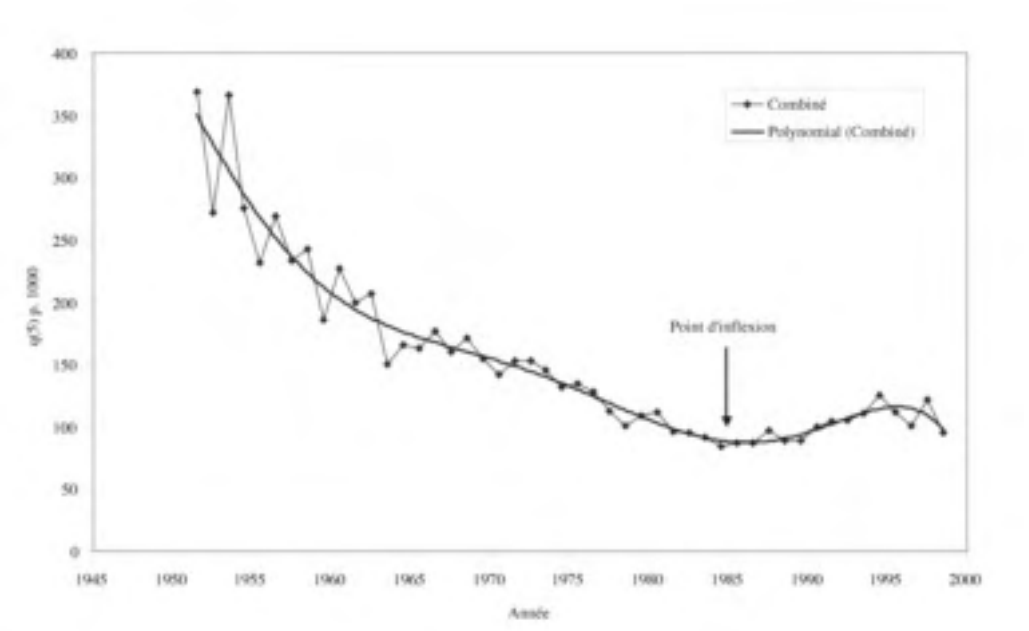
La seconde étape a consisté à agréger, le cas échéant, les différentes enquêtes d'un même pays. Cela a été fait en ajoutant, pour la même année et par groupe d'âge, les personnes-années vécues et les décès, de manière à recalculer les taux. Auparavant, on s'est assuré que les différentes enquêtes d'un même pays étaient compatibles, en plaçant dans un graphique les estimations de la mortalité infanto-juvénile par année. Aucune différence notable n'a été remarquée, et pratiquement toutes les

données ont été jugées compatibles, à l'exception du Nigeria, ce qui montre le haut niveau de consistance des enquêtes démographiques de type EMF/EDS. L'agrégation des différentes enquêtes d'un même pays repose sur le principe de conservation du nombre de décès observés. Ceux-ci s'ajoutent par âge et période, et le nombre de décès relatifs à l'estimation agrégée est la somme des décès observés dans les différentes enquêtes pour le même âge et la même période. Ceci permet de faire les calculs de tests statistiques et de tendances, qui ne dépendent que du nombre de décès observés et de la valeur de la mortalité infanto-juvénile.

2.3 Recherche des périodes de variation monotone

La troisième étape a consisté à rechercher les points d'inflexion (changement de pente) afin d'identifier les périodes de variations monotones de la mortalité infanto-juvénile. Ceci a été réalisé d'abord graphiquement en ajustant à la courbe des données agrégées par un polynôme de degré élevé (4 à 6). Chaque fois qu'une inversion de tendance a été relevée, un test statistique de Student a été appliqué pour savoir si la différence entre les valeurs observées au cours d'une période et la tendance précédente était significative. En particulier, les différences en début de période, où le nombre de décès est faible du fait de la troncature des histoires des maternités et non significatives, ont été ignorées. L'exemple de la figure 2 révèle les deux périodes monotones de l'évolution de la mortalité infanto-juvénile au Kenya : avant 1984, où elle baisse régulièrement, et après 1984 où elle remonte. Les différences par rapport à la tendance sont non significatives, hormis les points de 1954 et de 1964, et dues seulement aux variations d'échantillonnage.

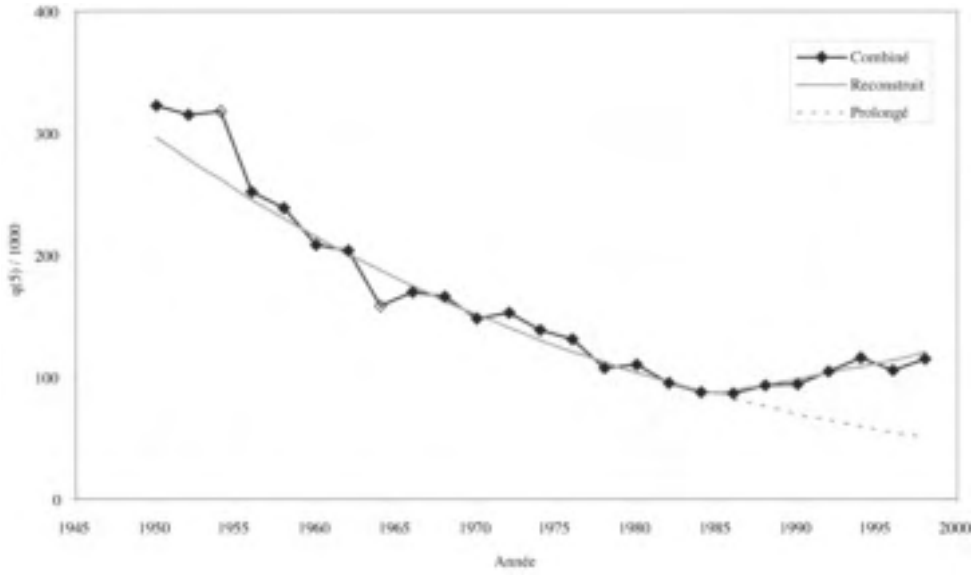
Figure 2 : Résultat de l'agrégation des différentes enquêtes et ajustement polynomial, Kenya.



2.4 Estimations de la mortalité par période de deux ans

Les données annuelles ont été utilisées pour assurer le maximum de précision pour les tendances, et pour identifier les points d'inflexion à l'année près, ce qui est important pour l'analyse des causes et circonstances des inversions de mortalité. Cependant, les données annuelles sont assez fortement affectées par les variations aléatoires, produisant des différences apparentes mais statistiquement non significatives. Pour la présentation des estimations brutes on a donc regroupé les données par période de deux ans, simplement en ajoutant les numérateurs (décès) et les dénominateurs (naissances). Les estimations par deux ans sont remarquablement stables dans les pays ayant plusieurs enquêtes et ceux dans lesquels le nombre de décès est important, comme le montre l'exemple du Kenya (figure 3).

