Statistics Norway
Statistisk sentralbyrå

João Dias Loureiro - INE,
Bjørn K. Wold - Statistics Norway,
Richard Harris - World Bank

Compendium from Workshop on Light Core Surveys for Policy Monitoring of National PRSPs and MDGs in Maputo, December 2005

Statistics Norway Division for Development Cooperation

João Dias Loureiro – INE,
Bjørn K. Wold – Statistics Norway,
Richard Harris – World Bank

Compendium from Workshop on Light Core Surveys for Policy Monitoring of National PRSPs and MDGs in Maputo, December 2005

Abstract
This paper reports from a workshop in Maputo in December 2005 aiming at presenting and discussing how a survey program built around a light core survey program may serve to monitor national poverty reduction strategies and similar policy programs as well as the Millennium Development Goal indicators. The workshop focused on five thematic building blocks, being household surveys for monitoring of national Poverty Reduction Strategy Plans (PRSPs) and MDGs, statistical models for estimation of poverty and small area statistics, data entry and data processing, analysis and dissemination, and the International Household Survey Network.

The workshop comprised a number of generic papers on survey approach and methods and a number of country level papers presenting experience and plans from 10 African countries and the audience represented a broad range of expertise, from national statistical offices, international organizations and scientific institutions.

The workshop made a series of detailed recommendations for household survey programmes & the role of light core surveys, for related IT issues on data entry and data processing and for related statistical modeling and analysis.

Acknowledgement: We are grateful to INE/ Mozambique which invited us all to Maputo and served as our host, to NORAD and World Bank, which funded the workshop and to INE/Mozambique, Statistics Norway and the World Bank which served as technical organizers of the workshop, to the long range of technical experts who made the series of interesting and inspiring presentations and to all participants whose active contributions ensured a successful outcome of the workshop.

Addresses: João Dias Loureiro, Instituto Nacional de Estatística, C.P. 493, Maputo, Mozambique, E-mail: info@ine.gov.mz ; Bjørn K. Wold, Statistics Norway, P.O.Box 8131 DEP, N-0033 Oslo, Norway. E-mail: bkw@ssb.no ; Richard Harris, World Bank, 1818 H street, Washington DC 20433, USA, E-mail rharris@worldbank.org .

Contents

1 Background

The objective of the workshop was to exchange experiences and to encourage national statistics offices to think strategically about monitoring their national PRSPs and the MDGs, especially on how to design of relevant household survey programmes based upon light core surveys or core survey modules. The workshop discussed a broad range of issues linked to the use of a core household survey such as the Core Welfare Indicators Questionnaire (or a core module based on the CWIQ but incorporated into other surveys) within a poverty monitoring system for monitoring PRSP and MDG indicators.

The workshop was held in Maputo over 2½ days, sponsored by the Instituto Nacional Estatistica of Mozambique (INE), Statistics Norway (SN) and the World Bank (WB).

Virtually all the National Statistics Offices (NSO) participating were part of the regional CWIQ programme and had undertaken at least one national CWIQ survey or pilot survey.

2 Workshop programme

The workshop programme followed five thematic building blocks, as follows:

1. Household surveys for monitoring of Poverty Reduction Strategy Plans (PRSPs) and MDGs
2. Statistical models for estimation of poverty and small area statistics
3. Data entry and data processing
4. Analysis and dissemination
5. The International Household Survey Network

The workshop also included a presentation of a Programmatic approach to planning and implementing statistics by Ms. Louise Fox, World Bank. This presentation stressed the needs for statistics to be a part of the PRSP process.

3 Household surveys for monitoring of PRSPs and MDGs, approach and country experience

3.1 Presentations

The presentations made were as follows:
1. A sustainable Household Survey System with a Core Survey for Poverty Monitoring for PRSPs and MDGs. By Bjørn K. Wold
2. Country experiences. By Country participants
3. The QWIQ survey designed as core of a household survey program. By Tim Merchant
4. Country experiences. By Country participants
5. Household Survey Programs - reflections on the Report for PARIS21 Task Force on Improved Statistical Support for Monitoring MDGs: "Household Surveys and the Millennium Development Goals". By Richard Harris
6. Country experiences. By Country participants

Statistics Norway and the World Bank introduced this topic during the first day of the workshop by presenting 3 background papers outlining issues around the development of poverty monitoring systems and household survey programmes and the role of light monitoring surveys such as the CWIQ. Issues highlighted are including the following ones:

- Developing capacity to respond to evolving data needs rather than produce particular surveys.
- Having the right tool for the job – its not feasible that one survey will meet all data needs.
- Developing a household survey system with different complimentary components.
- Developing a core household survey or module, based on scanning, fast processing and reporting, to include appropriate MDG & PRSP indicators.
- Using light core surveys to focus on indicators such as access, use and satisfaction with service delivery, to measure early responses to policy initiatives.

In between these papers countries - Mozambique, Angola, Ghana, Kenya, Lesotho, Malawi. Rwanda, Sierra Leone, South Africa and Zambia – made presentations highlighting their PRSP monitoring experiences. Common features included the following ones:

- NSOs were involved in PRSPs through developing and determining indicators.
- Most NSOs have a survey programme with an Integrated Household Survey or Household Budget Survey once every 4-5 years and stand alone CWIQs or Welfare Monitoring Surveys in between. (Mozambique had developed a core survey module based on the CWIQ which was to be included in all surveys)
- A number of NSOs reported quick production of results followed by dissemination workshops but slow formal publication of reports
- Most NSOs have or are planning to produce Poverty Maps combining Census data with a model predicting welfare levels produced from an Integrated Survey or Household Budget Survey dataset.
- Most NSOs are using or are planning to introduce a national socio-economic database based on the DEVINFO package.
- NSOs reported high demand for district level estimates

3.2 Discussion

During the discussion, it was stressed that sample size is important, but depends on the choice about the most important indicators. The choice of stratification variables depends on the chosen priority indicators. It is thus important to decide upon these. The need of finding a proper way to present error terms in an understandable way was highlighted.

Comments to the country presentations addressed a number of common as well as specific issues. Three general comments were

- There is a need to put more emphasis on the linking quantitative and qualitative methods.
- There is a problem with the demand side for data. Resources must be allocated to promote the demand for data among users. Doubtful whether there should be emphasis on more disaggregated *poduction* of data before the higher-level data are actually used.
- About the content of the "Core" surveys: An often-overlooked benefit of "Core" surveys is its importance as a capacity building tool, especially if conducted annually. When there is turnover in the statistical agencies, the lack of annual core surveys may imply that the competence in conducting and analyzing surveys are lost.

During the discussion the audience asked South Africa to elaborate on the duration of the expenditure reference period. South Africa explained that they had used a one month approach with 5 interviews,

no return visit as long as have information, combining recall (durable and semi-durable goods) and diary (non-durable goods) methods.

The design of the South Africa community survey was found different from other countries. Usually data are collected at community level providing information about the community as such. For South Africa, the "community survey" seems rather a way to collect summary household information. That might not be the best approach.

The close contact between two neighbour country statistical offices such as for Mozambique and South Africa has proved very useful and is in this case based upon a government agreement between the two countries. Close, and two-way contact, as well as a short flight. This approach was found an interesting example also by other countries such as Malawi and Zambia. This view was supported from Norway which had cooperated regularly by statistical offices from other Nordic countries over 50 years and now were in the position not only from borrowing methods from each other but also common development.

Several issues on how policy needs may guide the design of a survey program were raised, as follows:

- The need for different survey tools for different types of policy issues was raised. With the increasing demand for data this was found increasingly important.
- Will policy be satisfied by only getting poverty estimates every 5-year? r as frequent and detailed as possible.
- Not much information in poverty incidence. More important for government whether its policy has effect. Looking at the whole chain of policy implementation process.
- Policy process in Uganda as illustrative example. Heavy-duty surveys have given poverty estimates. Do such estimates take the interest away from the *causes* of poverty? Child malnutrition, inequality, malnutrition up while poverty rates are going down. Those near the poverty line are better off, but not those really poor.
- Uganda experience interesting. 90's try to do annual surveys. Then period increased, first to two, then three and then four years. Come up with an alternative way to get annual estimates without an annual HBS.

The need for annual information was stressed, but two opposite views were raised for how to gather information in the years in between large surveys. Some argued for other and simpler indicators, while others argued for annual (or bi-annual) light core surveys including a statistical poverty indicator model. It was stressed that in order to present real trends, 6-7 or at least 5 observations were needed.

It was however a full agreement of the support for annual light surveys to present rapid changes along the various non-monetary poverty dimensions.

The definition and content of the poverty concept was discussed. This was found a multi-dimensional concept. From the decision-makers point of view the issues is how to break the cycle of poverty. Other studies are also necessary. Qualitative studies to complement the traditional statistical information about poverty. Must talk about adult equivalence scales and intra-household distribution issues. Not sufficient only to have concern about poverty incidence.

Others argued that the high policy focus on poverty incidence requires the statistical community to develop and implement light core surveys with a statistical model for estimation of poverty incidence on an annual basis.

4 Statistical models for estimation of poverty

4.1 Presentations

The following presentations were made:
1. Poverty Correlates to Rank Households and Present Results by Quintiles. By John Ngwafon
2. A statistical model for simple, fast and reliable measurement of poverty. By Astrid Mathiassen.
3. How well do light surveys predict poverty? Preliminary results from a validation study. By Kenneth Simler
4. Poverty Monitoring – lessons from Mozambique. By Arndt Channing
5. Large samples versus small area estimation for district level. By Stein Opdahl

Fully-fledged consumption & expenditure survey are costly, both in term of administrative efforts and money. Hence, they are only conducted every 5 years or even less frequently in most African countries. Modeling of poverty may serve as an alternative for obtaining poverty estimates at regular intervals of time, without undertaking expenditure surveys. Poverty / welfare predictor models developed from Integrated Surveys or Household Budget Surveys are already used with Census data to produce poverty estimates for small areas.

The WB presented a paper reviewing experiences using these poverty predictor models to rank survey households into quintiles.

SN presented a new model with the objective to estimate and monitor national and sub-national poverty headcount ratios from annual “light surveys” conducted between rounds of the IS or HBS. The paper presented and discussed an approach to estimate the uncertainty (sampling error) of the estimated poverty headcounts produced from these models.

IFPRI presented preliminary results from an assessment in Mozambique on how well data from "light surveys" could be used to predict poverty based upon an earlier model without error terms. This validation study found that the model predicted poverty incidence reasonably well at the national level, but that sub-national performance was erratic.

Another presenter introduced household consumption panel datasets as a supplement to data from standard household surveys that are repeated with intervals of several years. This would yield a more regular flow of information on the evolution of poverty through time and enhanced information on the characteristics of poverty, such as transitions in and out of poverty.

Finally, a speaker from SN presented statistical theory for small area estimations. The session recommended further work to continue developing statistical models for poverty monitoring.

4.2 Discussion

It was questioned whether any country uses statistical poverty indicator models for estimation of their official measures and argued that it was better to rely on (more secure) frequent standard household budget surveys. The audience was then informed that Malawi has already started using statistical poverty indicator model estimates and that these estimates are presented as official statistics. But there are still challenges requiring further methodological work as for 1) Importance of fluctuations, 2) Changes in consumption patterns, and 3) Difference between identifying variables from HBS first, and then include in the CWIQ.

Some found the Nigeria and Mozambique cases too detailed with 39 and 11 models. There is a need more general model, using other criteria than R^2. The model approach was addressed in general and some found it relying too heavily on difficult accessible statistics. Why not a mixed model or a random coefficient model? That would be more flexible and avoid the problem of imposing the model to the data. Better to use methods that let the data talk.

But these questioned were again addressed and rejected. In Nigeria all states requested their own model. In fact this model does exactly what is demanded, it listened to the data and develop a model which fits the data. That will often lead to slightly different model parameters in different regions of a country and in urban versus rural areas. But it should be stressed that these are not a large number of models. In one country there is one model, but for each part of a country, state or province and urban versus rural areas different parameters will be estimated. Hence in the future we should rather talk about one model in different versions.

It was explained that not only correlates are used, but also theory in selecting the indicators. Useful with different "black boxes" to compare. The number of model-versions just reflects the actual differences in the economic reality. Random coefficients are one particular way of adjusting for changes. Most countries require several versions of a model due to consumption patterns. In the spirit of let the data talk, it is recommended using the same sub-division as the actual stratification of the data.

One speaker found the estimates from light surveys as of current too unreliable, and still not ready for policy advises. The problem may be the original HBS estimate. Too many shortcuts? Not sufficient with only a few visits, or even single visits during each month.

A addressed that this is work *in progress* , hence it is equally important to identify issues still needing further work as issues where the best approach is already identified.

A justification for panel data was requested, it is quite demanding for field work and is hardly utilized for analysis.

The small area estimation techniques remained unclear, but it was then presented as a parallel to poverty mapping and statistical poverty indicators models, as a model approach where a model was developed from a survey and the model applied upon census type of information. A general experience from poverty mapping is that estimates might be given for geographical areas and socio-economic groups down to around 10 000 households, for poverty indicator type of variables. The errors were said partly to be due to the design of the model and partly to be due to the variance of predictor.

Malawi, Zambia and Mozambique was requested to share the experience of tracking households for a panel survey. It was difficult to identify the general rate of households which were tracked. Changes were recorded both at household level and individual level affecting the composition of the household.

In Uganda a survey panel had been going on for several years with success. Variation in changes over those household was much lower than for "fresh" households. Experience had however showed that it was necessary to devote some resources to make sure to reach the same household.

Uganda also had a good track record in utilizing panel information to follow how households tended to move in and out of poverty. In Tanzania panel data had shown useful when trying to understand the impact of a number of policy changes. Analytically panel datasets have proved very useful. The need to add new household to avoid a sampling bias was however also stressed. The general experience from other countries varied, but worked out well in South Africa, Ghana, and Ethiopia.

One comment suggested that panel surveys should be carried out by other institutions than the national statistical offices in order not to compromise the main survey.

5 Data entry and data processing

5.1 Presentations

The following presentations were given:
1. Teleform. By Jim Otto
2. Documents (Eyes&Hands). By Per Schøning
3. Streamlined traditional manual keying. Dag Roll Hansen
4. Country experience in shortening process time by scanning, quality control, editing process, tabulation plans based upon pilot. By Berit Otnes
5. CWIQ approach to shortening process time by scanning, quality control, editing process, tabulation plans. By Jim Otto

Two software packages for data entry using optical scanners were presented to the workshop participants. The World Bank has developed a generic survey package (available on CD-ROM) for use in organizing and carrying out CWIQ surveys. The package includes a computer system for processing the generic CWIQ questionnaire. This system uses "*Teleform*" data capture software for data entry. Teleform includes modules for questionnaire design, scanning, the conversion of scanned images into data and operator verification of any errors detected during the conversion. The software supports optical mark reading (OMR) and optical character reading (OCR) including hand printed data. The generic system also includes functions for data base update, correction and validation and functions for data analysis such as data summaries, cross-tabulations and the calculation of sampling errors. In every survey, the generic system needs to be adapted by an experienced programmer for changes to the generic CWIQ questionnaire. Since 2001, The World Bank has conducted workshops for independent African survey consultants and national statistics office data processing staff to increase the capacity of countries in the Africa Region to assume responsibility for adapting the data processing system.

The software "*Documents*" (Eyes&Hands) has been used by Statistics Norway for general scanning purposes since 1994. In recent years it has been utilized in Uganda, Angola and Malawi. Optical Character Reading serves to speed up the data entry process in a cost efficient way, and facilitates the organization of the data entry procedure.

In addition, experiences from a well-organized and successful use of *traditional* data entry from INE/Mozambique were presented for the participants. This approach used the software CSPro,

Depending of the size and other characteristics of the survey at hand, it was concluded that both traditional data entry and scanning might be feasible methods for data entry. However, data entry is just one component of any survey data processing system and it must be well integrated with the other components. There is no benefit from having a highly efficient data entry system if there are no provisions for timely data validation, correction and analysis.

5.2 Discussion

Unfortunately no minutes were taken from the discussions. The presentations and the discussion was followed by Jim Otto administering a hands-on-experience for the participants with filling in forms which were then scanned. The error rates reported the following day turned out very low.

6 Data analyses and dissemination

6.1 Presentations

Three presentations were given, as follows:
1. CWIQ approach to timely reporting and dissemination. Richard Harris and Tim Marchant
2. Tracking Resource and Policy Impact. By Bjørn K. Wold
3. DevInfo - National Socio-Economic databases. John Gitau Mburu presented KenInfo followed by presentation of DevInfo from Mozambique.

SN presented a comprehensive study of the tracking of resources and policy impact in Malawi. This covered the education, health, water and sanitation, agriculture and urban informal sectors. The presentation focused on the use of indicators to make the linkages between the different levels in the results chain of input-output-outcome and to monitor trends. Eventually, it leads to the tracking of the impact of resources on national policy goals and targets. A similar report is currently being made for Uganda.

Kenya and Mozambique presented their national socio-economic databases. The databases are based on the UNICEF ChildInfo software v 3.5, but have been further developed in order to better cover national needs. The development of the databases and integration of new indicators is facilitated by a Technical Working group chaired by the NSO. The presentations highlighted the user-friendly software application within the database, based on Microsoft Office, for easy presentation of data in tables, graphs and maps

A short plenary discussion was held on the possibilities and obstacles for fast dissemination of results within the CWIQ/QUIBB approach. The general conclusion was that summary tables could be produced within a few weeks the end of field work, but that very often, long delays still occur before final reports are available. Such delays are usually caused by political, technical, resource and management constraints.

6.2 Discussion

The need to ensure and document the quality control of indicators was highlighted. It was told that while DevInfo did not include a system of quality control as such, discrepancies in data for the same indicator from different sources were highlighted by flagging.

The lack of an option for adjusting the MDG indicators to national versions in the DevInfo type of databases was addressed.

The set of indicators was supported, but there was a demand to include others, such as follows:
- an indicator for good governance
- an indicator for the *quality* of education
- an indicator for corruption

Both the presenters and other participants provided information that work is already ongoing on several of these issues, while still not implemented.

The need for timely reporting and dissemination using the QUIBB was stressed. The system was designed for summary tables to be released within weeks, why the delays? Who prepares the draft report? What are the formats of presentation? Are layout-professionals consulted? What about clearance procedures? What kind of dissemination is used, workshops, bulletins, flyers etc. ?

A number of the countries could reprt that tabulat ion reprts with methodol ogical information were poduced fast, but months could be needed for analysis reprts. Currently there was a shift towards a plicy for fast release. Bt a major problem was the lack of money available for report writing. Statisticians prefer to go to the fieldwork to get allowances. Report writing workshops was said to be the best approach, but some countries were introducing a bonus for report writing. Some highlighted that the need for approval from Washington contributed to the delays. Others stressed that trained staff is scarce on the ground and that they are over loaded with work even before the reports should be written. The QUIBB philosophy for short reports, flyers and fast release was reiterated.

The approval process was addressed and the likelihood of changing released figures. Some countries had changed already released data, but the threshold for doing this was really high in order not to confuse users. Small changes were however said to be more normal and did not confusing for users. The need to make data available for intermediate users i.e. researchers was addressed but not followed up.