Figure 3 : Mortalité par période de deux ans, et ajustement par modèle logit, Kenya



2.5 Calcul des tendances par période monotone / test des tendances

Une fois les périodes monotones identifiées, on peut recalculer les tendances par période. Ceci a été réalisé en partant du principe de l'équivalence période / cohorte. En effet, l'estimation de la mortalité infanto-juvénile pour une période donnée est équivalente à la probabilité de décéder entre 0 et 5 ans pour une cohorte fictive. On a donc considéré que les estimations successives de la mortalité par période sont équivalentes à des probabilités de décès entre 0 et 5 ans pour le même nombre de décès, et pour un nombre équivalent de naissances égal au rapport des décès au quotient de mortalité (N = D / q). Ceci permet donc de reconstruire des cohortes fictives de naissances et de décès par année, et d'utiliser un modèle linéaire logistique (Logit) pour calculer la tendance et son intervalle de confiance, en tenant compte de toutes les observations individuelles. On appelle « pente logistique » la pente de la régression du logit du quotient de mortalité sur la date en année; cette pente fournit une estimation de la vitesse de la baisse moyenne annuelle de la mortalité au cours de la période. Tous ces calculs ont été faits avec le module correspondant du logiciel SPSS. Cet ajustement permet en outre de faire un certain nombre de calculs supplémentaires : comparaison des pentes de la baisse de la mortalité; comparaison des valeurs observées et valeurs prédites par le modèle pour identifier des accidents de forte ou faible mortalité par rapport à la tendance, qui sont marqués par un point transparent sur le graphique; et prolongement des tendances précédentes pour es-

timer le parcours de la transition sanitaire en l'absence de crise. La figure 3 montre par exemple que seuls les deux points de 1954 et 1964 sont significativement différents de la tendance, ce qui est sans importance pour l'interprétation.

2.6 Séroprévalence VIH et impact attendu du sida pédiatrique

Le sida pédiatrique est devenu une cause majeure de mortalité dans de nombreux pays africains. Cette cause de décès s'ajoute aux causes classiques de décès des jeunes enfants, et peut à elle seule inverser les tendances de la mortalité toutes choses égales par ailleurs dans les pays à forte prévalence du sida. Il convient donc de rechercher les niveaux et tendances de mortalité en l'absence du sida, en soustrayant cette cause. Pour ceci on dispose des données sur la séroprévalence du VIH chez les femmes enceintes se rendant aux visites prénatales. Ces données sont les plus proches pour estimer l'impact du sida pédiatrique, puisque la plupart de ces cas proviennent de la transmission materno-fœtale.

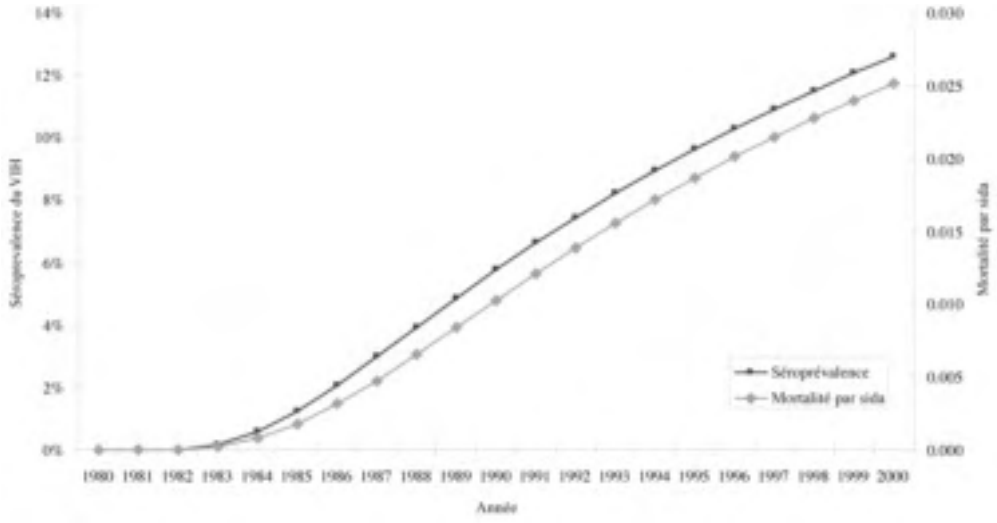
Pour estimer la mortalité par sida pédiatrique, on a d'abord reconstitué les tendances de la séroprévalence du VIH chez les femmes enceintes. Pour de nombreux pays on dispose de plusieurs points au cours de la période 1985-1997, qui sont fournis par la banque de données sur le sida tenue à jour par l'office du recensement américain *(US Bureau of Census / International Programs).*

Ces données sont souvent incomplètes et on ne dispose en général que de quelques points entre 1985 et 1997. Dans les pays où le cours de l'épidémie est naturel et en phase ascendante, on a ajusté un modèle simple mis au point sur les données de séroprévalence d'Afrique du Sud, la seule série complète pour la période 1990-2000. Cet ajustement se fait en deux temps: on transforme les proportions de femmes séropositives en logit, et la série des logit est ajustée par une fonction hyperbolique. Cet ajustement marche bien dans tous les pays en phase ascendante pour lesquels la série est consistante. Pour les pays où la séroprévalence diminue (Ouganda, RDC), on a d'abord ajusté la même fonction pour la période ascendante avant le maximum, puis une fonction linéaire-logistique pour la période descendante.

Cette série de séroprévalence chez les femmes enceintes permet de calculer l'impact attendu du sida pédiatrique. On a supposé une transmission materno-fœtale constante, et égale à 25%, sans tenir compte de la durée d'allaitement maternel. Pour les enfants infectés, on a supposé que 30% décédaient chaque année entre 0 et 5 ans, sans tenir compte du niveau de mortalité dans le pays. Ces paramètres

pourront être ajustés lorsque plus de données seront disponibles sur la mortalité par sida en Afrique. Ils correspondent à des estimations moyennes tirées de diverses études (Newell et al. 2004). Les résultats montrent une étroite corrélation entre les pourcentages de femmes enceintes séropositives et la mortalité par sida pédiatrique dans le cas du Kenya ou l'augmentation de la prévalence du sida est régulière entre 1980 et 1999 (figure 4). Dans les cas d'évolution complexe, avec augmentation suivi de diminution, comme en Ouganda, l'impact du sida pédiatrique est quelque peu amorti.

Figure 4 : Estimation de la mortalité par sida pédiatrique en fonction de la séroprévalence, Kenya

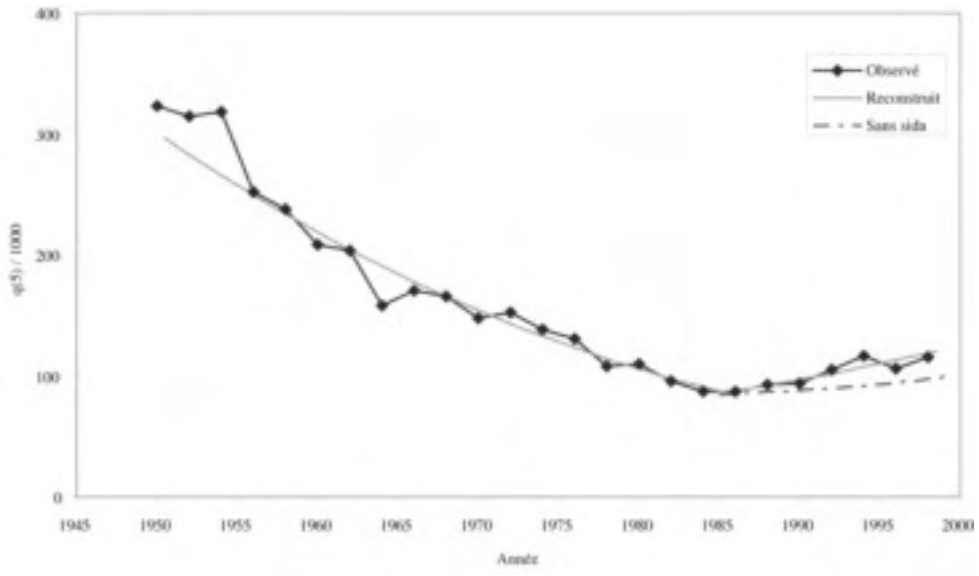


2.7 Tendances sans sida

Pour calculer les tendances sans sida, on a simplement supposé que les causes de décès s'ajoutent, donc que la mortalité sans sida est égale à la mortalité totale moins la mortalité attribuable au sida. Dans certains pays les données sur le sida sont de médiocre qualité, et dans d'autres la mortalité par sida ne représente qu'une faible partie de la mortalité totale, soit que celle-ci soit élevée, soit que la séroprévalence du VIH soit faible. Quand la mortalité par sida représentait moins de 10 % de la mortalité des moins de 5 ans on peut ignorer la correction. Les calculs de tendan-

ces sans sida ont donc été effectués en particulier pour tous les pays à forte prévalence, c'est à dire tous les pays d'Afrique Orientale et Australe, ainsi que dans trois pays d'Afrique Occidentale: Côte d'Ivoire, Mali et Burkina Faso. Ces calculs permettent aussi d'estimer si le sida explique tout ou partie de l'augmentation de mortalité, ou le cas échéant quelle serait la valeur de la mortalité sans sida.

Figure 5 : Estimation des tendances de la mortalité sans sida, Kenya



2.8 Comparaison avec d'autres données

Les estimations de mortalité, avant la correction sur le sida, ont été comparées aux tendances reconstruites à partir des données indirectes (Hill, 1991) et à la synthèse récente publiée par l'OMS (Ahmad et al. 2000), ainsi qu'aux estimations brutes fournies par les DHS et par d'autres enquêtes démographiques disponibles, en particulier les enquêtes réalisées par l'INSEE dans les années 1960 et 1970. Ces comparaisons permettent tout à la fois de valider les reconstructions, et de mesurer l'imprécision de certaines estimations indirectes. L'exemple du Kenya montre la bonne concordance entre les différentes estimations (figure 6).

Figure 6 : Comparaison avec les estimations indirectes et les autres données disponibles, Kenya.

Les estimations indirectes, qui proviennent des recensements de 1948, 1960 et 1970 sont remarquablement proches de la série reconstruite entre 1950 et 1980, et les données antérieures à 1950 sont pratiquement dans le prolongement. Les niveaux moyens des points situés entre 1948 et 1978 sont identiques (210 et 208 pour 1000), et les estimations indirectes tombent dans une fourchette de +/- 20% (+19% au recensement de 1970 et -19% à celui de 1948) par rapport à la tendance, ce qui donne une valeur de leur imprécision. Les tendances sont aussi approximativement les mêmes : les pentes logistiques sont respectivement de -4,3% et -3,8% entre 1948 et 1978. Les estimations directes sont pratiquement identiques à la série reconstruite, ce qui était attendu puisqu'il s'agit des mêmes données, avec un groupement d'années différent. Par contre la série reconstruite remonte beaucoup plus loin en arrière et, en utilisant des périodes de deux ans, au lieu de cinq ans, permet une meilleure estimation du point d'inflexion. La série de l'OMS, qui fait une synthèse entre les données directes et indirectes est moins précise : elle gomme le point d'inflexion, néglige l'augmentation de la mortalité après 1985, et sous-estime la pente entre 1955 et 1985 -2,6% au lieu de -4,3%, et sous-estime la mortalité en 1998. La série reconstruite sans sida indique que même sans le sida la mortalité n'aurait pas baissé au Kenya après 1985. En prolongeant la tendance précédente de 1950-1985, le niveau de mortalité attendu était de 51 en 1998 : on observe une surmor-

talité de +69 pour 1000, dont 22 pour 1000 est dû au sida et 47 pour 1000, soit environ les deux tiers, à d'autres causes. La reconstruction apparaît donc importante pour suivre l'évolution de la mortalité des jeunes enfants au Kenya et pour pouvoir en analyser les grandes causes.

3. Résultats

Les reconstructions concernent 32 pays mais avec une qualité de données inégale[4]. Les reconstructions commencent en général en 1950, sauf certains pays dont les données sont trop récentes et pour lesquels on ne dispose pas de données anciennes. Elles s'arrêtent au dernier point disponible, c'est-à-dire à la date de la dernière enquête, car compte tenu de la grande diversité d'évolutions récentes il est impossible de prévoir l'évolution future à partir des données antérieures. Ces premières analyses, qui seront suivies d'analyses plus poussées dans certains cas, permettent une typologie des tendances de la mortalité infanto-juvénile.

3.1 Tendances régulières

Parmi les 32 pays étudiés, sept ont connu une transition sanitaire régulière depuis 1950, c'est-à-dire une baisse pratiquement continue de la mortalité des jeunes enfants au cours des années pour lesquelles on dispose de données. Il s'agit des pays suivants : le Burkina-Faso, l'Ethiopie, la Guinée, le Libéria (jusqu'en 1986), le Mali, le Tchad, et le Togo. Ces pays avaient tous des niveaux élevés de mortalité en 1950, et n'ont connu que peu d'incident de parcours. Notons que la situation du Libéria n'est pas connue après 1986, et que du fait de la guerre civile des années 1990 il est probable que l'évolution ait divergé au cours des années récentes.

3.2 Crises conjoncturelles

Certains pays ont connu des petits accidents de parcours, des crises conjoncturelles qui ne durent que quelques années et sont sans conséquence sur les tendances générales. Ainsi, au Burkina-Faso la mortalité remonte en 1994-1995 en milieu rural, mais pas en milieu urbain. Au Lesotho la mortalité augmente significativement en 1962-1965. En Namibie, la mortalité augmente au moment de la crise d'indépendan-

4: Les résultats par pays ainsi que les tendances reconstruites par la même technique pour les milieux urbain et rural ont été publiés sous forme de DHS Working Papers 2005 No. 26 à l'adresse suivante http://pdf.dec.org/pdf_docs/pnade460.pdf

ce (1977-1980), pour reprendre un cours normal par la suite, avant l'arrivée du sida. Il en va de même au Zimbabwe, aussi à la période d'indépendance (1978-1982). Au Nigeria, la mortalité remonte pendant la période de la guerre du Biafra (1964-1968). Au Tchad, la mortalité remonte en 1980-1981. Au Togo, la mortalité augmente en milieu rural en 1993-1994.