A summary pointed to three types of constraints: political, technical problems, and resource and management constraints. They all need serious consideration and follow up. DevInfo is increasingly used for data dissemination on the continent. A working group may be needed for further planning and harmonization of data analyses.

7 The International Household Survey Network

7.1 Presentations

Two presentations were made:
1. Household Survey Network. By Olivier Dupriez
2. Demonstration of Micro Data Management Toolkit. By Olivier Dupriez

The WB gave a presentation on the International Household Survey Network (IHSN), which has been developed as a response to the demand at the following two international roundtables:

- A 2002 high level roundtable in Washington to improve monitoring highlighted the need for an action plan.
- The following 2004 high level roundtable in Marrakech initiated the Marrakech Action Plan for Statistics (MAPS). Here the IHSN was created with a steering committee including high level representatives DFID, ILO, PARIS21, UNICEF, UNSD, WB, and WHO with a secretariat initially in WB.

The justification for the IHSN is that household surveys are necessary but also complex, uncoordinated, expensive and underutilized. Hence IHSN focuses on activities increasing the availability of data, reducing duplication of work, conflicting agenda and competition of resources.

The two main activities of the IHSN are:

- An information system on household surveys. This information system is already in the pipeline. The objective is to develop a databank of recommended questionnaires modules. This should be a central survey catalogue i.e. a web based meta-data base (a list of surveys & censuses) with information about: size, when and where.
- Development of a micro data management tool kit i.e. a tool for documenting, preserving and dissemination of micro data according to international standards, including a best practices archive module picking ideas from: UK data archive, NSD - Norway and ICPSR university consortium - USA

The rationale for the tool kit was the huge investment behind each survey for data collection but the lack of investment in:

- Micro data long term preservation,
- Documentation of micro data and
- Micro data dissemination

Thus, the World Bank has developed a tool for best practices. The tool has two elements: A meta-data editor and a CD-ROM builder. Albeit still at beta test stage, international and national NSOs and research institutions may already use this tool for data documentation. The toolkit is actually in use in Sierra Leone, Bosnia Herzegovina, Tunisia and Ethiopia. Refer to www.surveynetwork.org/toolkit or send an e-mail to odupriez@worldbank.org to get a CD-Rom with the software. The toolbox may also be used to generate a tabulation report for a survey such as the QUIBB.

7.2 Discussion

The initiative of the IHSN was appreciated and supported, and information on who maintains the database and how the information channels are planned was requested. IHSN informed that the international organizations rather than the countries are to provide information. The audience found this a short coming, but IHSN was clear that including country level information would be too demanding.

8 Group discussions on challenges and recommendations

During the last session of the workshop, there was discussion of the thematic building blocks within sub-groups, which then reported back with their recommendations to the plenary meeting. The groups were as follows:
Group 1: The household survey system with a core survey
Group 2: Statistical models
Group 3: IT: data processing & archiving

8.1.1 Recommendations by group 1: The household survey system with a core survey

Recommendations by group 1 were as follows:
1. NSO should be involved in PRSP process because they deliver data put also budget lines for statistics
2. Involvement of stakeholders is essential (Education of stakeholders included)
3. Training/capacity building on analysis and reporting/dissemination is needed
4. Cross cutting and statistical advocating need to be addressed.

8.1.2 Discussion and recommendations by group 2: Statistical models

Statistical modeling is in progress in most countries, there is a need for a work group and a need for guidelines. The black box should be opened. The following issues need to be addressed:
1. Follow up on analyses
2. Harmonize modeling
3. Incentives for data analysis
4. Basic report, seeking of clearance, dissemination process (analysis and dissemination included in the budget)

8.1.3 Discussion and recommendations by group 3: IT: data processing & archiving

Recommendations by group 3 were as follows:
1. Cooperation across departments concerning data processing during the planning - a survey team
2. Developing a network between across countries for use of technology exchanging experiences
3. Adopt long-term documentation of data and strategies for micro data dissemination

The general recommendations from the groups discussions were as follows:

1. Create a work group on common statistical methods for modeling (explain the black box separately for the specialists and the laymen)
2. Survey budgets should include dissemination and analysis (including incentives and motivation).

9 Final recommendations

9.1 Recommendations

The main recommendations from the workshop concerning further work on light core surveys such as the CWIQ and household survey programmes for PRSP monitoring were as follows:

On household survey programmes & the role of light core surveys:
- NSOs should be more directly involved in the PRSP process.
- Stakeholder involvement and guidance in statistics is crucial.
- There is a need for more training and capacity building both in analysis, and in reporting and dissemination.
- NSOs need to be more involved in cross cutting coordination and statistical advocacy with the national planning authorities to ensure an optimal use of statistical expertise,
- NSOs need to promote the use of indicators such as access, use and satisfaction with service delivery, to provide early leading indicators for PRSP goals and targets.

On IT and data processing:
- Establish a network across countries for the exchange of technical and methodological experiences
- Improve and standardize documentation of data & develop strategies for micro data dissemination
- Ensure a close cooperation across departments to involve IT-specialists all the way from the planning stage, through implementation and up to final the data documentation.

On modeling and analyses:
- Statistical modeling for improved monitoring of national PRSPs is in progress in most countries. However, there is a need for harmonization and guidelines, preferably through:
    - Follow-up on analyses
    - Harmonized modeling approaches
    - Stronger incentives for data analysis

The participants expressed a demand for a follow up workshop within 1-2 years.

9.2 Closing remarks

President João Dias Loureiro of INE emphasized *the* need for follow-up on this seminar. Survey tools should be further improved. Official statistics should be a public good. We need to address the broader group of users. *Mr Richard Harris from the World Bank and Mr Jon Krold from Statistics Norway* thanked INE for hosting the workshop and participants for the presentations and the discussion. Upon request from the users they expressed a commitment for taking an initiative for a follow up work shop in around 2 years time.

9.3 Report for downloading from web-site

This report may be downloaded from the website of Statistics Norway www.ssb.no/en/int

10 Appendix I. Program for the Workshop

Time and Session	Activity/Presenters
December 12th	
8.30-9.00	Registration
9.00-9.15	Official opening. Joao Dias Loureiro, President of INE
9.15-17.00 Session 1: Household survey approaches to monitoring PRSP - country experiences	1. A sustainable Household Survey System with a Core Survey for Poverty Monitoring for PRSPs and MDGs. By Bjørn K. Wold 2. Country experiences. By Country participants 3. The QWIQ survey designed as core of a household survey program. By Tim Merchant 4. Country experiences. By Country participants 5. Household Survey Programmes - reflections on the Report for PARIS21 Task Force on Improved Statistical Support for Monitoring MDGs:"Household Surveys and the Millennium Development Goals". By Richard Harris 6. Country experiences. By Country participants 7. Discussion
December 13th	
8.30-11.00 Session 2: Statistical models, estimation of poverty and small area estimation	1. Poverty Correlates to Rank Households and Present Results by Quintiles. By John Ngwafon 2. A statistical model for simple, fast and reliable measurement of poverty. By Astrid Mathiassen. 3. How well do light surveys predict poverty? Preliminary results from a validation study. By Kenneth Simler 4. Poverty Monitoring – lessons from Mozambique. By Arndt Channing 5. Large samples versus small area estimation for district level. By Stein Opdahl 6. Discussion
11.20-15.10 Session 3: Data entry and data processing	1. Teleform. By Jim Otto 2. Documents (Eyes&Hands). By Per Schøning 3. Streamlined traditional manual keying. Dag Roll Hansen 4. Country experience in shortening process time by scanning, quality control, editing process, tabulation plans based upon pilot. By Berit Otnes 5. CWIQ approach to shortening process time by scanning, quality control, editing process, tabulation plans. By Jim Otto 6. Discussion
15.30-17.00 Session 4: Analyses & dissemination	1. CWIQ approach to timely reporting and dissemination. Richard Harris and Tim Marchant 2. Tracking Resource and Policy Impact. By Bjørn K. Wold 3. DevInfo - National Socio-Economic databases. By presenters from Kenya and Mozambique 4. Discussion
December 14th	
8.30-10.20 Session 5: Household Survey Network	1. Household Survey Network. By Olivier Dupriez 2. Discussion 3. Demonstration of Micro Data Management Toolkit. By Olivier Dupriez 4. Discussion
10.40-12.30 Session 6: Final discussions and conclusions	1. Group discussion on challenges & recommendations 2. Groups present reco0mmendatiuons to plenary - discussion 3. Closing Remarks

11 Appendix II. List of Presentations

11.1 Authors and formats of presentations

Title	Format	Presenter,Institution/Country
Eight Elements Forming a Sustainable Poverty Monitoring System for Policy Decisions	Ppt	Bjørn K Wold, Statistics Norway
The CWIQ survey, designed as core of a household survey program	ppt	Tim Marchant, World Bank
Household Survey Programmes for MDG monitoring, based on Household Surveys and the Millennium Development Goals by J. Munoz for PARIS21, at PARIS21 web-site: http://www.paris21.org/documents/1181.pdf	adobe	Richard Harris, World Bank
Kenya: Role and objectives of a 'light' core survey within a household survey program – presentation	Ppt	Kenya
Kenya: Role and objectives of a 'light' core survey within a household survey program – paper	word	Kenya
Light Core Monitoring Surveys	Ppt	Lesotho
Malawi Welfare Monitoring Survey 2005, Presentation 1	Ppt	Malawi
Malawi Welfare Monitoring Survey 2005, Presentation 2	Ppt	Malawi
Light Core Monitoring Surveys and Poverty Monitoring - the Sierra Leone Experience	Ppt	Sierra Leone
Enhance evidence-based decision-making in the key development areas of Poverty and HIV/Aids in South Africa	Ppt	South Africa
Zambia household survey program	Ppt	Zambia
Poverty Correlates/Predictors	Ppt	John Ngwafon, World Bank
A Statistical Model for Simple, Fast and Reliable Measurement of Poverty	Ppt	Astrid Mathiassen, Statistics Norway
How well do light surveys predict poverty? Preliminary results from a validation study	Ppt	Kenneth Simler, IFPRI
Poverty Monitoring: Lessons from the Experience in Mozambique	Ppt	Arndt Channing, World Bank
Large samples versus small area estimation for district level	Ppt	Stein Opdahl, Statistics Norway
Ghana Census based Poverty Maps	word	Ghana
Measuring poverty in South Africa	Ppt	South Africa
Simple Household Poverty Assessment Models for Malawi	word	Malawi
Data entry by scanning	Ppt	Per Schøning, Statistics Norway
Streamlined traditional manual keying	Ppt	Dag Roll-Hansen, Stat. Norway
Rwanda Experience and Practices in Monitoring National PRSP	Ppt	Rwanda
Tracking Resource and Policy Impact – a general approach and as applied in Malawi	Ppt	Bjørn K Wold, Statistics Norway
KenInfo database	word	Kenya
Adaptao da Technologia DevInfo para Mozambique	Ppt	Mozambique
Programmatic Approaches to Planning and Implementing Statistical Programs	Ppt	Louise Fox, World Bank

11.2 List of presentations annexed:

12 Appendix III. List of Participants

Nr.	Country/ Inst.	Name	Email address	Position
1	Angola	Mr. Kodi Samba Constantino	kodisamba@ine.gov.ao	Head of Depart of Social and Demogr. Stat
2	Angola	Mr. Zatandu Mbiki	mbiki@ine.gov.ao	Head of Department of IT
3	Denmark	Mr. Jesper Ellemose Jensen		Coordinating Officer / SCANSTAT
4	Ethiopia	Mr. Zenaslase Siyum Gebremedhin	shemuka@yahoo.com	
5	Ethiopia	Mr. Shimeles Mulugeta Gebreselassie	shemuka@yahoo.com	
6	Ghana	Mr. Stephen Narh	natatch@yahoo.com	Regional Statistician Great Accra
7	Ghana	Mr. Marfo Yentumi	mynkansah2000@yahoo.com	Regional Statistician Eastern Region
8	IFPRI	Mr. Ken Simler		Consultant Mozambique Min of Plan. & Dev.
9	Kenya	Ms. Mary Mildred Wanyonyi	mwanyoyu@cbs.go.ke	Economist/Statistician, poverty research & analysis
10	Kenya	Mr. John Gitau Mburu	jgmburu@cbs.go.ke	Economist coordinator KenInfo
11	Kenya	Mr. Josiah Waithaka Kaara	kaara2002yahoo.com	Economist, WMS I & II
12	Lesotho	Ms. Khahliso Mokati	k.mokati@bos.gov.ls	Senior Statistician Division of Economic Statistics
13	Lesotho	Mrs. Masentle Mphutlane	m.mphutlane@bos.gov.is	Senior Statistician Depart of Demography
14	Lesotho	Mr. Rampa Motloheloa	r.motloheloa@bos.gov.ls	IT Specialist
15	Malawi	Mrs. Berit Otnes	berit.otnes@malawi.net berit.otnes@ssu.no	Senior Statistical Adviser
16	Malawi	Ms. Mercy Kanyuka	nkanyuka@statistics.gov.mw	Deputy Commissioner
17	Malawi	Mr. Kelvin Saukila	ksaukila@statistics.gov.mw	Principal Statistician
18	Mozambique	Mr. João Dias Loureiro	joao.loureiro@ine.gov.mz	President
19	Mozambique	Mr. Manuel Gaspar	manuel.gaspar@ine.gov.mz	Vice-President
20	Mozambique	Mr. Arão Balate	arao.balate@ine.gov.mz	Director of Census and Survey
21	Mozambique	Mrs. Fatima Zacarias	fatima.zacarias@ine.gov.mz	Director of Demographic and Social Sta
22	Mozambique	Mr. Cassiano Chipembe	cassiano.chipembe@ine.gov.mz	Head of Demographic Studies
23	Mozambique	Mrs. Anastacia Honwana	anastacia.honwana@ine.gov.mz	Head of IT Department
24	Mozambique	Mr. Cristovão Muhaio	cristovao.muahio@ine.gov.mz	Head of Depart of Methodology and Sampling
25	Mozambique	Mr. Xadreque Maunze	xadreque.maunze@ine.gov.mz	Statistician
26	Mozambique	Mr. Elisio Mazive	elisio.mazive@ine.gov.mz	Statistician
27	Mozambique	Mr. Eugénio Matavel	eugenio.matavel@ine.gov.mz	IT Specialist

28	Mozambique	Mr. Anselmo Nhane	anselmo.nhane@ine.gov.mz	IT Specialist
29	Mozambique	Mr. António João Azize		Min. of Labor
30	Mozambique	Mrs. Cristina Matusse		Min. of Plan. and Dev.
31	Mozambique	Mr. Rogério João Nkomo		Head of Dep. of Studies & Stat. of Finance Public/Min. of Fin.
32	Mozambique	Mr. Ilídio Buduia	ilidio@mec.gov.mz	Head of Depart / Ministry of Education and Culture
33	Mozambique	Mr. António Vasco Sitói		Head of Depart of Information of Health/Ministry of Health
34	Mozambique	Mrs. Carla Matos	cmatos@dnsdee.misau.gov.mz	Head of Non Communicable Disease
35	Mozambique	Mr. Fernando Casimiro	fernando.casimiro@ine.gov.mz	Team Leader of Italian Project
36	Norway	Mr. Bjørn K. Wold	bkw@ssb.no	Head of Division for development cooperation
37	Norway	Mr. Geir Øvensen	gov@ssb.no	Senior Statistical Adviser
38	Norway	Mr. Stein Opdahl	sop@ssb.no	Senior Statistical Adviser
39	Norway	Mr. Per Schøning	per@ssb.no	Senior Statistical Adviser
40	Norway	Mrs. Astrid Mathiassen	mss@ssb.no	Senior Statistical Adviser
41	Norway	Mr. Dag Roll-Hansen	dag.roll-hansen@ine.gov.mz	Adviser at INE Moz/ SCANSTAT
42	Prude Univ. EUA	Mr. Arndt Channing	cardit@mpd.gov.mz	Consultant MPD
43	Rwanda	Mr. Louis Munyakazi	mzrlouism@yahoo.com	Director General, National Statistical Institute of Rwanda
44	Rwanda	Mr. Augustin Twagirumukiza	twaugustin@yahoo.com	Head of Surveys IT & Data Processing
45	Sierra Leone	Mr. Sheikh Rogers	trogers@statistics.sl	Manager, Special Projects and Analysis
46	Sierra Leone	Mr. Mohamed Koroma	mkkoroma@statistics.sl	Director of Data Processing
47	South Africa	Mrs. Annika Puide	annikap@statssa.gov.za	Adviser at Statistic South Africa (SCB)
48	South Africa	Mrs. Nthabiseng Makhatha	nthabisengm@statssa.gov.za	Manager, Social Statistics (Stasa)
49	Sweden	Mr. Lars Carlsson	lars.carlsson@ine.gov.mz	Team Leader at INE Moz / SCANSTAT
50	World Bank	Mr. Richard Harris	rharris@worlbank.org	SNR Statistician
51	World Bank	Mr. John Ngwafon	jngwafon@worldbank.org	MDE Specialist
52	World Bank	Ms. Louise Fox	lfox@worldbank.org	Lead Economist Poverty
53	World Bank	Mr. James Otto	jotto@worldbank.org	CWIQ DP Consultant
54	World Bank	Mr. Tim Marchant	tjmarchant@btinternet.com	Consultant
55	World Bank	Mr. Olivier Dupriez	odupriez@worldbank.org	Senior Economist/Statistic Development Data Group
56	World Bank/ Mozambique	Mrs. Katleen Van den Broeck	kvanenbroeck@worldbank.org	Consultant (Poverty Economist)
57	Zambia	Mr. Frank Kakungu	fkakungu@zamstats.gov.zm	Head of Living Conditions Monitoring Unit
58	Zambia	Ms. Nchimunya Nkombo	nkombon@yahoo.com	Head of Social Statistics

13 Appendix IV. Summary of Participants Evaluations of the Workshop

Table 1. Respondents by sex

	Male	Female	Total
Number	19	8	27

Table 2. Workshop ratings of logistics and proceedings

	Number of responses					
	1	2	3	4	5	
Dimension:	Unsatisfactory…-…Excellent					**Average score**
Logistics:						
Transportation[(1)]	0	0	3	6	10	4.4
Accommodation[(1)]	0	0	4	4	11	4.4
Meals& refreshment	0	1	3	14	9	4.1
Conference facilities[(2)]	0	0	3	13	10	4.3
Workshop Proceedings:						
Overall quality of presentations	0	0	2	19	6	4.2
Provision of documents / papers	0	1	4	12	10	4.2
Structure of workshop agenda	0	1	3	19	4	4.0
Time management of sessions	0	2	6	15	4	3.8
Facilitation & chairing	0	1	6	12	8	4.0
Effectiveness of discussions	0	0	6	16	5	4.0
Relevance of workshop	0	0	2	11	14	4.4
Overall satisfaction	0	0	2	13	12	4.4

(1) Not applicable for 8 participants from Mozambique
(2) 1 participant did not respond

Table 3. Overall satisfaction by workshop

Do you think the workshop achieved its objectives?			
	Yes	**No**	**Total**
Number	27	0	27