3.3 Inversions de tendances: événements politiques et économiques

Plusieurs pays ont connu des inversions de tendance pendant plusieurs années du fait d'une crise politique, d'une crise économique ou d'une guerre civile. Ainsi, en Angola l'évolution de la mortalité des jeunes enfants est favorable pendant la fin de la période coloniale et dans les années qui suivent l'indépendance (1975), mais s'inverse pendant la période de guerre civile des années 1980. Une évolution similaire s'observe au Mozambique, pratiquement aux mêmes périodes et pour la même raison (guerre civile). Au Ghana, la mortalité remonte nettement entre 1979 et 1983, période de crise politique sérieuse avant l'installation définitive de Jerry Rawlings. A Madagascar, la mortalité augmente pendant toute la période de la révolution Malagasy (1975-1986), période au cours de la quelle le PIB par tête va sérieusement diminuer. Au Nigeria, la mortalité remonte entre 1978 et 1988, surtout en milieu rural, alors qu'elle stagne en milieu urbain. En Ouganda, la mortalité remonte fortement pendant les années de présidence d'Idi-Amin (1971-1979) et les années qui ont suivi sa chute (1980-1984), jusqu'au retour à une situation politique stable avec l'arrivée au pouvoir de Museveni en 1986. Au Rwanda, comme au Burundi, la mortalité augmente dans la période qui suit l'indépendance, entre 1965 et 1976. Au Rwanda, la mortalité augmente à nouveau au moment de la guerre civile et du génocide (1994). En Tanzanie, la mortalité augmente au cours des années 1979-1985, et tout particulièrement en milieu urbain. En Zambie, la mortalité augmente entre 1975 et 1992, surtout en milieu urbain, suite à la chute des cours du cuivre sur les marchés internationaux en 1975, et la crise politique et économique qui a suivi. Cette hausse se vérifie même en faisant abstraction de l'impact du sida dans ce pays.

3.4 Stagnations dans les années 1980-1990

Plusieurs pays ont connu des périodes de stagnation de la mortalité dans les années 1980 et 1990. Ainsi, au Bénin, la mortalité cesse de baisser, en milieu urbain comme en milieu rural, entre 1979 et 1989 qui est cependant une période plutôt favorable au point de vue économique. En Centre-Afrique, la mortalité stagne entre

1977 et 1989, période défavorable sur le plan politique avec la succession de régimes ayant suivi l'essor puis la chute de Bokassa. Au Niger, la mortalité stagne en milieu rural entre 1972 et 1992. Au Soudan, la mortalité stagne au cours des années 1974-1984, et est même en légère augmentation en milieu rural.

3.5 Accident épidémiologique (paludisme)

Un seul cas, qualifié d'accident épidémiologique, est à noter. Il s'agit du Sénégal au cours de la période 1960-1970, où une faible remontée de la mortalité semble attribuable à une recrudescence du paludisme. Il faut noter que la mortalité par paludisme semble avoir augmenté à nouveau dans les années 1990, d'après des études ponctuelles, mais probablement sans grande incidence sur la mortalité générale.

3.6 Inversions de tendances dues au sida

Le sida a un fort impact démographique dans de nombreux pays, provoquant des inversions de tendance notables. Il s'agit des pays d'Afrique australe (Afrique du Sud, Botswana, Zimbabwe, Malawi, Zambie), de certains pays d'Afrique Orientale (Kenya, Ouganda, Tanzanie), et de certains pays d'Afrique Occidentale (Côte d'Ivoire). Au Zimbabwe, toute l'augmentation de la mortalité après 1986 semble attribuable au sida, alors que dans d'autres pays elle n'en rend compte que d'une partie (voir § 7)

3.7 Évolutions négatives récentes (sans sida)

L'évolution de la mortalité a été négative dans plusieurs pays au cours de la période récente du fait du sida, mais reste défavorable même quand on soustrait l'impact du sida pédiatrique. Ainsi, en Afrique du Sud, si on soustrait l'impact du sida, la mortalité des jeunes enfants cesse de baisser après 1993. Au Cameroun la mortalité nette de sida est estimée à 145 pour 1000 en 1998, alors que la tendance précédente (1950-1988) prévoyait 101 pour 1000. En Côte d'Ivoire, l'évolution récente sans sida reste positive, mais sa pente est très inférieure à ce qu'elle était avant 1983. Au Kenya, la mortalité sans sida augmente après 1985.

3.8 Évolutions divergentes des milieux urbain et rural

En général, la mortalité évolue de la même manière dans les milieux urbain et rural. Cependant on note plusieurs cas d'évolutions divergentes. Ainsi, au Burundi la mortalité rurale baisse entre 1976 et 1986 alors qu'elle augmente en milieu urbain. En Centre-Afrique elle augmente en milieu rural alors qu'elle continue à baisser en milieu urbain entre 1976 et 1990. Au Malawi, la mortalité augmente fortement en milieu urbain après 1984, alors qu'elle ne fait que stagner en milieu rural. Au Mozambique, la différence de mortalité des deux milieux diverge fortement pendant la période de guerre civile, qui a surtout affecté le milieu rural. Au Niger, la mortalité stagne en milieu rural dans les années 1972-1992 alors qu'elle continue à baisser en milieu urbain.

4. Discussion

Les données disponibles sur la mortalité des jeunes enfants en Afrique sub-Saharienne ne couvrent que certains pays et certaines périodes, et souffrent de fortes erreurs de sondage. Elles permettent cependant de reconstruire les grandes tendances et certains accidents de parcours depuis 1950 jusqu'à la dernière date disponible, en général dans les années 1990. Ces données montrent que la transition sanitaire, pourtant très nette au niveau du continent, a été émaillée de nombreux incidents et accidents de parcours. L'évolution récente est marquée par l'émergence d'un fort impact du sida pédiatrique dans de nombreux pays.

Ces évolutions renvoient aux principaux déterminants de l'évolution de la mortalité. Si la transition sanitaire en Afrique est loin d'être régulière, ce sont surtout les crises politiques, les guerres civiles, la mauvaise gestion de l'Etat, et les difficultés économiques y relatives qui en sont la cause. A cela s'ajoute l'émergence du sida, et dans certains cas les fluctuations du paludisme. Certaines augmentations récentes, dans les années 1990, qui persistent après soustraction de l'impact du sida, méritent une attention particulière. Il pourrait s'agir de l'effet de l'augmentation des inégalités et de la pauvreté, surtout en milieu urbain, comme le suggère les cas du Kenya, du Cameroun et de la Côte d'Ivoire.