Table 4. Rating of presentations

Which presentations did you find most interesting?	Number
All of them	4
Household survey approaches to monitoring PRSPs & country experiences	10
Statistical models for estimation of poverty and small area estimation	4
Data entry and data processing	5
Analysis & dissemination	1
Micro Data Management Toolkit	6
Total (Some respondents identified more than one presentation)	27

Table 5. Rating of modalities

List the 2 things you thought went best about the workshop?	Number
All	1
Exchange of country experiences & interaction	7
Group work and participation in discussions	11
Facilitation & organization	10
Quality & relevance of presentations & handouts	17
Total (Some participants did not respond to this question or only gave one answer)	46

Table 6. Rating of dissatisfaction

List the 2 things you were least happy with about the workshop?	Number
Time too short / time too short for discussion / inadequate time for interaction & participation / time for wrap up was too short	14
Some sessions too long / too many presentations in a session	3
Lack of workshop folders / files / name tags	3
Lack of activities in the evenings / after the workshop	2
Room too small	1
Room too cold	1
No lunch on day 3	1
Planning of flights to Mozambique	1
Panel discussion on options & challenges	1
Total (Some participants did not respond to this question or only gave one answer)	27

Table 7. Rating of potential improvements

Please provide any additional comments and suggestions for improvement?	Number
INE did a great job	1
More frequent / regular / follow-up meetings	8
More time needed	2
Some presentations needed more time, others less	1
Minimize presentations, more time for discussions	1
Closing should have re-emphasized the objectives of the workshop	1
More participants	1
More on reporting, storage and dissemination	1
Better transport arrangement to venue	1
Total (Some participants did not respond to this section)	17

14 Appendix V Presentations

14.1. A Sustainable Household Survey Based Poverty Monitoring System

A Sustainable Household Survey Based Poverty Monitoring System for Policy Decisions, Eight Elements Forming a program to be implemented under the auspices of a National Statistical Institute

Presented at Workshop on Light Core Surveys, Maputo, Mozambique, 12-14 December 2005
by Bjørn K. Wold, Statistics Norway, Kongensgt.6 N-0031 Oslo, Norway, bkw@ssb.no
based upon the following report: Wold, B K, D Roll-Hansen, A Mathiassen, and S Opdahl:
A Sustainable Household Survey Based Poverty Monitoring System, Doc. 2004/17, Statistics Norway
http://www.ssb.no/english/subjects/00/00/30/doc_200417_en/doc_200417_en.pdf

Statistisk sentralbyrå
Statistics Norway
1

Introduction

Survey programs should respond to policy issues. Across the large majority of developing countries two overlapping sets of policy issues are defined by:

- The Global Millennium Development Goals
- The National Poverty Reduction Strategy Papers

Statistisk sentralbyrå
Statistics Norway
2

Eight Elements for a Household Survey System for Annual Poverty Monitoring

1. Identification of national Poverty Reduction Strategies and nationally adapted Millennium Development Goals.
2. A 12 month Consumption and Expenditure Survey
3. A Statistical Model to estimate poverty and hunger related measures and error-terms
4. A core household survey based upon scanning and fast turn around technology
5. A 5-10 year household survey program including the core survey
6. Methodology for a 5-10 year household survey program including the core survey
7. International agencies and donors to accept a system with National Core surveys and their “own” surveys as Modules
8. Fast and policy relevant dissemination of trend data

Statistisk sentralbyrå
Statistics Norway
3

Oito Elementos num Sistema dos Inquéritos aos Agregados Familiares para Monitorização Anual de Pobreza

1. Identificação dum Estratégia Nacional de Redução da Pobreza e uma versão nacional dos Objectivos de Desenvolvimento do Milénio
2. Um inquérito de Despesas e Receitas por 12 meses
3. Um modelo estatistico para estimar pobreza, medições relacionadas a fome, e termos de erro
4. Um inquérito nuclear monitorando bem-estar baseado por leitura optica e tecnologia rapida 'pré-fabricada'
5. Um programa de inquéritos aos agregados familiares por 5-10 anos incl. o inquérito nuclear
6. Metodologia para um programa de inquéritos aos agregados familiares por 5-10 anos incl. o inquérito nuclear:
   - Ajustamento por estação
   - Estmativas das áreas pequenas
7. Organizações internacionais e doadores a aceitar um sistema composta por um inquérito nuclear nacional e os inquéritos 'próprios' com módulos
8. Disseminação rapida dos dados sobre tendências sendo pertinentes para política Identificação da Estratégia do Redução da Pobreza Nacional e a Versão de Nacional das Tarefas Desenvolvimentos do Milénio.

Statistisk sentralbyrå
Statistics Norway
4

Huit éléments formant un système soutenable de surveillance de la pauvreté pour décisions politiques, une programme à implémenter sous les auspices d'un institut national de statistique

1. Identification de stratégies nationales de réduction de la pauvreté et buts de développement Millenium adaptés au plan national.
2. Une surveillance de 12 mois de consommation et dépenses.
3. Un model de statistique pour estimer mesures relatives à la pauvreté et la famine, et termes d'erreur.
4. Une surveillance de paramètres centraux de prospérité (Core Welfare Monitoring Survey) basée sur scanning et technologie de revirement rapide
5. Une programme de surveillance de ménage de 5 - 10 ans, incluant la surveillance de paramètres centraux
6. La méthode de surveillance de ménage de 5 - 10 ans, incluant ajustements saisonniers et estimâtes de petit secteur géographique
7. Acceptation des agences internationales et de donneurs d’un système comprenant un Core Welfare Monitoring Survey national et leur "propre" surveillance comme modules
8. Dissémination rapide de data de tendances actuelles (trend-data) relevant pour la politique.

Statistisk sentralbyrå
Statistics Norway
5

1. Identification of national Poverty Reduction Strategies and nationally adapted Millennium Development Goals.

A household survey system comprising:

- The Millennium Development Goals
- The national Poverty Reduction Strategy Papers

Statistisk sentralbyrå
Statistics Norway
6

2. A 12 month Consumption and Expenditure Survey

- A 12 month consumption and expenditure survey, household budget survey or integrated survey conducted every 5th or 7-8th year

3. A Statistical Model to estimate poverty and hunger related measures and errorterms

- An income poverty indicator model
- A statistical stochastic model
- Based upon relationship between total consumption (as best proxy for income) and any other variable which is easy to measure and does not require a fixed context
- Giving estimates of probability of falling below a poverty line, a hunger line, or of total consumption
- Giving estimates of error terms
- Requiring 10-15 indicators
  (half or more are already included among standard variables)
- A revised model to be estimated for every consumption and expenditure survey
- Presented in another contribution at this conference

4A core household survey based upon scanning and fast turn around technology

- An annual light core survey
- Including MDG indicators
- Including PRSP indicators
- Including variables for the income poverty indicator model
- Based upon World Bank CWIQ approach with rigorous pilot
- Scanning
- Verification at data entry stage,
- Pilot to determine validity checking
- Pilot to determine tabulation program
- 3 months turn around from completion of field work

5. A 5-10 year modular household survey program including the core survey

- Year 1: A cons. and expend. survey with a core module
- 2-4 years: Sector surveys with core module
- 0-2 years: Self standing core survey
- Possible sector surveys with a core module:
  - Core survey + Labor Force Survey supported by ILO,
  - Core survey + Demographic and Health Survey (DHS) supported by USAID
  - Core survey + Multiple Indicator Survey supported by UNICEF,
  - Core survey + Demographic or Education survey of DHS type, supported by USAID
  - Core survey + Living Standard Measurement Study survey or an Integrated Household Survey supported by the World Bank,
  - Core survey + Informal sector survey, or
  - Core survey + Agricultural Survey for the small and medium scale farmers etc.

6. Methodology for a 5-10 year household survey program including the core survey

- Year 1: A cons. and expend. survey with a core module
- 2-4 years: Sector surveys with core module
- 0-2 years: Self standing core survey
- Core survey
  - scanning and predetermined tabulation plan, 3 months turn around
  - 2-3 months field work in one fixed season
  - large sample size for district estimates or smaller sample size with small area estimation
- Sector module surveys
  - traditional processing, 12 months turn around
  - 3 mon in any season or 12 mon
- **Requirements:**
  - **Seasonal adjustment**
  - **Small area estimates**

7. International agencies and donors to accept a system with National Core surveys and their "own" surveys as Modules

Refer to the next general presentation in this workshop and the report to the PARIS21 Task Force on Improved Statistical Support for Monitoring Development Goals, J Muñoz, J and K Scott: *Household Surveys and the Millennium Development Goals*

The document is still (December 2005) with quoting restrictions, but may be downloaded from:
http://www.paris21.org/documents/1181.pdf

8Fast and policy relevant dissemination of trend data

Not only data base publication but systematic presentation, such as by four steps:

- **INPUT DATA government finance statistics**
- **OUTPUT DATA sector administrative records as school attendance**
- **OUTCOME DATA survey data on status and achievement such as literacy and school completion**
- **IMPACT DATA survey data on poverty and hunger**

Refer to another presentation in the workshop and to

Wold, B K, S Opdahl, E Rauan, R Johannesen, I T Olsen, Report 2004/20: Tracking Resource and Policy Impact, Oslo, Statistics Norway http://www.ssb.no/english/subjects/00/00/30/rapp_200420_en/

Wold, B K, M Kanyuka, E Rauan, M Yute, M Mkwemba, S Opdahl og R Johannessen, Report 2005/27: Tracking Resource and Policy Impact in Malawi. Incorporating Malawi Poverty Reduction Strategy Paper Indicators, Millennium Development Goals and Poverty Monitoring Across Sectors, Oslo, Statistics Norway http://www.ssb.no/emner/00/00/30/rapp_200527/

Eight Elements Forming a Sustainable Poverty Monitoring System for Policy Decisions, a program to be implemented under the auspices of a National Statistical Institute

1. Identification of national Poverty Reduction Strategies and nationally adapted Millennium Development Goals.
2. A 12 month consumption and expenditure survey
3. A statistical model to estimate poverty and hunger related measures and error-terms
4. A Core Welfare Monitoring Survey based upon scanning and fast turn around technology
5. A 5-10 year household survey program including the core survey
6. Methodology for a 5-10 year household survey program including the core survey
   - Seasonal adjustment
   - Small area estimates
7. International agencies and donor to accept a system with a national Core Welfare Monitoring Survey and their "own" surveys as modules
8. Fast dissemination of policy relevant trend data

Remaining Challenges

- Test assumptions and reliability of statistical model in other countries
- Adapt and document standard approach for seasonal adjustments
- Adapt and document small area estimation methods
- Adapt other standard surveys to core survey + modular survey
- User-friendly dissemination

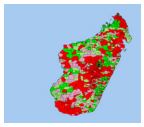
- We ask for cooperation, contact Bjorn K. Wold, Statistics Norway BKW@SSB.NO

14.2. The CWIQ Survey

The CWIQ survey, designed as core of a household survey program

Enhancing Statistical Capacity for Monitoring Development Goals

POPULATION CENSUS (Every 10 years)

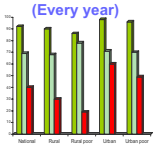


Poverty maps for locating the poor

INTEGRATED SURVEYS (Every 5 years)

Poverty profiles for measuring poverty

CORE INDICATORS SURVEYS (Every year)



Tracking access, use & satisfaction with public services

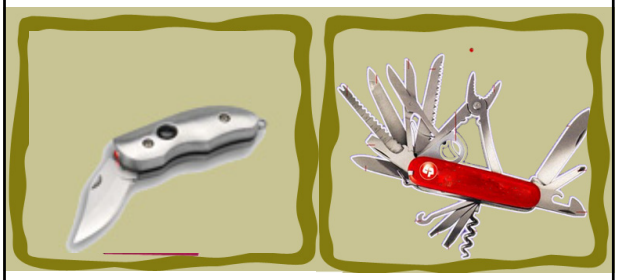
Plus Participative Poverty Assessments for Listening to the Poor

http://www.worldbank.org/AFR/STATS/

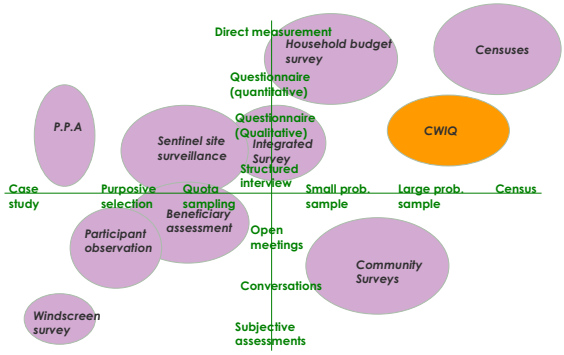
The CWIQ survey, designed as core of a household survey program

1. A discussion on tools
2. Monitoring the PRSP and the MDGs – the role of the core survey
3. What information should the core survey collect?
4. Examples of leading indicators and their use
5. A discussion on measuring poverty

The right tool for the job



Strengths and weaknesses of different surveys



Define the Job

A framework for monitoring the PRSP, and the MDGs

Goal: Increase Growth and reduce poverty

Impact → Impact on living standards - Lower poverty; reduced prevalence of malnutrition

5-10 yrs

Outputs → Goods and services generated

Annual

Inputs → Resources (financial, material and human) provided

Annual

The MDGs

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

Leading Indicators

- These may include:
  - Measures of early response to policy initiatives
  - questions on attitudes, opinions, levels of satisfaction
  - Measures of service delivery (awareness, access, use & satisfaction)
- Examples
  - Heath:
    - Goal: to reduce infant mortality
    - Leading indicators : Awareness and use of pre/post natal clinics, immunization rates etc.)
    - Supplementary info: reasons for non-use, satisfaction with services delivered
  - Agriculture
    - Goal: to increase production by (X) %
    - Leading Indicators: Use (and repeated use) of ag. Services (extension, fertilizer, credit etc.)
    - Supplementary: reasons for non-use, satisfaction with services delivered

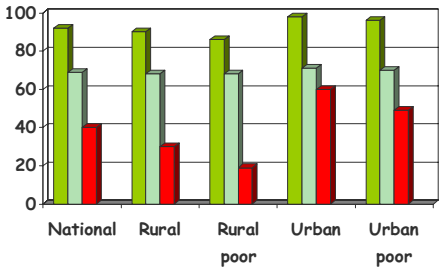
Using CWIQ to monitor economic wellbeing in Nigeria

Household economic situation compared to one year ago

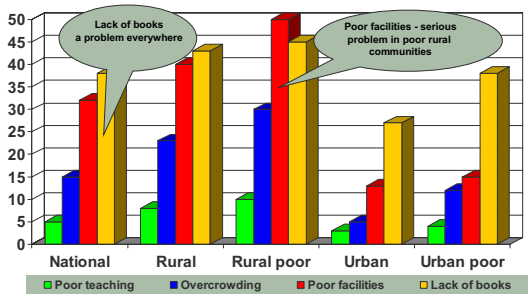
	Benue total	Margin of error (+/-%)	All Rural	Rural poor	All Urban	Urban poor
Worse now	27.6	5.5	30.6	23.4	14.6	23.8
Better now	49.5	5.4	46.7	59.2	61.5	65.7

Using CWIQ to monitor primary education in Ghana

- *Access to schools (within 30 minutes)*
- *Usage (enrollment rates)*
- *Quality of service (% satisfied)*



What are they complaining about?



What data does the CWIQ provide?

- Descriptive information about the household or individual (hh. size, SEG, asset ownership, welfare group, gender, age, educational attainment.
- Leading indicators linked to PRSP goals and MDG's (access, use and satisfaction)
- Used for generating time-series and making comparisons between population groups

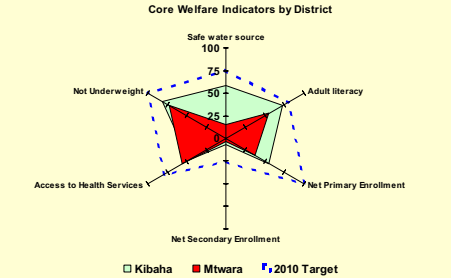
CWIQ generic questionnaire content

Interview information
Household members
Education
Health
Employment
Household assets
Household amenities
(Poverty Predictors)
Children under 5

CWIQ modules

- Gender Module (Nigeria)
- Flood damage module (Mozambique)
- CWIQ/MICS module (Mozambique)
- Consumption module (Rwanda)
- HIV/AIDS module (Lesotho)
- Mental Health module (Burundi)
- Community CWIQ (Tanzania)

Sample CWIQ outputs - Tanzania



A discussion about measuring poverty?

- Useful but risky
- Why is it useful?
  - ..as a classification/descriptive variable
  - ..as an outcome variable
- Why is it risky?
  - High potential error level
  - Easily misinterpreted
  - High political interest
- Options:
  - No measurement attempted
  - Use consumption correlates
  - Use asset ownership
  - Use light consumption model

Key messages

- A range of different monitoring tools are required. Purpose of the core survey is to provide early warning of changes in leading indicators over time and between groups
- Leading indicators include measures of access, use, and satisfaction
- Prompt and regular dissemination of results is essential to ensure feedback into budgeting and policy process
- Essential properties of the core survey include:
  - High quality data
  - Rapid results
  - Information that is easy to understand and use

14.3. Household Survey Programs

Household Survey Programmes

Regional Workshop on
CWIQ / 'light' Monitoring Surveys

Instituto Nacional de Estatistica (INE)
Mozambique,
Statistics Norway & the World Bank

Maputo, Mozambique
12th -14th Dec 2005

Presentation

- PARIS 21Paper on Household Surveys & the Millennium Development Goals
  - Role that household surveys play in the production of policy relevant information.
  - Main types of surveys; their purpose, use and characteristics.
  - Model household survey programme.
- *My own comments & observations*

Monitoring: Data Needs

INTERNAL

- Inputs
  Financial & physical resources provided
- Outputs
  Goods & services generated

EXTERNAL

- Outcomes
  Immediate benefits (access, use & satisfaction)
- Impact
  Impact on welfare & living standards

Information Sources

- Inter- linkages
  - Administrative – sectoral systems
  - Household surveys
  - Build the results chain
- Need to address both
  - Two halves of a coherent system
  - Previous experience shows unbalanced approach weakens system.

Role of Household Surveys

- Only source of some specific data required for policy making
  - Only source when administrative systems don't exist; a check when administrative systems are weak / incomplete e.g. infant mortality.
  - Only source for some indicators e.g. poverty incidence
  - Compliment administrative data
  - Identify impact, causality
- Active nature – reliable flow of new data within a reasonable period

What Makes Different Surveys Different ?

- Aims & Purpose (Table 1)
  - Monitoring progress (what?)
  - Understand & explain reasons behind the observations (Why?)
- Methodological differences (Table 2)
  - Sample size and geographic disaggregation
  - Frequency of data collection and reporting
  - Period of field operations
  - Duration of field operations & number of visits

What Makes Different Surveys Different? (cont.)

- Content (Table 3)
  - Most collect basic demographic data including basic data on education
  - Some MDGS can only be collected by certain surveys
- Resources (Table 4)
  - Depend on sample size but all have fixed costs so not as wide a range as might be expected
  - TA very different & country specific
- Quality control mechanisms
  - OK = Sampling theory, definitions of data needed, measuring techniques
  - Weak = qualitative & managerial aspects of implementation

Household Surveys & Monitoring

- No one single survey will meet all data needs – not feasible (Table 5)
  - Large samples for some indicators (rare events)
  - Small samples to reduce non-sampling errors for complex indicators
- Requires survey programme (Table 6)
- Monitoring not only goal of household surveys

Current State of Household Surveys

- Infrequently implemented
- Even less frequently used for policy making
- Variable quality

  -Timeliness, - Relevance

  -Accuracy - Cost-effectiveness
- Lack of comparability over time & surveys
- Lack of access by policymakers

Current state of household surveys (cont.)

- Vicious cycle
  - Under-performing statistical agencies produce no outputs that policy makers value
  - Policy makers have little incentive to increase budgets of NSOs
  - Policy makers do not develop an understanding of the potential value of data and have few skills to use data
- Need to better understand the causes.

What would be ideal?

- Goal

  To be able to respond to the evolving data needs of government & society
- Not to produce data X or survey Y but to have -
- CAPACITY to
  - Determine data needs as or before they arise
  - Identify appropriate tools to collect such data
  - Implement in a timely manner
  - Analyse data to generate valuable inputs for policy making process

Key needs

- Linking data producers & users.
- Providing adequate resources.
- Creating continuous Hhold Survey programmes
  - Create capacity
  - Economies of scale
  - Predictability will improve use of data
  - Defence against ad-hoc pressures though special surveys will be needed when the need arises
  - Will be more responsive and flexible as new data demands arise or priorities change

Generating a Hhold Survey System - Planning

- Plan within the development of the national statistical system as a whole
- Identify users and their information requirements (inc. those who would be users if the data was adequate /available).
- Assess
  - Alternative data sources / methodologies
  - Quality of different sources - Can this be improved? How?
  - Relative costs

Generating a Hhold Survey System - Process

- Lengthy
- Require substantial levels of consultation & negotiation
- Can itself be an educating & capacity building process
- Will need prioritization

Prototype Model

- A programme of Living Standard Measurement Surveys (LSMS) every 4 years
- A programme of Demographic & Health Surveys (DHS) every 4 years
- Core Survey
  - Suggested: A permanent quarterly labour force survey (LFS)

This would monitor all 27 household–based MDG indicators (Table 7)
Though monitoring MDG / PRSP indicators is not the only goal.

What does this mean for us?

- *Need for a better understanding of what works, what hasn't worked.*
- *Need to develop better linkages between producers / users*
  - *Build capacity on both sides, together*
  - *Within context of developing M&E system*
  - *Allow time*

What does this mean for us? (cont.)

- *LSMS / Integrated Survey or HBS + LFS*
- *DHS*
- *Core Survey*
  - *LFS in Africa context ?*
  - *Adds little to monitoring of MDGs/ PRSPs*
  - *Need survey with expanded sample to monitor 'what' for different household groups*
  - *CWIQ / QUIBB (Access, use & satisfaction)*

14.4. Role and objectives of a 'light' core survey within a household survey program - Kenya, i) PowerPoint presentation, ii) Word Document

Kenya

WORKSHOP ON 'LIGHT' CORE MONITORING SURVEYS

Hotel Cardoso, Maputo, Mozambique, 12th - 14th December 2005

Kenya

Role and objectives of a 'light' core survey within a household survey programme;

How is the survey programme and content set? Who decides on priorities? Who is consulted?

- Kenya has not focused on a light core survey but administered the broad indictor based Welfare Monitoring Survey series (WMS I-III). Currently, the country is undertaking the Kenya Integrated Household Budget Survey (KIHBS) expected to end in April 2006.
- Involve various subject matter specialists from key stakeholders. Ie Donor community, line ministries, research institutions such as universities, NGOs, etc. Further to this, the KIHBS questionnaire was posted in the CBS website for comments from the general public.

Kenya

Role and objectives of a 'light' core survey within a household survey programme cont'd;

What linkages do you have with the PRSP monitoring unit?

- Since 2003, when the NARC government came into power, they produced the Economic Recovery Strategy for Wealth and Employment Creation (ERS) which has three main pillars:
- Economic growth
- Equity and poverty reduction which mainly addresses the aspirations of the PRSP and the MDGs.
- Governance
- The MDGs are entrenched within the sectors.

Kenya

Role and objectives of a 'light' core survey within a household survey programme cont'd;

Are you involved in any of the discussions over the choice of PRSP indicators?

- Yes; In collaboration with the Monitoring and Evaluation Department.

How do you decide on which PRSP / MDG indicators to monitor each year, frequency, degree of dissagregation (sample size)?

- There are those monitorable ERS and MDG indicators which have been identified. These are incorporated within the household based surveys undertaken by the CBS. Examples include:
  - Access to safe drinking water and adequate sanitation
  - Literacy, gender equity etc

Kenya

Role and objectives of a 'light' core survey within a household survey programme cont'd;

Is there a programme for the monitoring of each PRSP / MDG indicator?

- There is a GOK department charged with the responsibility of M&E of the ERS and MDGS in the country. This is the MED in the MP&ND.

Do you find some indicators difficult to measure accurately?

- Yes; eg
- HIV/AIDS orphans. This is because the accurate number of deaths associated with HIV/AIDS
- Proportion of population below $1 (PPP) per day

Kenya

Role and objectives of a 'light' core survey within a household survey programme cont'd;

- Poverty gap ratio {incidence x depth of poverty}
- Proportion of population below minimum level of dietary energy consumption
- Proportion of households with access to secure tenure

Role and objectives of a 'light' core survey within a household survey programme cont'd;

Which surveys have been conducted in recent years or are planned for the immediate future?

- The ongoing 2004/05 KIHBS (includes a PPA component)
- The 2005 literacy survey - ongoing
- APRM (Governance) survey (completed)
- International Comparison of Prices (ICP)- ongoing

Planned surveys

- Agriculture and Livestock census survey
- Small and Micro-Enterprise (SME) survey
- Census of Industrial Production
- The 2009 Population and Housing Census
- Other planned activities are outlined in the CBS national strategic plan (see www.cbs.go.ke)

Role and objectives of a 'light' core survey within a household survey programme cont'd;

Have any 'light' monitoring surveys such as the CWIQ or WMS been carried out (year, content, sample size)

- CWIQ – piloted in 1997 preceding the WMS III.
- WMS I - 1992 (12,050 Households)
- WMS II - 1994 (10,860 Households)
- WMS III - 1997 (12, 580 Households and excluded the Arid and Semi-Arid Parts of Northern Kenya)
- KIHBS - 2004/05 (13, 430 Households)

Role and objectives of a 'light' core survey within a household survey programme cont'd;

How do you decide which PRSP indicators are appropriate to monitor using the 'light' core survey and which are not?

- N/A

How do you meet the needs of policy analysts who want cross-sectional information from households rather than indicators?

??????

How do you ensure consistency in definition and question wording across surveys?

- We maintain indicator definition within the household surveys for results comparability. The definitions comply with the international (UN) standards and domesticated to the country situation.

The use of statistical models to identify the poor and poorest households and to derive estimates for small areas;

- Do you use statistical modeling techniques to identify the poor or poorest households in your 'light' core monitoring survey?
- **Yes**
- How well have these worked?
- **Very well. In fact, Kenya has produced two volumes of the Poverty Monitoring Reports (The Geographic Dimensions of Well Being in Kenya).**

The use of statistical models to identify the poor and poorest households and to derive estimates for small areas;

the results believed and used?

- **Yes. The associated point estimate errors are very low. The results are used for identification and targeting of the poor. The government uses the poverty index statistics to allocate Constituency Development Funds (CDF) and the Constituency Education Bursary fund.**
- Do you use modeling techniques to provide estimate for small administrative areas (second administrative level below national level) e.g. if the government system is national – regional - district, then to provide estimates for districts and below

The use of statistical models to identify the poor and poorest households and to derive estimates for small areas;

- - Do you have a poverty map with estimates of poverty incidence for small areas?
- **Yes, at the sub-district levels. We have poverty estimates for the divisions, locations, sub-locations (urban) and constituency level.**
- - Are the results believed and used? **Yes**
- - Do you use the poverty map technique for anything other than poverty incidence?

The use of statistical models to identify the poor and poorest households and to derive estimates for small areas;

- **Yes; Trying to superimpose on the maps, other socio-economic variables such as IMR, Education, water, schools infrastructure etc**
-
- Have you conducted any large sample surveys to give estimates for small administrative areas?
- **Yes; WMS II**
- - What problems have you encountered?
-
- - Are you under pressure to conduct large sample surveys? For which topics?

Data entry and data processing for timely dissemination

How long between the end of field work and the producing final tables for any 'light' monitoring surveys? How does this compare to other surveys?

- Optical Mark reader scanning was undertaken for the first time ever during data capture of the 1997 WMS III. The process of data capture took long and was very tedious due to low quality paper and low precision of the mark readers in the questionnaires.

Data entry and data processing for timely dissemination cont'd

Were the questionnaires scanned? Has this technology been adapted for other uses within the Office?

- Scanning technology was used again during the 1999 Population and Housing Census. This time the character recognition scanning was successfully done.

Were the data processing system and tabulation programmes ready and tested before the end of field work?

- The tabulations as per the scanning process had already been put in place. However, they were never put into use since the World Bank withdrew funding midstream. The withdrawal of funding support culminated in CBS being unable to produce even a basic report from the survey.

Reporting & Dissemination

How long between the production of final tables and the publication of reports for any 'light' monitoring surveys? How does this compare to other surveys?

- The process of basic report production takes a relatively short time depending on the immediate availability of both financial and technical support.

Do you publish survey reports, statistical abstracts, summary reports, flyers?

- Yes; survey reports; flyers (ie 1998/99 labour force survey)

Reporting & Dissemination

Do you target reports for particular users?

- The popular version of a report is produced to cater for the less technical public

Do the survey results get into the PRSP annual progress report?

- The survey results feed into the M&E annual progress reports.

Do you do further data analysis done after the initial presentation of results?

- Further analysis is done. Ie poverty analysis

Do your indicators produce believable trends over time?

- Yes; ie WMS series

Kenya

Reporting & Dissemination

How are the data sets stored? Are your surveys archived?

- They are stored by the research, data processing and publication division. The division is charged with the responsibility of archiving the survey data sets and findings. There are also advanced plans to set up an IMIS in the department.

How do researchers get access to the datasets for further analysis?

- Requests have to be made to the Director of Statistics

Kenya

THE END

THANK YOU

WORKSHOP ON 'LIGHT' CORE MONITORING SURVEYS

Hotel Cardoso, Maputo, Mozambique, 12th - 14th December 2005

A. Role and objectives of a 'light' core survey within a household survey programme;

How is the survey programme and content set? Who de cides on priorities? Who is consulted?

Kenya has not focused on a light core survey but administered the broad indictor based Welfare Monitoring Survey series (WMS I-III). However, in 1997, a CWIQ was piloted in a few districts alongside the WMS III questionnaire. Currently, the country is undertaking the Kenya Integrated Household Budget Survey (KIHBS) expected to end in April 2006.

Involve various subject matter specialists from key stakeholders. Ie Donor community, line ministries, research institutions such as universities, NGOs, etc. Further to this, the KIHBS questionnaire was posted in the CBS website for comments from the general public.

- What linkages do you have with the PRSP monitoring unit?

Since 2003, when the NARC government came into power, they produced the Economic Recovery Strategy for Wealth and Employment Creation (ERS) which has three main pillars:

- **Economic growth**
- **Equity and poverty reduction which mainly addresses the aspirations of the PRSP and the MDGs.**
- **Governance**

The MDGs are entrenched within the sectors.

- Are you involved in any of the discussions over the choice of PRSP indicators?

Yes; In collaboration with the Monitoring and Evaluation Department.

- How do you decide on which PRSP / MDGindicators to monitor each year, frequency, degree of dissagregation (sample size)?

There are those monitorable ERS and MDG indicators which have been identified. These are incorporated within the household based surveys undertaken by the CBS. Examples include:

- **Access to safe drinking water and adequate sanitation**
- **Literacy, gender equity etc**

- Is there a programme for the monitoring of each PRSP / MDGindicator?

There is a GOK department charged with the responsibility of M&E of the ERS and MDGS in the country. This is the MED in the MP&ND.

- Do you find some indicators difficult to measure accurately?

Yes; eg

- **HIV/AIDS orphans. This is because the accurate number of deaths associated with HIV/AIDS**
- **Proportion of population below $1 (PPP) per day**
- **Poverty gap ratio {incidence x depth of poverty}**
- **Proportion of population below minimum level of dietary energy consumption**
- **Proportion of households with access to secure tenure**

Which surveys are in the household survey programme? Wh ich surveys have been conducted in recent years or are planned for the immediate future?

- **The ongoing 2004/05 KIHBS (includes a PPA component)**
- **The 2005 literacy survey - ongoing**
- **APRM (Governance) survey (completed)**
- **International Comparison of Prices (ICP)- ongoing**

Planned survey

- **Agriculture and Livestock census survey**
- **Small and Micro-Enterprise (SME) survey**
- **Census of Industrial Production**
- **The 2009 Population and Housing Census**

Other planned activities are outlined in the CBS national strategic plan (see www.cbs.go.ke)

- Have any 'light' monitoring surveys such as the CWIQ o r WMS been carried out (year, content, sample size)

CWIQ – piloted in 1997 preceding the WMS III.
WMS I - 1992 (12,050 Households)
WMS II - 1994 (10,860 Households)
WMS III - 1997 (12, 580 Households and excluded the Arid and Semi-Arid Parts of Northern Kenya)
KIHBS - 2004/05 (13, 430 Households)

- How do you decide which PRSP indicators are appropriate to monitor using the light'core survey and which are not?

N/A

- How do you ensure consistency in definition and question wording across surveys?

We maintain indicator definition within the household surveys for results comparability. The definitions comply with the international (UN) standards and domesticated to the country situation.

- How do you meet the needs of policy analysts who want cross-sectional information from households rather than indicators?

B. The use of statistical models to identify the poor and poorest households and to derive estimates for small areas;

Do you use statistical modeling techniques to identify the poor or poorest households in your light'core monitoring survey?
Yes

- How well have these worked?

Very well. In fact, Kenya has produced two volumes of the Poverty Monitoring Reports (The Geographic Dimensions of Well Being in Kenya).

- Are the results believed and used?

Yes. The associated point estimate errors are very low. The results are used for identification and targeting of the poor. The government uses the poverty index statistics to allocate Constituency Development Funds (CDF) and the Constituency Education Bursary fund.

Do you use modeling techniques to provide estimate for small administrative areas (second administrative level below national level) e.g. if the government system is national –regional - district, then to provide estimates for districts and below

- Do you have a poverty map with estimates of poverty incidence for small areas?

Yes, at the sub-district levels. We have poverty estimates for the divisions, locations, sub-locations (urban) and constituency level.

- Are the results believed and used? **Yes**

- Do you use the poverty map technique for anything other than poverty incidence?

Yes; Trying to superimpose on the maps, other socio-economic variables such as IMR, Education, water, schools infrastructure etc

Have you conducted any large sample surveys to give estimates for small administrative areas?
Yes; WMS II

- What problems have you encountered?

- Are you under pressure to conduct large sample surveys? For which topics?

C. Data entry and data processing for timely dissemination;

How long between the end of field work and the producing final tables for any light' monitoring surveys? How does this compare to other surveys?
Optical Mark reader scanning was undertaken for the first time ever during data capture of the 1997 WMS III. The process of data capture took long and was very tedious due to low quality paper and low precision of the mark readers in the questionnaires.

- Were the questionnaires scanned? Has this technology been adapted for other uses within the Office?

Scanning technology was used again during the 1999 Population and Housing Census. This time the character recognition scanning was successfully done.

- Were the data processing system and tabulation programmes ready and tested before the end of field work?

The tabulations as per the scanning process had already been put in place. However, they were never put into use since the World Bank withdrew funding midstream. The withdrawal of funding support culminated in CBS being unable to produce even a basic report from the survey.

D. **Reporting & Dissemination**

How long between the production of final tables and the publication of reports for any light'monitoring surveys? How does this compare to other surveys?
The process of basic report production takes a relatively short time depending on the immediate availability of both financial and technical support.

How do you report and disseminate results from your light'monito ring surveys?

- Do you publish survey reports, statistical abstracts, summary reports, flyers?

Yes; survey reports; flyers (ie 1998/99 labour force survey)

- Do you target reports for particular users?

The popular version of a report is produced to cater for the less technical public.

- Do the survey results get into the PRSP annual progress report?

The survey results feed into the M&E annual progress reports.

- Do you do further data analysis done after the initial presentation of results?

Further analysis is done. Ie poverty analysis

Do your indicators produce believable trends over time?

Yes; ie WMS series

Do you have particular problems getting consistent results for the same indicator over time? Which indicators?

Do you have survey results which differ from results given by administrative systems in sector ministries - how are these issues resolved?

How are the data sets stored? Are your surveys archived?

They are stored by the research, data processing and publication division. The division is charged with the responsibility of archiving the survey data sets and findings. There are also advanced plans to set up an IMIS in the department.

How do researchers get access to the datasets for further analysis?

Requests have to be made to the Director of Statistics

14.5. Lesotho 2002 CWIQ

Light Core Monitoring Survey

Workshop

Lesotho 2002 CWIQ

CWIQ 2002 and other suveys

- Bureau of Statistic(BOS) is a member of poverty monitoring sub-group of PRSP and has a number of responsibilities.
- BOS is involved in discussions over choice of PRSP indicators
- BOS provides data to monitor PRSP indicators

Surveys conducted by BOS

BOS conducts a number of surveys from time to time.(e.g. APS,HBS,CWIQ,LFS,LDS,EMICS etc)

- User-Producer workshops decide which PRSP indicators are appropriate and ensures consistency in definition and question wording.

Identification of poor or poorest households

- The wealth of the households was expressed in terms of quintiles in 2002 CWIQ. Thus, identifying the poor or poorest households.
- The quintiles were constructed for small areas e.g. Districts, Rural and Urban place of residence.

Recent Large Sample Surveys

- BOS conducted Lesotho Demographic Survey in 2001 and 2002/2003 Household Budget Survey.
- Estimates were derived for districts, rural, urban and ecological zones in these two large sample surveys.
- Lesotho is preparing for April 2006 Population and Housing Census.

Survey workplan CWIQ 2002

- About 5000 Households were randomly selected in all 10 administrative districts of the country.
- Data collection was planned to last for 1 month
- Completed questionnaires were retrieved daily from enumeration areas and submitted to BOS headquarters for scanning.

Data Entry and Data Processing

1. Software and Hardware used

- Teleform software was used to scan Questionnaires
- One fujitsu scanner
- Two computers
    1. For scanning questionnaires
    2. Editing inconsistencies
- Printer for printing of errors

3. Data processing systems

- Teleform was installed on one PC as well as a scanner, and were tested prior to actual scanning.
- The CWIQ data processing was developed by World Bank Africa Statistics Division.
- Scanning and production of preliminary tables took about 6 weeks, whilst with manual data entry, it takes longer to produce tables.