Cette étude présente plusieurs avantages sur les essais antérieurs. Par rapport aux estimations indirectes, qui par ailleurs sont anciennes, on obtient une meilleure estimation des niveaux, des tendances, et surtout des points d'inflexion. Par rapport aux estimations de l'OMS, fondées sur une combinaison des données directes et indirectes, on obtient souvent des tendances différentes, qui indiquent des phénomènes

particuliers. Enfin, la possibilité de soustraire l'impact du sida permet de mieux cerner certaines évolutions récentes.

Cette reconstruction n'est qu'un préalable à l'étude détaillée des évolutions dans chaque pays, dont la complexité apparaît déjà à la lecture des figures par pays. La transition sanitaire en Afrique sub-Saharienne a été rendue complexe par l'interférence de plusieurs phénomènes dont les principaux semblent être l'instabilité politique, la mauvaise gestion de l'Etat et l'émergence du sida. Si la diversité des situations du sida dans les pays africains est maintenant bien documentée, celle des situations politiques et économiques mérite une attention particulière.

Références

Ahmad OB, Lopez AD, Inoue M. (2000). The Decline in Child Mortality: a Reappraisal. Bulletin of the World Health Organization, 78(10), 1175-1191.

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

Feachem, R., and Jamison, D. eds. (1991). Disease and Mortality in sub-Saharan Africa, World Bank / Oxford University Press, New-York, NY, USA.

Hill, A. (1991). Infant and Child Mortality: Levels, Trends and Data Deficiencies. In: R. Feachem and D. Jamison, eds., Disease and Mortality in sub-Saharan Africa, World Bank / Oxford University Press, New-York, NY, USA, 37-76.

Hill, A. and Hill, K. (1988). Mortality in Africa: Levels, Trends, Differentials and Prospects. In: E. van de Walle, P.O. Ohadike, and M.D. Sala-Diakanda, eds. The State of African Demography, IUSSP-Derouaux, Liège, Belgium, 67-101.

Mandjale, AE. (1985). Mortalité infantile et juvénile en Afrique: niveaux et caractéristiques, causes et déterminants. CIACO. Louvain la Neuve, Belgique.

McKeown T. (1976). The Modern Rise of Population. New York, Academic Press.

Newell ML. et al. (2004). Mortality of Infected and Uninfected Infants Born to HIV-infected Mothers in Africa: a Pooled Analysis. Lancet, 364:1236-1243.

Preston SH. (1980). Causes and Consequences of Mortality Declines in Less Developed Countries During the Twentieth Century. In: R.A. Easterlin, ed. Population and Economic Change in Developing Countries. University of Chicago, 289-353.

Preston, S.H. (1976). Mortality Patterns in National Populations. New York, Academic Press.

Schoffield R. and Reher D. (1991). The Decline of Mortality in Europe. In: R. Schoffield, D. Reher, Alain Bideau eds. The Decline of Mortality in Europe. Oxford, Clarendon Press, 1-17.

Stolnitz G. (1955). A Century of International Mortality Trends, pt. 1, Population Studies, 9(1), 24-55.

Stolnitz GJ. (1965). Recent Mortality Trends in Latin America, Asia and Africa: Review and re-interpretation. Population Studies 19(2), 17-138.

United Nations, Population Division. (1982). Levels and Trends of Mortality Since 1950. UN Population Study No. 74.

van de Walle, E., Ohadike, P.O., Sala-Diakanda, M.D., eds. (1988). The state of African demography, IUSSP-Derouaux, Liège, Belgique.

Annexe. Liste des pays et des enquêtes

Code	***Nom du pays***	***Enquêtes***		
AO	Angola	MICS, 1996		
BJ	Bénin	WFS, 1982	DHS, 1996	DHS, 2001
BW	Botswana	DHS, 1988		
BF	Burkina Faso	DHS, 1993	DHS, 1999	
BI	Burundi	DHS, 1987		
CM	Cameroun	WFS, 1978	DHS, 1991	DHS, 1998
CF	République Centre Africaine	DHS, 1994		
TD	Tchad	DHS, 1997		
KM	Comores	DHS, 1996		
CI	Côte d'Ivoire	WFS, 1980	DHS, 1994	DHS, 1999
ET	Ethiopie	DHS, 2000		
GA	Gabon	DHS, 2001		

GH	Ghana	WFS, 1980	DHS, 1993	DHS, 1999	
GN	Guinée	DHS, 1999			
KE	Kenya	WFS, 1978	DHS, 1993	DHS, 1999	
LS	Lesotho	WFS, 1977			
LR	Liberia	DHS, 1996			
MG	Madagascar	DHS, 1992	DHS, 1997		
MW	Malawi	DHS, 1992	DHS, 2000		
ML	Mali	DHS, 1987	DHS, 1995	DHS, 2001	
MZ	Mozambique	DHS, 1997			
NA	Namibie	DHS, 1992	DHS, 2000		
NE	Niger	DHS, 1992	DHS, 1998		
NG	Nigeria	WFS, 1982	DHS, 1990	DHS, 1999	
RW	Rwanda	WFS, 1983	DHS, 1992	DHS, 2000	
SN	Sénégal	WFS, 1978	DHS, 1986	DHS, 1993	DHS, 1997
ZA	Afrique du sud	DHS, 1998			
SD	Soudan	WFS, 1979	DHS, 1989		
TZ	Tanzanie	DHS, 1991	DHS, 1996		
TO	Togo	DHS, 1988	DHS, 1998		
UG	Uganda	DHS, 1988	DHS, 1995		
ZM	Zambie	DHS, 1992	DHS, 1996	DHS, 2001	
ZW	Zimbabwe	DHS, 1988	DHS, 1993	DHS, 1999	

Editorial policy

The African Statistical Journal was established to promote the understanding of statistical development in the African region. It focuses on issues related to official statistics as well as application of statistical methodologies to solve practical problems of general interest to applied statisticians. Of particular interest will be exposition of: how statistics can help to illuminate development and public policy issues like poverty, gender, environment, energy, HIV/AIDS, etc; development of statistical literacy; tracking national and regional development agenda; development of statistical capacities and effective national statistical systems; and the development of sectoral statistics e.g. educational statistics, health statistics, agricultural statistics, etc.

In addition to individual academic and practicing statisticians, the Journal should be of great interest to a number of institutions in the region including National Statistical Offices, Central Banks, research and training institutions and sub-regional economic groupings, and international development agencies.

The Journal serves as a research outlet and information sharing publication among statisticians and users of statistical information mainly in the African region. It publishes, among other things:

- articles of an expository or review nature that demonstrate the vital role of statistics to society rather than present technical materials,
- articles on statistical methodologies with special emphasis on applications,
- articles about good practices and lessons learned in statistical development in the region,
- opinions on issues of general interest to the statistical community and users of statistical information in the African region,
- notices and announcements on upcoming events, conferences, calls for papers, and
- recent statistical developments and anything that may be of interest to the statistical community in the region.

The papers which need not contain original material, should be of general interest to a wide section of professional statisticians in the region.

All manuscripts will be reviewed and evaluated on content, language and presentation.

Ligne éditoriale

Le Journal statistique africain a été établi pour favoriser la compréhension du développement statistique dans la région africaine. Il se concentre sur des questions liées aux statistiques officielles aussi bien que l'application des méthodologies statistiques pour résoudre des problèmes pratiques d'intérêt général pour les statisticiens de métier. L'intérêt particulier est de montrer comment les statistiques peuvent aider à mettre en exergue les problèmes de développement et de politique publique tels que la pauvreté, le genre, l'environnement, l'énergie, le VIH/ SIDA, etc. ; le développement de la culture statistique ; la prise en compte des questions de développement régional et national; le développement des capacités statistiques et des systèmes statistiques nationaux efficaces; et le développement des statistiques sectorielles comme les statistiques d'éducation, de santé, des statistiques agricoles, etc.

En plus des universitaires et des statisticiens de métier, le Journal devrait revêtir un grand intérêt pour les institutions de la région, notamment les offices nationaux de statistiques, les banques centrales, les instituts de recherche et les organisations économiques sous-régionaux et les agences internationales de développement.

Le Journal constitue un document de recherche et d'information entre les statisticiens et les utilisateurs de l'information statistique, principalement dans la région africaine. Il publie entre autres :

- des articles sur le plaidoyer en matière de statistique qui démontrent le rôle essentiel des statistiques dans la société plutôt que la présentation des outils techniques,
- des articles sur les méthodologies statistiques, avec un accent particulier sur les applications,
- des articles sur les meilleures pratiques et les leçons tirées de la région,
- des avis sur des questions d'intérêt général pour la communauté statistique et les utilisateurs de l'information statistique dans la région africaine,
- des informations et des annonces sur les prochains événements, les conférences, les appels à contribution pour des papiers, et
- les développements statistiques récents et tout autre aspect susceptible d'intéresser la communauté statistique dans la région.

Les articles, qui n'ont pas besoin de contenir du matériel original, devraient intéresser une grande partie des statisticiens professionnels dans la région.

Tous les manuscrits seront passés en revue et évalués sur le contenu, la langue et la présentation.

NOTES TO AUTHORS

Submission

Manuscripts in English or French should be sent by email to the Co-Chairpersons, Editorial Board, at c.lufumpa@afdb.org and bkiregyera@yahoo.com with a copy to statistics@afdb.org. and africanstat.journal@ubos.org

Title

The title should be brief and specific. The title page should include the title, the author's name, affiliation and address. The affiliation and address should be given as a footnote on the title page. If the manuscript is coauthored, the same information should be given for the coauthor(s).

Summary, Key Words and Acknowledgements

A short summary of about 150 words must be included at the beginning of the manuscript together with up to 6 key words used in the manuscript. The key words should not repeat words used in the title. Acknowledgements, if any, should be inserted at the bottom of the title page.

Sections

Sections should be numbered. Subsections may be used.

Tables and Figures

Tables and figures should be numbered and given a title. These should be referred to in the text by number, not by page or indications such as "below" or "above".

Equations

Any equations in the paper should be numbered. The numbers should be placed to the right of the equation.

References

A list of references should be given at the end of the paper. The references should be arranged alphabetically, and for the same author chronologically. The references should give author's name and year of publication, title and details of the publication – name of Journal. Use a,b,c, etc to separate publications of the same author in the same year.

Examples

Kish, L. (1988a). Multipurpose Sample Designs, Survey Methodology, 14, 19-32.

Kish, L. (1988b). A Taxonomy of Elusive Populations, Proceedings of the Section on Survey Research Methods, American Statistical Association, 44-46.

Herzog, A.R. and Dielman, L. (1985). Age Differences in response Accuracy for Factual Survey Questions, Journal of Gerontology, 40, 350-367.

In the text, the author's surnames only should be given, followed by the year of publication in parentheses e.g. Kish (1988a). For three or more authors, only the first surname should be given, followed by et al. Abbreviations ibid, opt. cit. should not be used.

NOTES AUX AUTEURS

Soumission

Les manuscrits en anglais ou en français doivent être envoyés aux présidents du comité de rédaction par email aux adresses suivantes c.lufumpa@afdb.org et bkiregyera@yahoo.com avec copie à statistics@afdb.org et africanstat.journal@ubos.org

Titre

Le titre devrait être bref et détaillé. La page de titre doit inclure le titre du papier, le nom de l'auteur, l'affiliation et l'adresse. L'affiliation et l'adresse doivent figurer comme note de bas de page. Si le manuscrit est produit par des coauteurs, la même information doit être donnée pour les coauteurs.

Résumé, mots clés et reconnaissance

Un résumé court d'environ 150 mots doit être inclus au début du manuscrit ainsi qu'environ 6 mots clés utilisés dans le manuscrit. Les mots clés ne doivent pas répéter les mots utilisés dans le titre. Les signes de reconnaissance, s'il y en a, doivent être insérés en bas de la page titre.

Section

Les sections doivent être numérotées. Des sous-sections peuvent être employées.

Tableaux et graphiques

Les tableaux et les graphiques doivent être numérotés et comporter un titre. Ceux-ci devraient être mentionnés dans le texte par le nombre correspondant, et non par une indication de page ou par d'autres indications telles que "ci-dessous" ou "au-dessus de".

Équations

Toutes les équations dans le papier doivent être numérotées. Les nombres doivent être placés à la droite de l'équation.

Références

Une liste de références doit être fournie à la fin du papier. Les références doivent être classées par ordre alphabétique, et pour le même auteur chronologiquement. Les références doivent donner le nom de l'auteur et l'année de publication, le titre et autres détails concernant la publication. Employer a, b, c, etc.. pour séparer les publications du même auteur au cours der la même année.

Exemples

Kish, L. (1988a). Multipurpose Sample Designs, Survey Methodology, 14, 19-32.

Kish, L. (1988b). A Taxonomy of Elusive Populations, Proceedings of the Section on Survey Research Methods, American Statistical Association, 44-46.

Herzog, A.R. and Dielman, L. (1985). Age Differences in response Accuracy for Factual Survey Questions, Journal of Gerontology, 40, 350-367.

Ne doivent figurer dans le texte que les noms de famille des auteurs, suivi de l'année de la publication entre parenthèses par exemple Kish (1988a). Pour trois auteurs ou plus, seulement le premier nom de famille devraient être donnés, suivi des autres. Les abréviations, comme ibid, opt, cit ne doivent pas être employées

Acknowledgements

The Editorial Board would like to express its appreciation to all authors who submitted papers for publication in the journal and to the following people who reviewed the papers that are published in this issue of the Journal:

Peter Lanjouw, Development Economics, The World Bank, 1818 H. St. N.W., Washington D.C. 20433, U.S.A.