Reporting and dissemination

- Final reports are produced and are accessible through Bureau of Statistics information office - library.
- Dissemination workshops are held in all 10 districts of the country.
- Statistics website is used as another means of dissemination.

Storage of data sets

- Datasets are copied onto CDs and stored in a fireproof safe.
- Copies of datasets are stored on the server.
- Copies are made on backup tapes and stored in a safe.

Accessibility of datasets or reports

- Reports through Bureau of Statistics information office - library.
- Data can be available on CDs through arrangements with Bureau of Statistics for further analysis.

Advantages and disadvantages of scanning

1. Advantages
- Quick Data entry and tabulation
- Use of limited human resources

2. Disadvantages
- Expensive software
- Expensive scanning equipment

Problems encountered

- Late retrieval of questionnaires from the field

Future Plans

- Scan all surveys and censuses.
- Use CDs for dissemination.
- Use BOS website for dissemination.

14.6. Welfare Monitoring Survey 2005 NSO Malawi, i) presentation a, ii) presentation b

Welfare Monitoring Survey 2005 (WMS)

NSO, Malawi

Experience in shortening process time by scanning, editing process and tabulation plans

- Background of the WMS
- Plans for scanning and tabulation
- Status today
- Problems encountered
- Changes before next round
- Conclusion

WMS: Background

- Developed from CWIQ and other surveys
- Flexible design, contents adapted to data needs in the country, (MPRS and other indicators)
- Independent scanning software, ”Eyes and Hands”, can read characters, so ”bubbling” is not necessary

Plan for scanning and editing: Concurrently with the fieldwork

- Start scanning and editing (validation) 1-2 weeks after start fieldwork, mid-August
- Finish 1-2 weeks after end fieldwork, end-September
- Cross-checking and data cleaning in SPSS 1-2 weeks: finish mid-October

Plan for designing, programming and running tables

- Design tabulation plan, 1 week in May
- Programming tables: Start mid-August (on preliminary data)
- Write introductory and background parts of Statistical Abstract in August
- Workshop for running and editing tables and writing of text, second half of October
- Abstract ready for publishing in November, 1-2 months after end of fieldwork

Status

- Fieldwork completed on schedule
- Scanning/cleaning delayed about 2 weeks

Status (continued):

- A draft Abstract presented to NSO management 24 November,
- and at in-house dissemination 1 December,
- lot of useful comments, which implied more work,
- So final publication/dissemination postponed till January

Experiences, and problems encountered in scanning/editing

- Scanning and editing on the whole went according to plans, but:
- Not enough staff with knowledge of software (Eyes and Hands, SPSS)
- Problems with numbering of clusters, some manual editing necessary
- Printing quality somewhat inferior
- Paper jams caused by damage to some questionnaires

Problems in scanning (continued)

- The questionnaire had too many pages, white on black numbers (unreadable when scanned), and answering categories were not numbered
- E&H view function cumbersome
- Server crash

Experiences, and problems in designing, programming and running tables

Table design workshop successful:

- propositions for tables from each module
- standardised background variables
- variable specifications

were discussed and agreed upon through a combination of group work and plenary sessions

Experiences in designing, programming and running tables (cont.)

- New programs had to be made, because of extensive changes between pilot and main survey
- Lack of qualified staff: A small group was responsible for training, planning, supervision and managing of fieldwork, plus supervision of scanning/editing, in addition to programming and running tables and writing report

Changes before next round

- Shorten and simplify questionnaire, improve its scanning properties
- Do a new listing/ new sample
- 2-3 extra persons should be trained as experts in E&H
- Network so the E&H software can be used on more than one computer. Or an extra licence should be purchased.
- The programmer (external) should be present during first week to check that the programs are working

Changes before next round (continued)

- Measures to insure careful treatment of questionnaires (plastic envelopes?)
- 2-4 extra persons trained as experts in SPSS programming (tables and file management)
- The whole team should be trained in SPSS, preferrably combined with training in analysing/writing about survey data

Conclusions

- Despite (or because of) the problems, struggling with the data has been a good learning experience
- With a little extra expert advice at crucial points things would have worked even better
- The staff at the agricultural division is much better equipped to do the next survey than they were before

Thank you!

MALAWI WELFARE MONITORING SURVEY 2005

Presented by
Kelvin Saukila,
National Statistical Office

Presentation Plan

- Introduction
- Survey Coverage
  - Identification of issues
- MDGs and MPRS indicators
- Comparability
  - Integrated Household Surveys (1 & 2)
  - CWIQ 2002
- User – Producer contact

INTRODUCTION

- NSOs effort to provide relevant information for monitoring the welfare status of Malawians
- Designed to collect the minimum amount of information necessary for the identification and classification of vulnerable groups of households within society
- Follow – up to CWIQ 2002
- Conducted in 2005 between July and November
- 5,400 households covered

SURVEY COVERAGE

- Identification of issues for the WMS
  (Refers to topics covered in the WMS 2005)
  - Characteristics of household members
  - Education
  - Health and nutrition
  - Employment
  - Crop production
  - Housing conditions and amenities
  - Poverty predictors
  - Child information
    - Births and anthropometrical measures
    - Malaria prevention, treatment and vaccination
  - HIV/AIDS knowledge

MDG and MPRS indicators

- MDG Indicators covered in the WMS 2005:-
  - Indicator 6: Net primary enrolment rate
  - Indicator 8: Literacy rate of 15-24 years old
  - Indicator 10: Ratio of literate women to men, 15-24 years old
  - Indicator 17: Proportion of births attended by skilled health personnel
  - Indicator 19b: HIV/AIDS knowledge
  - Indicator 45: Unemployment rate of the 15-24 years old

MDG and MPRS indicators - contd

- MPRS indicators covered in the WMS 2005:-
  - Major impact targets
    - Female literacy rate
    - Proportion sick/ injured last 2 weeks
    - Labour force participation rate
  - Rural infrastructure (pillar 1)
    - Proportion with safe drinking water
    - Proportion with improved sanitation facilities
  - Basic education (pillar 2)
    - Female gross enrolment

Comparability

- WMS 2005 Questionnaire design
  - Based on CWIQ and IHS-2 questionnaire contents
    - IHS-1 conducted in 1998
    - IHS-2 conducted in 2004
    - CWIQ conducted in 2002
- WMS 2005 Results
  - Compared to other previously conducted household surveys in Malawi viz:-
    - MDHS
    - IHS (1&2)
    - CWIQ
  - Comparisons in line with MDG and MPRS indicators

User – Producer contact

- WMS 2005 Planning
  - Meeting with all major stakeholders done prior to finalising the questionnaire
    - To solicit ideas on the users' areas of needs
      - NAC
      - Min of Health and Population
      - Min of Education
      - Min of Agriculture and Food Security etc
    - Questionnare circulated to all users and producers
    - User – producer workshop conducted

Conclusion

- WMS has provided Indicators for monitoring MDGs and MPRS
- The indicators are comparable with other survey results i.e. CWIQ 2002, IHS-1&2
- NSO, Malawi has gained valuable experience in conducting WMS as a light core survey

THANK YOU FOR LISTENING

14.7. Light Core Monitoring Surveys and Poverty Monitoring - the Sierra Leone Experience

WORKSHOP ON

'LIGHT' CORE MONITORING SURVEYS

MAPUTO,MOZAMBIQUE.

12TH -14TH DECEMBER 2005

'LIGHT' CORE MONITORING SURVEYS AND POVERTY MONITORING:
THE SIERRA LEONE EXPERIENCE.

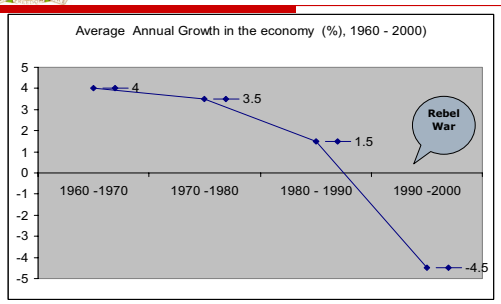
PRESENTATION BY:

MOHAMED KING-KOROMA
SHEIK A.T.ROGERS
(STATISTICS SIERRA LEONE)

Presentation Outline

- Background
- Poverty Monitoring in Sierra Leone
  - Data sources
  - Institutional framework
  - SL's progress towards achieving MDGs
  - Some SL-PRSP goals (2005 – 2007)
- Sierra Leone CWIQ pilot (2004)
- Data archiving and dissemination (DevInfo).
- Challenges for SSL
- Next Steps

Background



Background cont'd

- Interim PRSP – 2001
  - Restore national security & good governance.
  - Re-launch the economy after a decade of conflict (1991 – 2001)
  - Provide basic social services to most vulnerable groups – displaced/refugees, women/children and aged.

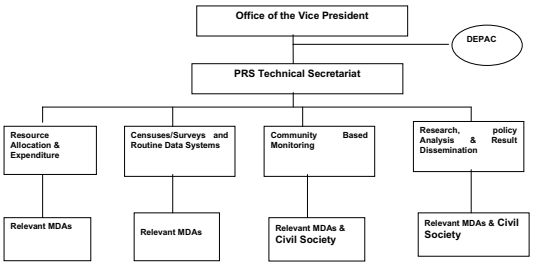
Background cont'd

- Full PRSP (2005) : 'A National programme for Food Security, Job Creation and Good Governance'.
- Achieving the SL-PRS core objecivtes would mean:
  - Achieving high and sustained broad-based growth
  - Providing essential and economic services to the poor, and
  - Improving governance at all levels.

Monitoring poverty

- Data Sources :
  - Quantitative (SLIHS 2003/04); CWIQ ,etc.
  - Qualitative :
    - PPAs (Participatory Poverty Assessment)
    - FGD (Focus Group Discussion)
    - PLF (Participatory Learning Forums).
  - SL – PRSP used both quantitative and qualitative data sources.

SL– PRSP: Institutional Framework for Monitoring and Evaluation



Monitoring poverty cont'd

- **Working groups** :
  - Resource Allocation and expenditure
    - Budget allocations
    - Expenditures
    - Public Expenditure tracking Surveys (PETS) - tracking of financial and other resources to service delivery points
  - Censuses/Surveys & routine data systems
    - Integrated HH Surveys,CWIQ, MICS,etc.
    - District/Local Council data systems.

Monitoring poverty cont'd

- Community based monitoring ; assess progress and impact of PRS on the poor (civil society).
  - District budget oversight committees
  - Community report cards
  - FGDs, PPAs,etc.
- Dissemination, research & policy analysis
  - Poverty monitoring database
  - Targeted information/seminars/workshops/media etc. (including Data Packs)
  - Capacity building in data access and use
  - Geographical information systems, including poverty mapping
  - Poverty and Social Impact Analysis

Monitoring poverty cont'd

Millennium Development Goals	Status of Achievement (in % unless otherwise indicated)		
	1990	2000	2015 Target
Population living in extreme poverty	57.0	89.0	28.5
Universal primary education (net primary school enrolment for school-age children)	52.0	42.0	100
Proportion of children under five- years that are underweight	24.0	27.0	12.0
Proportion of population without access to safe drinking water	35.0	57.0	17.5
Proportion of children under one-year immunized against measles	79.0	60.0	100
Infant mortality rate 1,000 live birth	302	284	100
Maternal mortality ratio 100,000 live birth	700	2,000	175
HIV prevalence rate among adults 15-49 years	-	4.9	0.0

Monitoring poverty cont'd

- **Some SL-PRS Goals (2005 – 2007)**
  - Increase the accessibility and affordability of service from 45 to 60 percent of the population by 2007.
  - Reduce maternal, infant and under five mortality by one-third by 2007
  - Reduce morbidity and mortality, due to malaria and other communicable diseases by 50 percent by 2007.
  - Reduce the incidence of HIV/AIDS and other sexually transmitted diseases.
  - Reduce malnutrition, especially among children and women.
  - Increase enrolment in primary school by ----- percent in 2007
  - Increase the level of literacy among 15-24 year olds by ---- - percent by 2007.
  - Eliminate gender disparity in primary education by 2007

Sierra Leone CWIQ Pilot

- Objective :
  - to determine CWIQ's suitability as a national poverty-monitoring instrument that can provide benchmark and annual progress information on poverty reduction at an affordable level.
  - Conducted by SSL with technical assistance provided by two consultants
    - Mr. Jonah Coker (Survey Statistician)
    - Yaw Antwi-Adjei (Data Processing)

Sierra Leone CWIQ Pilot cont'd

- Survey areas :

	Kono District		Western Area	
	Urban	Rural	Urban	Rural
EAs/Clusters	10	10	10	10
HHs/EA	25	25	25	25

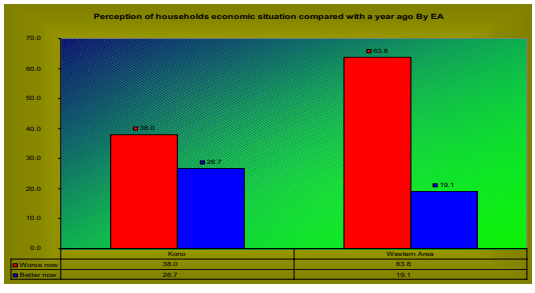
Sierra Leone CWIQ Pilot cont'd

- Survey Instrument :
  - Generic CWIQ questionnaire supplied by the World Bank with few modifications in Sections C & H
- Field work and field returns
  - Start date **July 26, 2004**
  - First batch of returns July 28, 2004
  - Last batch of returns **August 09, 2004**

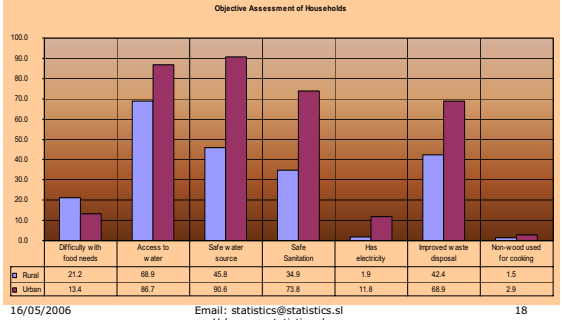
Sierra Leone CWIQ Pilot cont'd

- Data processing
  - TELEform – automated data capture software
  - Microsoft Office 2000 – Excel & Access
  - Three desktop computers
  - One workgroup document scanner
- Completion of data capture and validation : **10 August 2004**

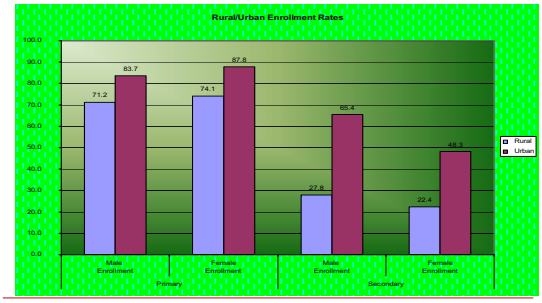
SL- CWIQ Pilot – some outputs



SL- CWIQ Pilot – some outputs



SL- CWIQ Pilot – some outputs



Data Archiving & dissemination

- DevInfo – already adopted as a platform for presentation and archiving of national statistics and other survey outputs. Work in progress
- GDDS – SL metadata posted since May 2003.Updated regularly.
- SSL website : www.statistics.sl
- Simultaneous dissemination of annual/monthly publications :
  - Statistical digest
  - National Accounts
  - Consumer price Index (regional).
  - External trade statistics
  - Demography and Social Statistics (outputs from Census).

Challenges for Poverty monitoring (SSL)

- Human resources :
  - Strong data collection capabilities but weak capacity in programming, data analysis and reporting.
  - Much of data processing done by contract staff.
  - Weak in – house IT support capacity
  - Staff attrition
  - Material : enough PCs (from 2004 census) but not enough CWIQ equipment and software
- Finance ; dwindling financial allocation from central government; negative implications on training particularly in ICT,study tours and conduct of some routine activities.

Next Steps

- Training to improve capacity in the areas of survey design ,survey data analysis and reporting.
- Build upon existing capacity to conduct PRS surveys in Year One of implementation.
- Plan for CWIQ/WMS round I in early 2006.

THANK YOU

statitics@statistics.sl

www.statistics.sl

14.8. Enhanced Evidence-based Decision-making – SouthAfrica

SIDA/DFID/CIDA and Statistics South Africa Agreement

Enhance evidence-based decision-making in the key development areas of Poverty and HIV/Aids in South Africa

Four Subprojects

- Developing capacity for Poverty Research and Analysis
- Improvinging the understanding of the impact of HIV-Aids
- Statistical Training, Methods and Tools
- Development of Methodology for Income and Expenditure Survey

Project Goal

To provide statistical information on Poverty within nationally agreed concepts and definitions

Framework for a Poverty Monitoring System at Stats SA

- Literature review
- South African Policy Documents addressing Poverty reduction
- Guidelines from International Organizations
- Interview with Stakeholders

Stakeholders

- Researchers
- Ministries
- NGOs
- Provinces

Issues

- Use of Statistics for Poverty Analysis within organization
- Sources of Data for Poverty Analysis
- Collection on Poverty Data within organization – Surveys, Admin...
- Data Gaps

Stakeholders Views

- A Poverty Line
- National Poverty Reduction Strategy Paper
- Coordination of Regional Monitoring Systems
- Annual measure of Poverty and Inequality
- Administrative Data
- Poverty Data at small Area Level
- Data Quality

The Road forward

- The Role of Statistics South Africa in relation to other SA Bodies
- One Poverty Survey rather than Poverty Modules in several Surveys
- Periodicity
- Geographical Level
- Administrative Data – Access & Quality

Future Goal

- A Coherent Information System
- A Poverty Data Base

14.9. LIGHT core monitoring surveys - Zambia

Workshop on 'LIGHT' core monitoring surveys

Maputo Mozambique, 12-14th December 2005

Zambia

By
Frank Kakungu
Nchimunya Nkombo

Outline

- Introduction
- Living Conditions Monitoring Surveys (LCMS)
- Roles and Objectives of LCMS
- Contents of LCMS
- Links with PRSP and MDGs
- Selection of indicators
- Use of Statistical models
- Data processing
- Reporting and Dissemination
- Challenges

Introduction

- In the 1980s Zambia introduced structural adjustment programme (SAP) and was vigorously embarked on in 1991.
- The SAPs were having adverse effects on the poor and vulnerable subgroups in the population and required safety nets.
- This led to the introduction of social dimensions of adjustment, which were aimed at mitigating the negative effects of SAP.
- Thus it became necessary to monitor the welfare of the people.
- In 1991, the government in collaboration with World Bank launched the first welfare monitoring survey known as Social Dimensions of Adjustment (SDA) Priority Survey I (PS1).

Welfare monitoring and assessments in Zambia

- Six countrywide surveys have been conducted to measure the living standards of people since 1991. These are:
  - 1991 Priority Survey I(PSI)
  - 1993 Priority Survey II (PSII),
  - 1996 Living Conditions Monitoring Survey I (LCMSI),
  - 1998 Living Conditions Monitoring Survey II (LCMSII) ,
  - 2002/3 (LCMSIII) also known as Integrated Household Budget Survey,
  - 2004 Living Conditions Monitoring Survey IV (LCMSIV).

Role and Objectives of the Surveys

- The overall objective is to measure socio-economic performance of government's policy, programmes and project implementation.
- Specific objectives include:
  - to provide a quick identification of policy targets groups,
  - and to provide a mechanism where key socioeconomic variables can be easily and regularly produced to describe and monitor the well being of different groups of households.

Contents of the Surveys

- It places emphasis on some basic indicators. These are
  - Education
  - Health
  - Nutrition
  - Consumption and expenditure
  - Housing
  - Asset ownership
  - Income
  - Food Production
  - Income generating activities
  - Access to facilities
  - Copying strategies

Links with PRSP and MDGs

- In order to monitor the PRSP progress a list of indicators was selected.
- The selection of indicators took into account the need to monitor Zambia's progress towards the achievement of the Millennium Development Goals.
- Central Statistical Office (CSO) takes a leading role in providing statistical indicators through the LCMS'.
- This has been the core monitoring mechanism for indicators on poverty, welfare and access/utilization of social services.

Selection of Indicators

The selection of indicators is demand led

- There is a User Producer committee.
- The Users-Producer Committee consists technocrats appointed from various sector ministries and Civil Society Organization.
- The committee undertakes the role of guiding the CSO in the development and standardization of survey instruments.

Use of Statistical Models

- Zambia plans to undertake Poverty Mapping using the Census and Survey data
- This is meant to give Poverty estimates at sub national levels such as, District and Ward levels.
- An attempt has been made to give district poverty estimates using the 2004 LCMS.

Data Processing

- Data entry is usually done in provincial centres using Census and Survey processing (CSPro) software.
- This software has components for data entry and verification and another component for consistency checks.
- For advanced processing and tabulation SAS is normally used.
- Scanning technology was used to process Census 2000.
- For surveys scanning was only used in a pilot CWIQ survey that was undertaken in 1997.
- Only data entry programmes are tested before the end of field work

Reporting and Dissemination

- Results have shown some consistency over the years.
- Upon completion of the survey, results are presented to stakeholders in a workshop setup.
- Results are published in form of reports.
- Summaries of the results are posted on the CSO web site.
- Various articles are presented in the CSO Monthly bulletin.
- Data sets are archived on CD ROMs.
- Full reports and CD ROMs are available for sale at subsidized prices.

ZambiaInfo

- Zambia has adapted Devinfo data base and it is called ZambiaInfo.
- ZambiaInfo contains indicators from the Censuses and Surveys undertaken by CSO.
- It also contains indicators from statistical units in the line ministries such as Education and Health.
- The database is saved on CD ROMS which are distributed to the public at no fee.

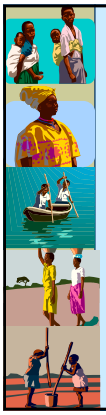
Challenges

- Some indicators are difficult to measure, such as Hunger statistics
- Demand for district estimates is high
- Non availability of panel data to measure trends accurately
- Limited data handling skills by various users hinders further data analysis.
- Differences on some indicators for line ministries because of different catchments area definitions used e.g. Statistics on Health

Challenges Cont.

- Tabulation programmes are normally not tested.
- Results usually published six months field work.
- Zambia has not conducted any Core Welfare Indicators Questionnaire (CWIQ) survey. However, there are plans to conduct CWIQ surveys in between the longitudinal surveys.

14.10. Poverty Correlates/Predictors



POVERTY CORRELATES/PREDICTORS

LIGHT SURVEYS WORKSHOP

DECEMBER 12-14, 2005

MAPUTO, MOZAMBIQUE

John Ngwafon, World Bank, Washington D.C

INTRODUCTION

Typical Survey Program:

YEAR	SURVEY TYPE
YEAR 1	IS/LSMS/HBS
YEAR 2-4	CWIQ, LFS
YEAR 5	IS/LSMS/HBS
-	
-	
-	
YEAR 10	POPUL. CENSUS

INTRODUCTION

In order to develop Sound, and Appropriate Poverty Monitoring Policies and Programs, a clear understanding of the characteristics of the poor, causes of Poverty and where the poor are located, is required.

PRSP implementation agencies seek to capture information as to how the poor and vulnerable groups are faring during times of rapid social and economic change. The Poverty Monitoring System is an instrument for monitoring progress of Poverty Oriented policies, programs and projects in terms of impact, effectiveness and efficiency.

INTRODUCTION

What Follows after the Surveys:

- **Develop a Poverty Line and produce your Poverty Indicators: (P_0, P_1 and P_2). For the sake of the presentation, I will assume that this is a CONSUMPTION Based Poverty Line;**
- **Develop other Baseline Indicators to be Monitored, along with MDGs, PRSP and IDA indicators.**

INTRODUCTION

Statistical Offices collect information on different dimensions of household Living Standards, covering: Consumption, Assets, Education, Health, Agricultural activities, Income and Child Anthropometrics.

Household Surveys are Key Instruments often used, Particularly detailed Living Conditions Surveys such as: the LSMS, IS and HBS which provide a very rich source of Information of Living Conditions in the Country.

INTRODUCTION

The LSMS and IS surveys are very expensive to conduct, consequently such surveys may not be feasible each year.

It raises the issue of whether it would be possible to monitor evolution of poverty, based on much smaller surveys that collect information on a relatively small set of indicators, which are collectively Good predictors of consumption Poverty or any poverty indicator.

INTRODUCTION

If such a Reliable Set of Indicators could be Identified, these light surveys could be conducted in years when Conventional Consumption-based Surveys cannot be conducted. This could still allow the evolution of poverty to be studied.

This Presentation aims at setting the stage on the Theoretical underpinnings and Methodology of Identifying Reliable Indicators for Consumption and Computing predictors that could be used to Rank Households by their Standard of Living.

INTRODUCTION

Several Analysts have used Different Names for these Indicators:

(1) Andrew McKay in 2001 produced similar indicators for Uganda which he called: POVERTY CORRELATES

(2) IFPRI for Malawi (2002) called the Indicators: PROXY MEANS TESTS

(3) Fofack for Ghana (1999) called them: POVERTY PREDICTORS

INTRODUCTION

The Methodology in each Case is basically the same. They usually involve Identifying a Small set of Poverty correlates from a much larger set of prospective candidates, based on statistical techniques of CORRELATION AND MULTIVARIATE REGRESSION ANALYSIS.

The techniques assess how well a set of indicators perform collectively in predicting CONSUMPTION-based Standard of Living Measures, depending on the Reliability of their prediction.

METHODOLOGY: SELECTION OF VARIABLES

STEP 1.

Identify potential poverty correlates from the consumption survey. UGANDA STARTED WITH 250 POTENTIAL CORRELATES COVERING DIFFERENT ASPECTS OF HOUSEHOLD LIVING CONDITIONS.

STEP 2.

For each Potential Correlate, its association with Consumption is considered

METHODOLOGY: SELECTION OF VARIABLES

STEP 3.

The more important correlates are then considered collectively, with a view of identifying 10-20 correlates which most accurately predicted Household Consumption.

THIS IS BASED ON MULTIPLE REGRESSION ANALYSIS, ELIMINATING THOSE CORRELATES FOUND TO BE LESS SIGNIFICANT.

METHODOLOGY: SELECTION OF VARIABLES

STEP 4.

Poverty Correlates may be Geographic-specific in the sense that a characteristic that is a good poverty correlate in one region, may be a poor correlate in another.

The selection of variables should be conducted separately for the following groups: NATIONAL, URBAN, RURAL and maybe some sub-regional classifications such as Zones, East, North, South and West etc.

METHODOLOGY: SELECTION OF VARIABLES

STEP 5.

In the Identification of Poverty Correlates care should be taken to avoid characteristics that are difficult to measure in a light survey, even though that characteristic might be a very good predictor of poverty. For example:

VALUE OF A HOUSEHOLD ASSET.

METHODOLOGY: VARIABLES SELECTION

STEP 6.

Variables Selected Should be based on several criteria:

- **❑ Their Correlation to Total Consumption;**
- **❑Diversity in the type of household variable. A range of information which reflects different aspects of household welfare;**
- **❑The variable would be collected in light surveys like the Core Welfare Indicators Survey (CWIQ).**
- **❑ REF: Turn to P.16 of the document "Simple Household Poverty Assessment Models for Malawi".**

DEVELOPING THE MODELS

STEP 1.

Ordinary Least Squares (OLS) Regression procedure is used to construct the model.

The DEPENDENT VARIABLE in the regression is the Welfare Indicator of Interest. This would usually be the Total Consumption as computed from the information in the Integrated Survey. Nutritional intake could also be used in a nutrition or food security application of this method.

DEVELOPING THE MODELS

STEP 2.

The NATURAL LOGARITHM of Total Consumption is the preferred dependent variable because it improves the predictive power of the models.

The Distribution of the Logged consumption across households in the sample better approximates a Normal Distribution than the simple Unlogged indicator.

DEVELOPING THE MODELS

STEP 3.

Household surveys typically include modules on: Demography, Education, Employment, Income Sources, Consumption, Living Conditions and a range of information on household members.

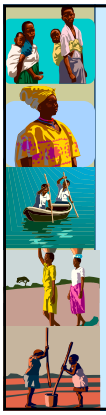
The EXPLANATORY VARIABLES in the right-hand side of the model can include any of the variables listed above, as well as other characteristics of the household available from other sources such as the Census.

MODELS

NATIONAL, URBAN and RURAL

Details of how these models are constructed are found in the handouts.

Participants are encouraged to try to replicate some of the results from their own National surveys. If participants are interested in a National Workshop to go through the steps, this can be arranged



HOW WELL do the correlates predict Standard of Living

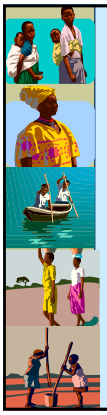
This can be assessed in two ways

(1) The coefficient of determination R^2

The higher R^2, the better the prediction;

(2) An assessment of how well Quintile groups from the prediction match those from the Consumption Based measure by simple Cross tabs.

(3) Generally models have not done very well in this aspect.



EVALUATION OF RESULTS SO FAR

Although models appear to fit well on Statistical Criteria (R square), their performance in predicting which quintile Group a Household is Located in, are Disappointing;

(1) The Predictions from these models should not be used to predict value of Poverty Indices;

(2) The models in principle could still be used to predict changes in the average consumption based standard of Living for Groups of the Population, e.g by location, by type of economic activity, by gender etc.

WHAT TO DO IN A CASE WHERE AN IS/LSMS SURVEY DOES NOT EXIST

There are cases where the Detailed Consumption survey is not available; We faced this situation in Lesotho 2002 CWIQ

(1) The Predictions used "PRINCIPAL COMPONENT" analysis to construct the predictors; CAUTION: Methodology has not been tested;

(2) The details are available as a short Write-up

NEXT STEPS

(1) Prepare a Handbook and User's Guide on the Methodology;

(2) Conduct Workshops to build Capacity on how to develop the Predictors.

14.12. A Statistical Model for Simple, Fast and Reliable Measurement of Poverty

A Statistical Model for Simple, Fast and Reliable Measurement of Poverty

by Astrid Mathiassen
Division for Development Cooperation
Statistics Norway
December 13, 2005

Statistisk sentralbyrå

Outline

- Present an approach for predicting the headcount ratio by combining an expenditure survey and a light survey
- Present some experiences and preliminary results from some analyses
- Further improvements?

Background

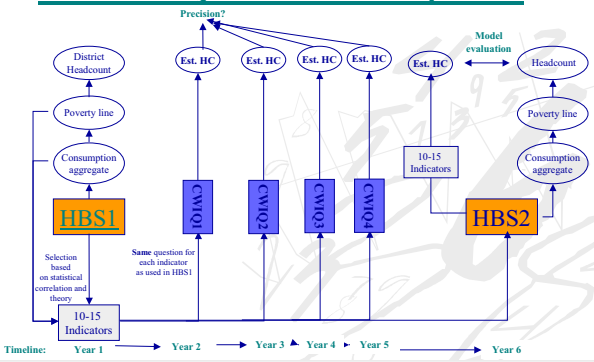
- Increased national and international demand for annual monitoring of poverty (PRSP, MDG)
- Traditional source of poverty data: Full-fledged Household Budget Survey (HBS)
- But: Due to very high financial and administrative costs, a full HBS is typically conducted only every 5-10 years
- However: Annual lower cost "light surveys" (e.g. CWIQ-surveys) are common.
- The approach taken is to estimate annual regional/ district poverty headcount from the "light survey" with its corresponding uncertainty, without undertaking a full expenditure survey

How "light surveys" can be used to produce annual poverty data

- Use HBS to identify 10-15 poverty indicators that can be included in the "light survey".
- Estimate the model parameters in an expenditure model based on data from HBS
- Use indicators from the light survey together with the model to predict the headcount ratio

– Poverty estimates based on the probability of being poor with corresponding estimates of the standard error of the predictor

The Poverty Prediction Sequence



Statistisk sentralbyrå

Identifying poverty predictors

- If a light survey is already set through; use common variable in two surveys (Mozambique case)
- If a light survey is about to be set through, use HBS to freely identify variables to be included (Malawi case)
- Criteria for indicators; reliable, easy and fast to measure, correlated with household expenditure

The predictor for the headcount ratio is obtained by;

- Estimate a multivariate regression model between expenditure and indicators selected from HBS1
- Predict expenditure per capita for the households in light survey
- Calculate the probability that a households predicted expenditure falls below the poverty line by inserting in a probability function.
  - Even households with high predicted expenditure have a non zero probability of being poor
- Predict the headcount ratio by calculating the weighted average of all individual probabilities of being poor

Uncertainty in prediction of headcount ratio

- The (quantifiable) variance of model prediction is due to;
  - sampling of indicators in the light survey
  - uncertainty in estimated model parameters
  - the predictor is biased
  - the headcount ratio in the population deviates from expected headcount ratio
- Factors that we are *not* able to quantify
  - Misspecification of model/model parameters change over time

In contrast

- The variance of HBS prediction is *solely* due to;
  - sampling error in HBS; only a sample of all households are used to estimate average expenditure

Results from Testing

- The model predicts well within sample
  - When using half sample to predict for the other half
- The standard errors of the model-based Headcount Ratio prediction may even be smaller than the standard errors of the HBS Headcount Ratio prediction
  - Somewhat counter-intuitive: The reason is that we apply a model to find the average poverty ratio (Headcount) from the "Light Survey", and compare this to starting from "scratch" in a HBS.

Some experiences

- Apply separate models for Rural and Urban
- IAF and IFTRAB, Mozambique, joint set of identical poverty predictors is small. In particular, the rural model suffers
- Experiences; good indicators are housing, demography and education. In addition typically variables like; did you eat meat last week, how many meals did you eat, do the children have more than one pair of shoes.

Some results from Malawi

Region	**IHS2-04** HC estimate (st.error)	**WMS-05** HC estimate (st.error)	No. Observations in WMS
Rural North	56 % (2,2)	52 % (2,6)	870
Rural Central	47 % (1,2)	48 % (2,0)	1500
Rural South	64 % (1,1)	63 % (2,0)	2070
Urban All	25 % (1,6)	25 % (2,4)	700

Some results from Mozambique

Province	IAF 2002/03	IFTRAB 2005	No of observations IFTRAB
Rural North			
Rural Central			
Rural South			
Urban North			
Urban Central			
Urban South			

Road Ahead

- Further testing of model
  - Use two subsequent HBS and test stability of parameters in model; compare poverty predicted by model to poverty based on expenditure estimates
  - Malawi IHS1 and IHS2?
- Adjust for seasonal variation
- Random coefficients modeling, Heteroscedasticity
  - Framework for stochastic model parameters, heterogeneity
- Other applications; e.g. a hunger model

14.13. How well do light surveys predict poverty - Mozambique

How well do light surveys predict poverty?

Preliminary results from a validation study

Kenneth Simler
International Food Policy Research Institute
&
Ministério da Planificação e Desenvolvimento

Workshop on Light Core Surveys
Maputo
12–14 December 2005

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

Outline

- Objectives of study
  - Can CWIQ be used to predict poverty?
  - More generally, how well do models stand up over time? (Validation)
- Brief review of methodology
- Results
  - compared to consumption in same survey
  - compared to consumption in later survey (preliminary)
- Some pertinent issues
- Conclusions

Objectives: Predicting welfare levels

- Common approach

to rank households by consumption levels.

- Where we diverge:
  - dividing by quintiles (relative poverty)
  - estimating poverty indices

Validation: Testing the predictions

- Some very strong assumptions.
- Compare prediction with actual measured consumption (and poverty) in subsequent survey, i.e.,

Methodology: Data sources

- 1996–97 National Household Survey (IAF)
  - source of information for statistical models
- 2000–01 CWIQ/QUIBB
  - poverty prediction when consumption is not collected
- 2002–03 National Household Survey (IAF)
  - compare predictions from models with actual consumption data collected
- All are representative at national, provincial, and urban/rural levels. Sample size from 8250 to 13,790.
- Geo-referenced data on spatial characteristics (average rainfall, temperature, elevation, road networks)

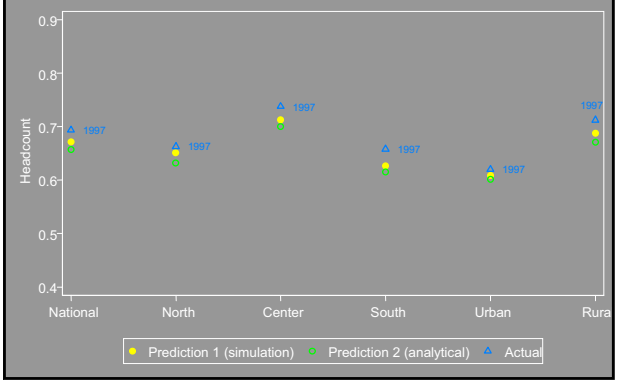
Methodology: Model estimation

- Estimated separate models for each province and Maputo city.
- Variables included: assets, literacy and education, household composition, housing characteristics, geographic characteristics, community characteristics
- Overall adjusted $R^2 = 0.508$
- Two methods of estimating consumption and poverty:
  - Elbers, Lanjouw & Lanjouw (ELL) – Simulation methods used for poverty mapping
  - Hentschel et al. (1988) & Datt (2005) – Analytical solution that is the "grandfather" of ELL.