David Roberts, OECD-TOUR 1326, 2 rue André Pascal, 75775 Paris Cedex 16, France

Tony Williams, Senior Adviser, PARIS 21 Secretariat, 2 Rue Andre'-Pascal, 75775 Paris, Cedex 16, France

Tim Merchant, Clock house, Berwick St. James, Salisbury, Wiltshire SP3 4TN, UK

Graham Eele, Senior Statistician, World Bank, 1818 H Street NW, Washington DC 20433 USA

Wold Bjørn Kjetil, Statistics Norway, P.O.B 8131 Dep, NO-0033 Oslo, Norway

Remerciements

Le comité de rédaction voudrait exprimer sa gratitude à tous les auteurs qui ont soumis des articles pour publication dans le journal et aux personnes suivantes qui composent le comité de lecture des articles qui sont publiés dans cette édition du journal.

Peter Lanjouw, Economie du Développement, Banque Mondiale, 1818 H. St. N.W., Washington D.C. 20433, U.S.A.

David Roberts, OCDE-TOUR 1326, 2 rue André Pascal, 75775 Paris Cedex 16, France

Tony Williams, Conseiller Senior, Secrétariat PARIS 21 St, 2 Rue Andre'-Pascal, 75775 Paris, Cedex 16, France

Tim Merchant, Clock house, Berwick St. James, Salisbury, Wiltshire SP3 4TN, UK

Graham Eele, Statisticien Senior, Banque Mondiale, 1818 H Street NW, Washington DC 20433 USA

Wold Bjørn Kjetil, Statistics Norway, P.O.B 8131 Dep, NO-0033 Oslo, Norway

Upcoming events
Evénements en vue

Celebration of African Statistics Day
Luanda Angola, December 4-8, 2006

We are pleased to announce that the Angolan Government has graced the African Statistics Community by inviting them to Angola to celebrate the African statistics day which although is the 18th of November, will be celebrated in Angola in the week of the 4 to 8 December. In this celebration, we will have the African countries represented by their heads of statistics offices, the ECA represented by the Executive Secretary Dr Abdulie Janneh, the AfDB represented by the President Dr Kaberuka, the African Union Commission represented by the Chairperson Dr Konare. Key events of the celebrations will include an AfDB workshop on the preliminary results of the International Comparison Program and presentations of the development account program for SADC Countries by the United Nations Statistics Division. These two eminent activities will contribute decisively to the harmonization of statistics in Africa. A Luanda declaration will also be signed at the conclusion of the celebrations. This will focus both on Angola and Africa in matters of statistical development, especially the undertaking of a census in Angola and elsewhere in on the continent. African statistics are on the move!

Nous avons le plaisir de vous annoncer que le Gouvernement Angolais a honoré la Communauté Statistique Africaine en l'invitant en Angola pour célébrer la journée de la statistique Africaine. Cette journée qui normalement est célébrée le 18 Novembre le sera cette fois-ci pendant la semaine du 4 au 8 Décembre. Durant cette commémoration, les pays Africains seront représentés par leurs Directeurs des Instituts Nationaux de Statistique, la CEA par son Secrétaire Exécutif Dr Abdulie Janneh, la BAD par son Président Dr Kaberuka et l'Union Africaine par son Président Dr Konaré. Les deux événements-clés de cette célébration sont : (i) l'atelier de la BAD sur la présentation des résultats préliminaires du Programme de Comparaison Internationale et (ii) la présentation du programme « development account » de la Division Statistique des Nations Unies en faveur des pays de la SADC. Ces deux importantes activités vont résolument contribuer à l'harmonisation de la statistique en Afrique. Une déclaration de Luanda sera signée à l'issue de ces festivités. Elle concernera les questions relatives au développement de la statistique en Angola et en Afrique de façon générale, notamment la réalisation d'un recensement en Angola et ailleurs sur le continent. La statistique Africaine est en marche !

ICP Retreat for the Generation of Purchasing Power Parities (PPP):
Tunis, Tunisia, From 16 October – 15 November 2006

Retraite du PCI sur le calcul des parités des pouvoirs d'achat (PPA):
Tunis, Tunisie, du 16 Octobre au 15 Novembre 2006

The main objectives of the retreat are: (1) the calculation of purchasing power parities; (2) the preparation of the first publication of ICP results for the Africa Region; (3) the recommendations concerning the second publication of ICP results.

Les principaux objectifs de la retraite sont: (1) le calcul des parités de pouvoir d'achat ; (2) la préparation de la publication des premiers résultats du PCI pour la région africaine ; (3) les recommandations concernant la publication de la deuxième série des résultats du PCI.

Regional Workshop on the African Comparison of PPP-adjusted GDPs and Price Level Differentials: Luanda, Angola, from December 4-8, 2006

Atelier Régional sur la Comparaison Africaine des PIB déflatés par les PPA et le Différentiel des Niveaux des Prix: Luanda, Angola, du 4 au 8 Décembre, 2006

The main objective of the workshop is to review and endorse:

- the detailed GDP estimates for RMCs participating in the ICP-Africa
- the main results of PPP-adjusted GDPs and Price Levels generated by the ICP-Africa

Country participants are also expected to discuss the plan of publication of ICP results for the reference years 2005 and 2006, at country, sub-regional, regional and world levels.

L'objectif principal de l'atelier est de passer en revue et approuver :

- *les estimations détaillées du PIB pour les pays participant au PCI-Afrique ;*
- *les résultats principaux du PIB déflatés par les PPA et les niveaux des prix générés par le PCI-Afrique.*

On s'attend également à ce que des participants des pays discutent le plan de publication des résultats du PCI pour les années de référence 2500 et 2006, aux niveaux national, sous-régional régional et mondial.

African Symposium and FASDEV III: Kigali, Rwanda, January, 2007

Symposium Africain et FASDEV III: Kigali, Rwanda, Janvier, 2007

Design of the National Strategy for the Development of Statistics in Africa

The design of the National Strategy for the Development of Statistics (NSDS) has picked pace in Africa with support from the African Development Bank (ADB) and other development partners. The Marrakech Action Plan for Statistics (MAPS), which aims to improve statistics at national and international level, set 2006 as the target date by which all low-income countries should have developed an NSDS with a view to producing better statistics for national and international use by the time of the next Millennium Review in 2010. The design of the NSDS is also the overarching strategy for the Reference Regional Strategic Framework for Statistical Capacity Building in Africa (RRSF) which was endorsed by a meeting of Directors of National Statistical Offices from all over Africa and the Forum on African Statistical Development early this year.

ADB has also set aside resources to assist regional member countries to update and/or design their statistical plans using the NSDS principles. These principles have been developed by PARIS21 (Partnerships in Statistics for Development in the 21st Century) in collaboration with partners. ADB has in recent past formerly launched its assistance to the NSDS process in four sub-regional workshops, namely

- the COMESA workshop held in Nairobi, Kenya from 15 to 18 May 2006 for Directors of National Statistical Offices (NSOs) and Poverty Reduction Strategy (PRS) Units from nine COMESA member states of Djibouti, Burundi, Egypt, Ethiopia, Kenya, Madagascar, Rwanda, Uganda and Sudan. The workshop was attended by Directors of NSOs and Heads of Poverty Reduction Strategy (PRS) Units from the countries participating in the workshop.