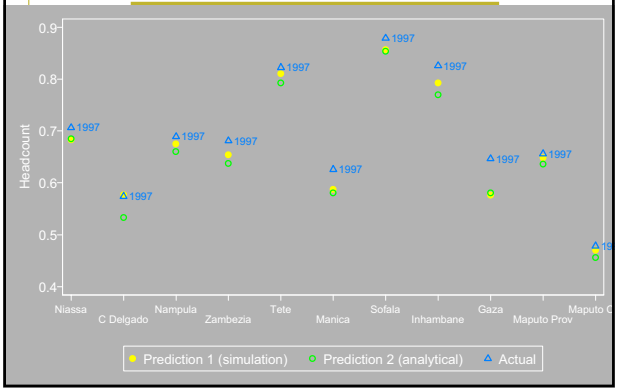
Results: Within-sample (quintiles)

Actual quintile	Predicted quintile: Poorest	2	3	4	Richest
Poorest	58	27	12	4	0
2	24	32	25	16	3
3	11	24	29	25	12
4	6	12	23	32	27
Richest	2	5	12	24	58

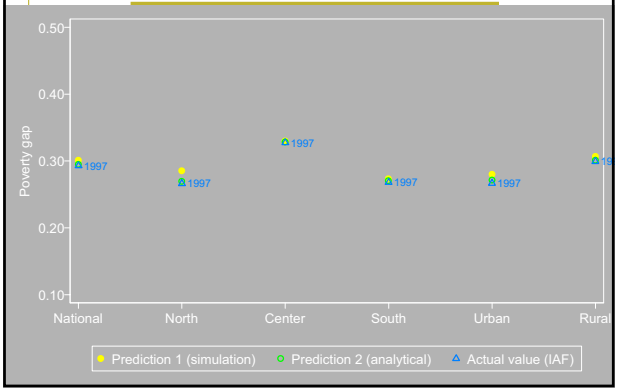
Results: Within-sample (headcount)



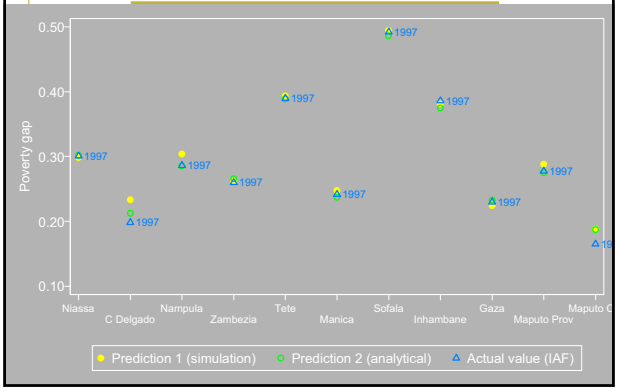
Results: Within-sample (headcount)



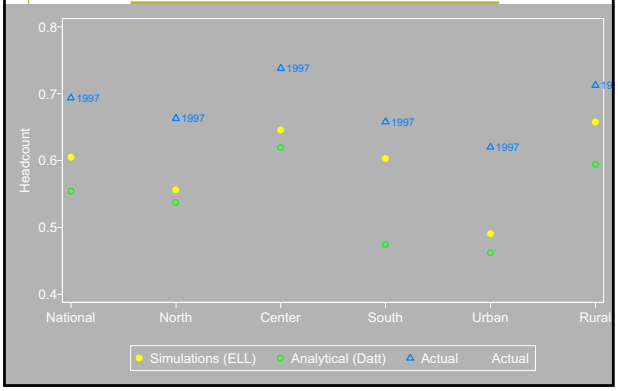
Results: Within-sample (poverty gap)

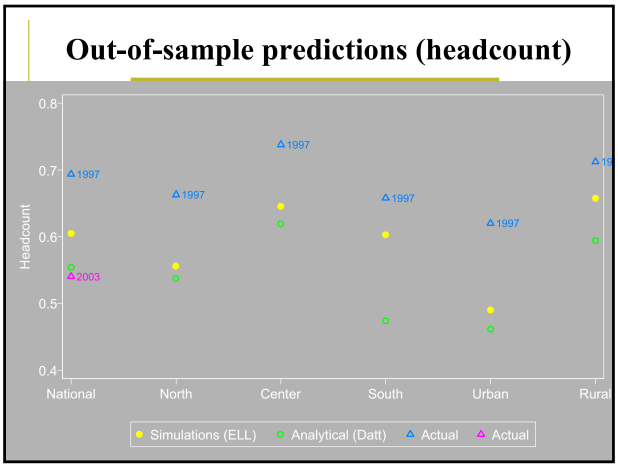


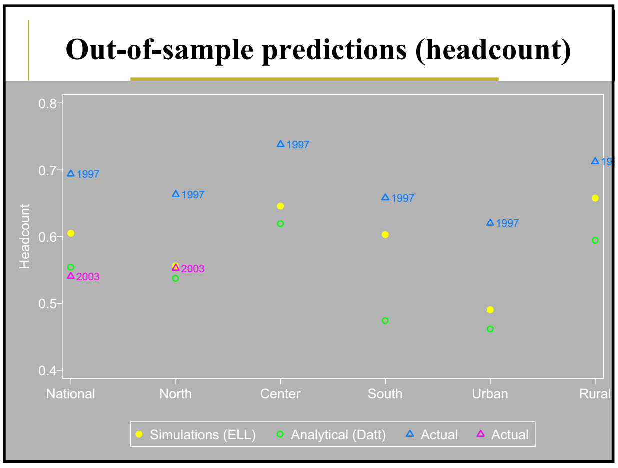
Results: Within-sample (poverty gap)

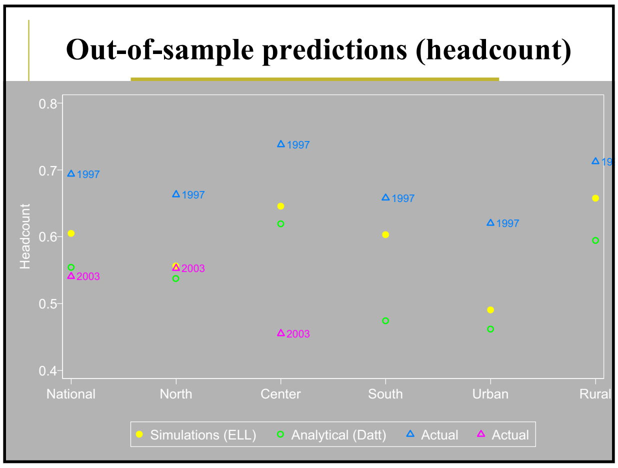


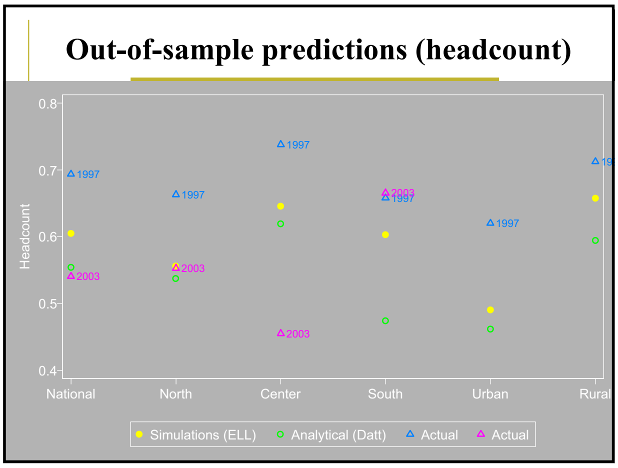
Out-of-sample predictions (headcount)

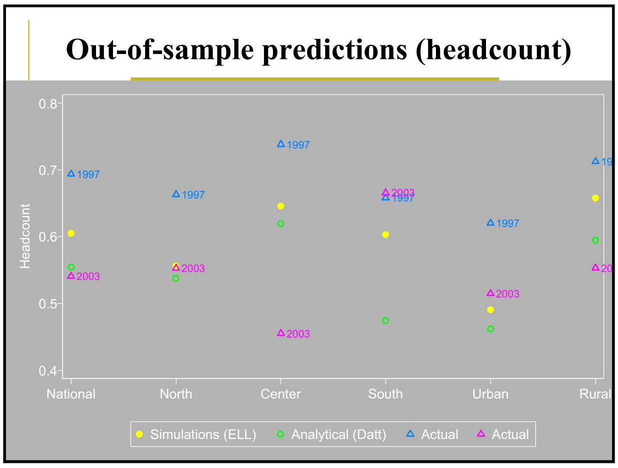


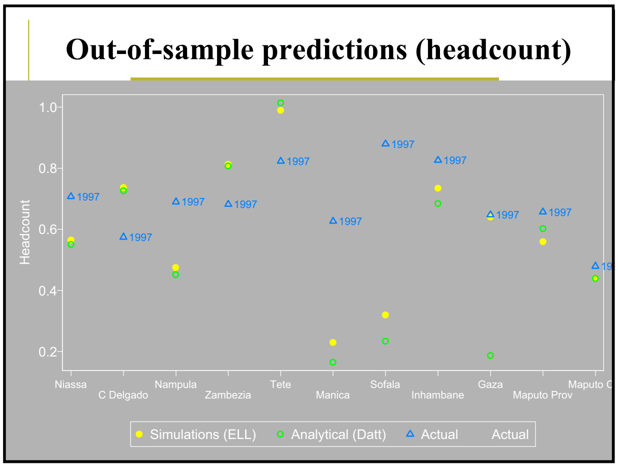

Out-of-sample predictions (headcount)
Headcount
0.8
0.7
0.6
0.5
0.4
National
North
Center
South
Urban
Rural
Simulations (ELL)
Analytical (Datt)
Actual
Actual


Out-of-sample predictions (headcount)
Headcount
0.8
0.7
0.6
0.5
0.4
National
North
Center
South
Urban
Rural
Simulations (ELL)
Analytical (Datt)
Actual
Actual

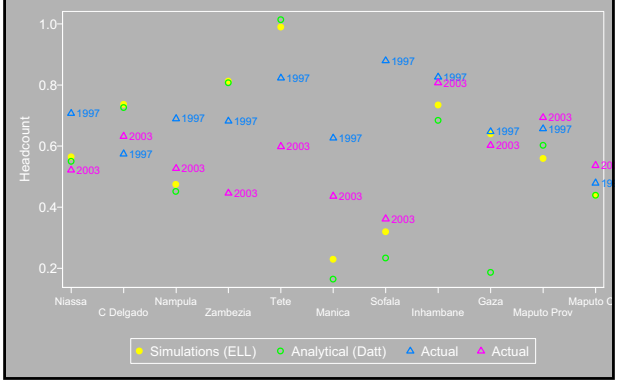

Out-of-sample predictions (headcount)
Headcount
0.8
0.7
0.6
0.5
0.4
National
North
Center
South
Urban
Rural
Simulations (ELL)
Analytical (Datt)
Actual
Actual


Out-of-sample predictions (headcount)
Headcount
0.8
0.7
0.6
0.5
0.4
National
North
Center
South
Urban
Rural
Simulations (ELL)
Analytical (Datt)
Actual
Actual

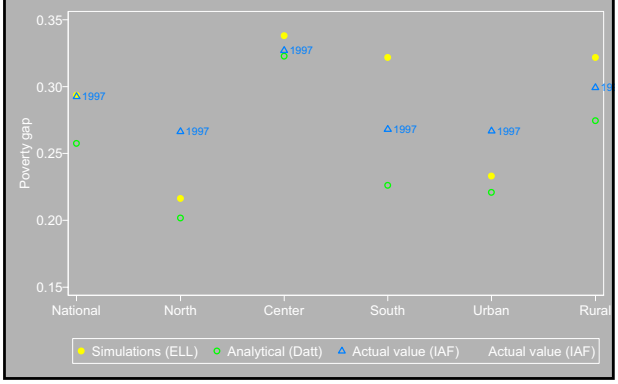

Out-of-sample predictions (headcount)
Headcount
0.8
0.7
0.6
0.5
0.4
National
North
Center
South
Urban
Rural
Simulations (ELL)
Analytical (Datt)
Actual
Actual


Out-of-sample predictions (headcount)
Headcount
1.0
0.8
0.6
0.4
0.2
Niassa
C Delgado
Nampula
Zambezia
Tete
Manica
Sofala
Inhambane
Gaza
Maputo Prov
Simulations (ELL)
Analytical (Datt)
Actual
Actual

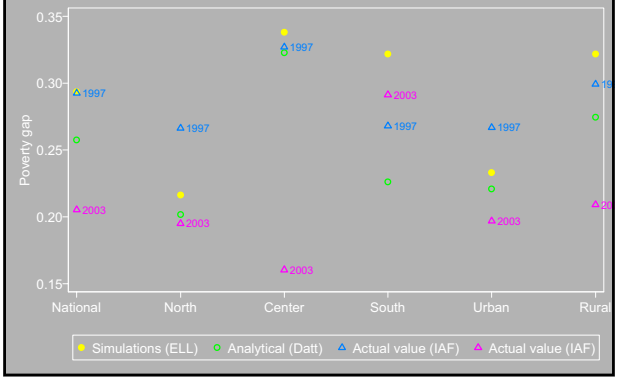
Out-of-sample predictions (headcount)



Out-of-sample predictions (poverty gap)



Out-of-sample predictions (poverty gap)



Out-of-sample predictions (poverty gap)

Out-of-sample predictions (quintiles)

Actual quintile	Predicted quintile: Poorest	2	3	4	Richest
Poorest	32	26	16	13	12
2	23	26	21	17	14
3	19	20	21	20	20
4	17	18	22	23	21
Richest	9	10	20	27	34

Precision of estimates

	IAF97	Std Err	IAF03	Std Err
National	67.2	1.1	58.3	4.3
North	65.1	1.6	54.5	4.0
Center	71.3	1.8	62.0	7.1
South	62.6	1.8	57.1	11.9
Urban	60.9	1.6	47.9	4.5
Rural	68.8	1.2	63.2	5.0
Niassa	68.3	3.8	54.0	9.8
Cabo Delgado	57.7	3.5	73.9	4.8
Nampula	67.5	2.3	46.0	6.4
Zambézia	65.4	3.2	78.5	10.7
Tete	81.1	2.3	97.7	1.7
Manica	58.7	3.8	19.0	14.6
Sofala	85.7	2.0	29.6	19.3
Inhambane	79.3	2.3	70.0	9.3
Gaza	57.6	3.7	56.9	39.9
Maputo Province	64.8	3.5	55.1	3.1
Maputo City	47.0	2.6	43.0	2.1

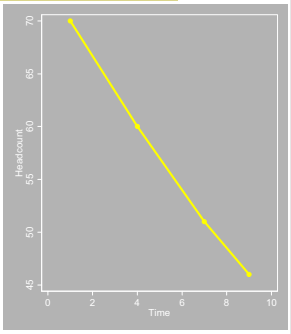
Some conclusions & important issues

- Predicts poverty reasonably well at national level, but sub-national performance is erratic.
- Weak predictive power even for the more modest goal of predicting quintiles
- Likely that the coefficients have changed over time (area for future research)
- Are the Xs capturing inter-annual changes that affect consumption?
- Issues of statistical significance, especially for small changes (e.g., short intervals)

25

Broader issues

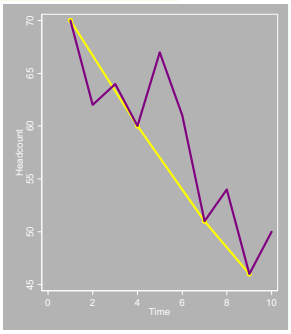
- How does one monitor poverty when it is subject to inter-annual fluctuations?
- Interpreting "blips". What do you do if poverty doesn't fall, or rises? More of a problem with short-interval surveys.
- Volatile sub-national components of an often more stable national picture.



26

Broader issues

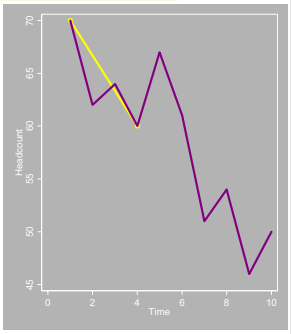
- How does one monitor poverty when it is subject to inter-annual fluctuations?
- Interpreting "blips". What do you do if poverty doesn't fall, or rises? More of a problem with short-interval surveys.
- Volatile sub-national components of an often more stable national picture.



27

Broader issues

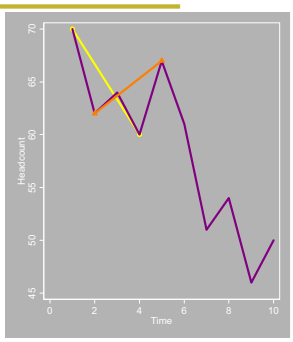
- How does one monitor poverty when it is subject to inter-annual fluctuations?
- Interpreting "blips". What do you do if poverty doesn't fall, or rises? More of a problem with short-interval surveys.
- Volatile sub-national components of an often more stable national picture.



28

Broader issues

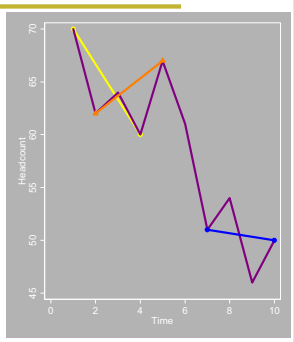
- How does one monitor poverty when it is subject to inter-annual fluctuations?
- Interpreting "blips". What do you do if poverty doesn't fall, or rises? More of a problem with short-interval surveys.
- Volatile sub-national components of an often more stable national picture.



29

Broader issues

- How does one monitor poverty when it is subject to inter-annual fluctuations?
- Interpreting "blips". What do you do if poverty doesn't fall, or rises? More of a problem with short-interval surveys.
- Volatile sub-national components of an often more stable national picture.



30

Poverty monitoring & poverty dynamics

- Poverty monitoring tends to look at changes in poverty in the population.
- One perspective on poverty dynamics is to examine changes in poverty/well-being of households over time (longitudinal).
- Chronic and transitory poverty have different causes, and different policy implications.
- Need to decide where to invest resources, considering significant opportunity cost of whatever choice is made.

31

Thank you

Muito obrigado

Merci

32

How well do light surveys predict poverty?
Preliminary results from a validation study

Kenneth Simler
International Food Policy Research Institute

Workshop on Light Core Surveys
Maputo
12–14 December 2005

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

14.14. Poverty Monitoring: Lessons from the Experience in Mozambique

PURDUE UNIVERSITY

Poverty Monitoring: Lessons from the Experience in Mozambique

Channing Arndt
Ministry of Planning and Development
Purdue University

Poverty Monitoring: Purposes

- To inform current and future policy.
- To evaluate performance of past policy.
  - Government
  - Donors

Some Facts on Mozambique

- 1992: "poorest country in the world."
- Rapid economic growth (from a low base) since 1992.
- Substantial aid inflows:
  - About 50% of government budget
  - About 15% of GDP
  - About $3 billion over the first PRSP period 2001-05 (five years).

Poverty in Mozambique

- Consumption poverty:
  - 1996-97: absolute poverty rate of 69%.
  - 2002-03: absolute poverty rate of 54%.
- Improvements in other indicators:
  - Rural income (Ag income survey)
  - Assets (CWIQ)
  - Health (DHS)
  - Service delivery (Administrative plus survey data)

Implications

- Monitoring information has strongly influenced the new PRSP (PARPA II)
- Current PRSP projects between $4-$5 billion in foreign assistance over the period 2006-2009 (four years).
- Enhanced importance of budget support.
- Off budget assistance projects coming on budget.
- INE and collaborators deserve credit.
- Even higher stakes in monitoring outcomes looking forward.

Concerns

- Massive over-reliance on a single household consumption survey in 2002-03.
  - No consensus on poverty evolution prior to publication of report despite:
    - CWIQ (including poverty predictors from the 2000 QUIBB)
    - National accounts etc.
- Set for heavy reliance on the 2008-09 household survey (PRSP target of 45% in absolute poverty by 2009 using a consumption metric).
- Repeated cross sections at five year intervals fail to provide adequate information for poverty analysis.