- the ECOWAS workshop held in Accra, Ghana from 29 May to 1st June 2006, 2006 for five English-speaking member states of Gambia, Ghana, Liberia, Nigeria and Sierra Leone in Accra. The workshop was attended by Directors of NSOs and Heads of Poverty Reduction Strategy (PRS) Units from the countries participating in the workshop.

- the SADC workshop held in Durban, South-Africa from 10 to 12 July 2006. This workshop was attended by included Directors of NSOs, NSDS team leaders, Heads of Poverty Reduction Strategy Units and Heads of Statistics Units in Ministries of Health and Agriculture in SADC member countries of Botswana, Lesotho, Malawi, Mauritius, Mozambique, Namibia, South Africa, Swaziland, Tanzania, Zambia and Zimbabwe. This was the first time professionals in sectors had been invited to such a workshop because of the new emphasis being placed on integration of sectoral concerns and issues into the NSDS.

- the AFRISTAT workshop held in Bamako, Mali from 28 August to 1 September 2006. The workshop participants included Directors of NSOs, NSDS team leaders, Heads of Poverty Reduction Strategy Units and Heads of Statistics Units in Ministries of Health and Agriculture in SADC member countries

The other objectives of the workshops were to: (i) empower key actors in National Statistical Systems with knowledge about the NSDS processes and how to go about them, and (ii) facilitate sharing of experience among countries about the NSDS process. During the workshop, each recipient country team was assisted to prepare an NSDS design "roadmap" (if the country did not have one already) and an associated budget.

Elaboration de la stratégie nationale de développement de la statistique en Afrique

Le processus d'élaboration de la stratégie nationale de développement de la statistique (SNDS) a marqué le pas en Afrique avec l'appui de la Banque de développement africaine (BAD) et d'autres partenaires au développement. Le plan d'action de Marrakech pour la statistique (MAPS), dont l'objectif est d'améliorer les statistiques aux niveaux national et international, a fixé 2006 comme date butoir à laquelle tous les pays à faible revenu devraient avoir élaboré une SNDS en vue de produire de meilleures statistiques pour l'usage national et international avant la prochaine revue du millénium en 2010. L'élaboration de la SNDS est également la stratégie fondamentale du cadre stratégique régional de référence (CSRR) pour le renforcement des capacités statistiques en Afrique qui a été approuvée par une réunion des directeurs des instituts nationaux statistiques des pays africains et le Forum sur le développement statistique en Afrique (FASDEV) au début de cette année.

La BAD a également mobilisé des ressources pour aider les pays membres régionaux à mettre à jour et/ou concevoir leurs plans statistiques en utilisant les principes de la SNDS. Ces principes ont été développés par PARIS21 (Partenariat statistique pour le développement au 21e siècle) en collaboration avec des partenaires. La BAD a récemment lancé officiellement son assistance au processus SNDS à travers quatre ateliers sous-régionaux, à savoir :

- L'atelier du COMESA tenu à Nairobi, Kenya, du 15 au 18 mai 2006 qui a réuni des directeurs des INS et des responsables des Unités de stratégie de réduction de pauvreté (SRP) de neuf Etats membres du COMESA : Djibouti, Burundi, Egypte, Ethiopie, Kenya, Madagascar, Rwanda, Ouganda et Soudan.

- L'atelier de la CEDEAO tenu à Accra, Ghana, du 29 mai au 1er juin 2006, pour cinq Etats membres d'expression anglaise : Gambie, Ghana, Libéria, Nigeria et Sierra Leone. L'atelier a réuni des directeurs des INS et des responsables des unités SRP des pays participants.

- L'atelier de la SADC tenu à Durban, Afrique du Sud, du 10 au 12 juillet 2006. Cet atelier a réuni les directeurs des INS, les coordonnateurs des SNDS, les responsables des unités SRP et des unités statistiques aux ministères de la santé et de l'agriculture dans les pays membres de la SADC suivant: Botswana, Lesotho, Malawi, Ile Maurice, Mozambique, Namibie, Afrique du Sud, du Swaziland, Tanzanie, Zambie et Zimbabwe. C'était la première fois que des professionnels dans ces secteurs avaient été invités à un tel atelier en raison de la nouvelle emphase placée sur l'intégration des préoccupations et des questions sectorielles dans la SNDS.

- L'atelier des pays membres d'AFRISTAT a eu lieu à Bamako, Mali, du 28 août au 1er septembre 2006. Les participants à cet atelier comprenaient des directeurs des INS, des coordonnateurs de SDNS et des chefs des unités de SRP des pays suivants : Bénin, Burkina Faso, Cap vert, Cameroun, Centrafrique, Comores, Congo, R.D. Congo, Côte d'Ivoire, Gabon, Guinée, Guinée Bissau, Mali, Mauritanie, Niger, Sao Tomé et Principe, Sénégal, Tchad et Togo. La réunion s'est focalisée sur la planification stratégique, la gestion axée sur les résultas et l'intégration des préoccupations et des questions sectorielles dans la SNDS.

Les autres objectifs de ces ateliers étaient de : (i) renforcer les connaissances des acteurs clés dans les systèmes statistiques nationaux sur le processus de la SNDS et comment les mettre en œuvre, et (ii) faciliter le partage d'expériences entre les pays sur le processus de la SNDS. Au cours de l'atelier, chaque équipe des pays bénéficiaires a été assistée pour élaborer une « feuille de route » (si le pays n'en disposait pas auparavant) et un budget associé.

ADB/AFRISTAT NSDS Workshop

Photo de groupe des participants lors de l'ouverture officielle de l'atelier organisé par la BAD et AFRISTAT sur l'élaboration de la SNDS à Bamako du 29 août au 1er septembre 2006. Le séminaire a été ouvert par le Ministre du Plan et de l'Aménagement du Territoire du Mali. (*Group picture at the official opening of the ADB/AFRISTAT NSDS workshop held in Bamako, Mali from 28 August to 1 September 2006. The workshop was officially opened by The Minister for Planning and Regional Development of Mali)*

Launch of NSDS process in Tanzania

His Excellency, Dr Ali Mohamed Shein, Vice President of the United Republic of Tanzania officially launched the NSDS Process in Dar es Salaam, Tanzania on 6 October, 2006 (*Lancement du processus d'élaboration de la SNDS par le Vice-Président de la République Unie de Tanzanie le 6 Octobre 2006)*

Launch of NSDS document and opening statistics house in Uganda

His Excellency, Yoweri Kaguta Museveni, the President of Uganda officially launched the NSDS document and opened Statistics House in Kampala, Uganda on 12 October, 2006

Lancement du Document de la SNDS et ouverture officielle de la Maison de la Statistique par le Président de l'Ouganda le 12 Octobre 2006

Opening of the ADB/SADEC NSDS workshop in Durban, South Africa, 10-13 July 2006.

(Left-right): Mr. Ackim Jere, Senior Programme Manager – Statistics, SADC Secretariat, Mr. Charles Lufumpa, Ag. Director, Statistics Departement, ADB, Mr. Pali Lehohla, Statisticion General, South Africa)