HH Consumption Panel Data

- At reasonably close intervals should be PART of the monitoring package in Mozambique
- Why?
  - Monitoring demand is very large.
  - Large underlying variability in living standards with implications for:
    - Policy
    - Monitoring.
  - Quite possible from the data collection perspective (ag income survey illustrated that this year).
  - Well set up for analysis.

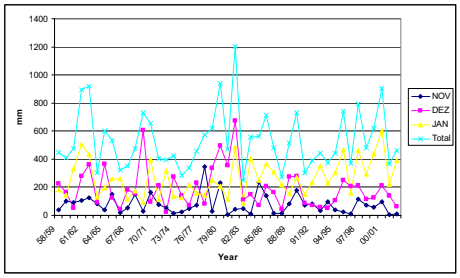
Variability in living standards

- Panel data on household consumption is the best method available for tracking variation in living standards.
- Due to the absence of panel data, there is very little direct evidence of variation in living standards experienced by households.
- Indirect evidence: many items important to the determination of living standards vary a lot.

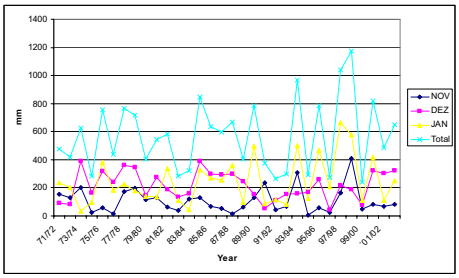
Occupation Shares

	All	Women	Men
Agriculture	79.6	89.3	67.5
Mining	0.4	0.1	0.7
Manufactures	0.8	0.1	1.6
Construction	2.2	0.1	4.7
Transport	1.1	0.1	2.3
Commerce	7.5	5.5	10.0
Service	5.2	3.3	7.6
Govt	3.3	1.5	5.5
Working	100	100	100
LFP Rate	80.6	82.0	78.8

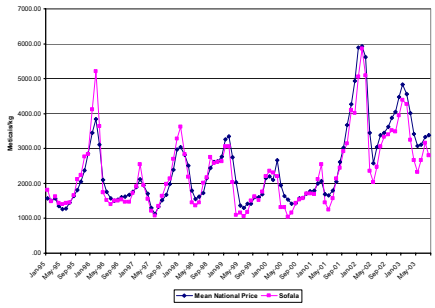
Rainfall in Quelimane



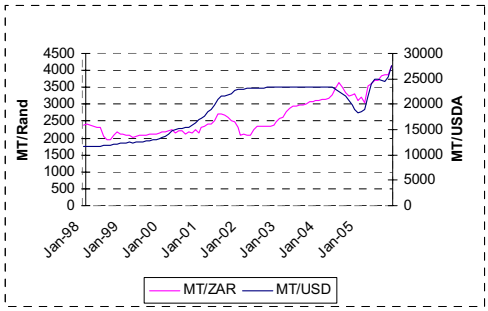
Rainfall in Beira



Maize Price Series



Nominal Exchange Rates



Poverty Headcount

	1996-97	2002-03	Difference
Nacional	69.4	54.1	-15.3
Urbano	62.0	51.5	-10.5
Rural	71.3	55.3	-16.0
Niassa	70.6	52.1	-18.5
Cabo Delgado	57.4	63.2	5.8
Nampula	68.9	52.6	-16.3
Zambezia	68.1	44.6	-23.5
Tete	82.3	59.8	-22.5
Manica	62.6	43.6	-19.0
Sofala	87.9	36.1	-51.8
Inhambane	82.6	80.7	-1.9
Gaza	64.6	60.1	-4.5
Maputo Prov	65.6	69.3	3.7
Maputo Cid	47.8	53.6	5.8

Correlation between 1996-97 and 2002-03 = 0

Conclusion: Living standards vary

Implications for Policy

- Extent of living standard variation
- Transitory poverty
- Chronic poverty
- Distribution of chronic and transitory poverty
- Livelihood strategies

Panel data on consumption is required to answer these questions.

Implications for Monitoring

- Dangerous to tie one's fortunes too tightly to one indicator (e.g., poverty headcount) from one survey in one year.
- What if the 2008-09 HH survey occurs in a bad year?
- Provincial outcomes likely more variable than national outcomes. [Risk of running from good year (2002-03) to bad year (2008-09?). Consider Zambezia and Nampula.]

Panel data communicates the underlying variability and gives better information on trends.

Data Collection

- Ag income (TIA 2005) demonstrates that panels are quite feasible.
- Intelligent sampling for panels. See paper on sub-sampling at www.mpd.gov.mz (DNEAP) also forthcoming in JRSS Series C.

Analysis Constraint

- Significant advances in analytical capacity.
- Technical advances
  - Poverty comparison across space and through time.
  - Reproduceability of results and reuseability of existing code.

Summary

- Regular panel data on consumption should form a PART of the poverty monitoring strategy.
  - Value for policy analysis
  - Reduction in monitoring risks
  - Other limits (financial, data collection, analysis) can be surpassed in the Mozambican context.
- CWIQ indicators also a part.

Owns a Functioning Bicycle

	IAF96	QUIBB00	IAF02
National	13.3	27.3	28.1
Rural	14	30.7	31.8
Urban	10	17.9	19.4
Niassa	24.1	47	56.9
Cabo Delgado	14.8	24.9	24.1
Nampula	10.9	23.3	26.7
Zambézia	13.9	46.8	38.7
Tete	20.3	37.1	27.9
Manica	18.3	25.9	38.5
Sofala	11.9	25.4	35.5
Inhambane	7.8	12.5	11.7
Gaza	14.4	15.3	16.7
Maputo Province	9.4	9.6	10.2
Maputo City	2.6	9.1	7.8

Correlation in differences: -.67

Owns a Functioning Radio

	IAF96	Census 1997	QUIBB00	IAF02
National	28.9	28.9	49.6	45.5
Rural	23.6	21.4	43.7	41.5
Urban	53.9	51.1	65.4	54.9
Niassa	27.3	22.6	39.2	43.0
Cabo Delgado	24.5	21.3	42.3	43.0
Nampula	16.5	20.8	49.4	48.3
Zambézia	22.5	20.9	48.9	39.4
Tete	29.4	26.3	41.2	45.1
Manica	35.1	34.3	53.7	63.6
Sofala	25.1	36.2	54.2	52.3
Inhambane	38.1	32.1	41.7	32.9
Gaza	42.7	37.3	43.6	34.1
Maputo Province	46.7	49.1	64.1	53.4
Maputo City	77.2	73.3	80.8	61.8

Correlation in differences = -.59

Thank you

Headcount Estimated from QUIBB 2000-01

	IAF 1996–97	QUIBB 2000–01	IAF 2002–03
National	69.4	60.7	54.1
Rural	71.3	64.5	55.3
Urban	62.1	51.6	51.5
Niassa	70.6	72.4	52.1
Cabo Delgado	57.4	50.7	63.2
Nampula	68.9	61.7	52.6
Zambézia	68.1	60.3	44.6
Tete	82.3	75.7	59.8
Manica	62.6	35.3	43.6
Sofala	87.9	81.5	36.1
Inhambane	82.6	69.3	80.7
Gaza	64.7	56.9	60.1
Maputo Province	65.6	49.4	69.3
Maputo City	47.8	41.0	53.6

Recommendation

- Follow a sub-sample of the IAF households through time (develop a panel data set).
- Benefits:
  - A more regular flow of information on poverty evolution through time.
  - Enhanced information on the characteristics of poverty. For example, do the poor tend to stay poor or do many people tend to move in/out of poverty?
  - Capacity building.

14.15. Large samples vs small area estimation

Large samples versus small area estimation for district level

A presentation held in Maputo, Mosambique
December 12-13, 2005

Stein Opdahl
Senior Adviser
Division for Development Cooperation
Statistics Norway
sop@ssb.no +47 920 13114

Statistics Norway

Estimation for small / local areas

Four basic types of strategies :

- Direct estimators
- Covering (or nearly unbiased estimators)
- Prediction (or regression estimators)
- Composite estimators

Statistics Norway

Sample:
Sample Plan and Accuracy

The accuracy to be obtained is dependent on three factors:

- The distribution (form and variance) of the actual variable
- The design of the sample (number of stages, stratification)
- The sample size

Statistics Norway

Sample:
Sample Plan and Accuracy

There are many determinants, factors and constraints in figuring sample size:

- Expected (or anticipated) size of survey estimates or indicators
- Expected (or anticipated) size of target population(s)
- Expected household size
- Margin of error tolerated
- Level of confidence wanted
- "Design effect" in cluster surveys
- Expected non-response rate
- Numbers of clusters or PSUs
- Cluster size (number of households per sample cluster)
- Number of estimation "domains" wanted
- Survey budget

Statistics Norway

Sample:
Sample Plan and Accuracy

Example :

- Binomially distributed variable
- Normal distribution as proxy
- Multistage sampling design
- Error tolerance:
  (a) Confidence level α =0.05
  (b) Relative error margin

Statistics Norway

Sample:
Sample Plan and Accuracy

Minimum sample size required; sampling unit = unit of analysis

$$n = \frac{(3.84)(1-p)(1.1)(Deff)}{(p)(V^2)}$$

Household sample; analysing individuals **n =** $n_h \bar{h}$

$$n_h = \frac{(3.84)(1-p)(1.1)(Deff)}{(p)(\bar{h})(V^2)}$$

Statistics Norway

Nearly unbiased estimate

$$\bar{x}'_a = \left(\sum_{j=1}^{J} \left(\frac{n_{aj}}{n_{.j}}\right) x'_j\right) / n_{a.}$$

x'_j = the survey unbiased estimate of the total or aggregate level of X in stratum j.

n_{aj} = the number of persons in stratum j that belong to area a.

$n_{.j}$ = the total number of persons in stratum j.

$n_{a.}$ = the total number of persons in area a.

J = the total number of strata in the survey.

Statistics Norway

Small Area Estimation:

Regression estimators

Variable $\hat{\bar{Y}}_j$

Basic model:

$$y_{ij} = \alpha + \beta \bar{X}_j + u_j + e_{ij}$$

where

- y_{ij} is the survey variable of interest for individual/case i within area of interest j
- $\bar{X}_j$ is the (known) population mean for the covariate in area of interest j
- α and β are the regression parameters for intercept and slope respectively
- u_j is a random area-level term that is assumed to have expectation 0 and variance σ_u^2.
- e_{ij} is a random individual-level term with expectation 0 and variance σ_e^2.

Statistics Norway

Small Area Estimation:

Synthetic Estimator

Estimator:

$$\hat{\bar{Y}}_{j,SYNTH} = \hat{\alpha} + \hat{\beta}\bar{X}_j$$

Bias:

The synthetic estimator is not design-unbiased, since what it estimates is the underlying expected value for any area with the same covariate values, not the real value for the area in question.

Statistics Norway

Small Area Estimation:

Synthetic Estimator

Mean Square Error:

$$MSE(\hat{\bar{Y}}_{j,SYNTH}) = \sigma_u^2 + \mathrm{var}(\hat{\alpha} + \hat{\beta}\bar{X}_j)$$

I) Real variance in particular areas around the underlying expected values

II) Sampling variance of the synthetic estimator

Confidence interval:

$$\hat{\bar{Y}}_{j,SYNTH} \pm 2\sqrt{\widehat{MSE}(\hat{\bar{Y}}_{j,SYNTH})}$$

Statistics Norway

Small Area Estimation:

A Generalised Regression estimator

$$\hat{\bar{Y}}_{j,GREG} = \bar{y}_j + \hat{\beta}(\bar{X}_j - \bar{x}_j) = \hat{\beta}\bar{X}_j + (\bar{y}_j - \hat{\beta}\bar{x}_j)$$

- The GREG adjusts the sample mean to take account of any differences in the sample and population area means and covariates, i.e. it provides an adjustment to the sample area mean which adjusts for an unrepresentative sample

Statistics Norway

Small Area Estimation:

A Composite estimator

$$\hat{\bar{Y}}_{j,COMP} = \hat{\gamma}_j \hat{\bar{Y}}_{j,GREG} + (1 - \hat{\gamma}_j)\hat{\bar{Y}}_{j,SYNTH}$$

- This estimator uses area data (where available) to improve SYNTH estimation.

Statistics Norway

Small Area Estimation:

Other issues

- The linear relationship and the Normal variance assumption may not be valid or reasonable ---- Transformation of data
- Selection of covariates - criteria for selection/omission
- Diagnostics – checks for model misspecifications
- Non-response bias
- Calibration
- Changes in administrative units
- GIS technology – georeferenced data
- Potential impact of the ecological fallacy

Statistics Norway
Statistisk sentralbyrå

GHANA CENSUS-BASED POVERTY MAP: District and Sub-District level Results[1]

Harold Coulombe

Ghana Statistical Service
Department for International Development, UK
May 2005

Abstract

This paper documents the construction and presents the main results of a Ghanaian poverty map based on the GLSS4 survey and the Census 2000. The methodology takes advantages of detailed information found in the survey and the exhaustive coverage of the census. It permits the calculation of poverty indicators at a very low level of desegregation; sub-district in the case of Ghana. In the current paper district level poverty figures are presented. Council level estimates are also available.

Introduction

1. This paper documents the construction of a poverty map based on data on the fourth round Ghana Living Standards Survey (GLSS4) and the Housing & Population Census 2000. Based on a recently developed methodology, it permits the calculation of poverty indicators at very low levels of aggregation, using the detailed information found in the survey and the exhaustive coverage of the Census. Results at district level as well as at the town and area council level are presented and analyzed.

2. In the past decade poverty profiles have been developed into useful tools to characterize, assess and monitor poverty. Based on information collected in household surveys, including detailed information on expenditures and incomes, those profiles present the characteristics of the population according to their level of - monetary and non-monetary - standard of living, help assessing the poverty reducing effect of some policies and compare poverty level between regions, groups or over time.

3. While these household-based studies have greatly improved our knowledge of welfare level of households in general and of the poorer ones in particular, the approach has a number of constraints. In particular, policy makers and planners have need of finely disaggregated information in order to implement their anti-poverty schemes. Typically they need information for small geographic units such as city neighbourhoods, towns or villages. Telling a Ghanaian policy maker the neediest people are in the Savannah region would not be too impressive as that information is well known and not useful since it would be too vague; telling them in which villages or towns or even districts the poorest households are concentrated would be more convincing! Using regional-level information often hides the existence of poverty pockets in otherwise relatively well-off region which would lead to poorly targeted schemes. Having better information at local level would necessarily minimize leaks and therefore permit more cost-effective and efficient anti-poverty schemes. Poverty indicators are needed at a local level as spatial inequalities can be important within a given region.

4. This paper presents results at regional, district and council levels. The methodology used have been developed by Elbers, Lanjouw and Lanjouw (2002, 2003) and should be seen as more sophisticated than other methods as it uses information on household expenditure, is fully consistent with poverty profile figures, and permits the computation of standard errors of those poverty indicators. Since those types of poverty maps are fully compatible with poverty profile results, they should be seen as a natural extension to the Poverty Profile, a way to operationalise poverty profile results. The current poverty map would reach its full potential once a series of applications under consideration would be undertaken.

5. The remaining of this paper is structured as follow: we first present the methodology in layman words, follow by a description of the data used. The paper ends by a discussion of the results and on furthers work to undertake. A more technical presentation of the methodology can be found in annex, along with some detail results.

Methodology

6. The basic idea behind the methodology is rather straightforward. First a regression model of adult equivalent expenditure is estimated using GLSS survey data, limiting the set of explanatory variables to those which are common to both that survey and the latest Census. Next, the coefficients from that model are applied to the Census data set to predict the expenditure level of every household in the Census. And finally, these predicted household expenditures are used to construct a series of welfare indicators (e.g. poverty level, depth, severity, inequality) for different geographical subgroups.

7. Although the idea behind the methodology is conceptually simple, its proper implementation requires complex computations. Those complexities are mainly coming from the need to take into account spatial autocorrelation (expenditure from households within the same cluster are correlated) and heteroskedasticity in the development of the predictive model. Taking into account those econometric issues insure unbiased predictions. A further issue making computation non-trivial is our willingness to compute standard errors for each welfare statistics. Those standard errors are important since they would tell us how low we can disaggregate the poverty indicators. As we disaggregate our results at lower and lower level, the number of households on which our estimates are based decrease as well and therefore yields less and less precise estimates. At a given level, the estimated poverty indicators would become too imprecise to be use with confidence. The computation of those standard errors would help us to decide where to stop the disaggregation process. The methodology used is further discussed in annex 1.

[1] This project was funded by the Department for International Development (DfID), along with some initial funding by the World Bank. The work was undertaken by Harold Coulombe (consultant) with full support from the Ghana Statistical Service staff, in particular Dr. Grace Bediako, Dr. Nicholas Nsowah-Nuamah, Dr. K.A. Twum-baah, Ms Rosalind Quartey, Mr. KB Danso-Manu and Ms. Jacqueline Anum. Technical guidance from Peter Lanjouw, Berk Özler and Johan Mistiaen (World Bank) as well as full support from Richard Harris (World Bank, previously DfID) were also welcomed. We also want to thank participants at the "Ghana at the Half Century" conference held in Accra in July 2004. The author can be reached at hcoulombe@videotron.ca.

Data

8. The construction of such poverty map is also very demanding in terms of data. The uttermost requirement is a household survey having expenditure modules and a population and housing census. If not already done, a monetary-based poverty profile would have to be constructed from the survey. The household-level welfare index and the poverty line from such poverty profile would be used. Apart from household-level information, community level characteristics is also useful in the construction of poverty map as differences in geography, history, ethnicity, access to markets, public services and infrastructure, and other aspects of public policy can all lead to important differences in standard of living, defined in monetary terms or not. Fortunately all that information was available in the case of Ghana.

Census:

9. The latest Housing and Population Census was conducted in spring 2000. The questionnaire is relatively detailed but does not contain any information on neither incomes nor expenditures. At the individual level, it covers demography, education and economic activities. At the household level, dwelling characteristics are well covered. The Census database turns out more than 18.9 million individuals grouped into 3.7 million households. The Census field work grouped households into around 26,800 enumeration areas (EAs) of 138 households each on average.

10. Along with the housing and population census a *facility census* was conducted in every single locality. Those "localities" go from tiny sub-EAs settlements to large urban neighbourhood having many EAs. There is around 89 000 "localities" in the facility census database. The information collected includes the existence in the locality of a post office, telephone, traditional healing centre, hospital, maternity/clinic, and primary, JSS and SSS schools. If any of those facilities was not found in the locality, the distance to the nearest one was asked.

GLSS4 Survey:

11. The fourth round of the Ghana Living Standard Survey is the latest national survey having collected expenditure data at household level. Having been administrated in 1998/99, it is also the most appropriate survey time wise. The survey dataset was also enhanced by including information from the facility census. This required a tedious matching exercise to link the Enumeration Areas (EAs) used as sampling units (clusters) in the GLSS - which were based on the 1984 Census - with the 2000 Census EAs.

12. The welfare index to be used in our regression models (expenditure per equivalent adult in real terms) is the same as the one used in the Government-sponsored poverty profile based on GLSS4. Using the same welfare index would ensure full consistency between the latest poverty profile (GSS, 2000; Coulombe and McKay, 2003) and the new poverty map. It will also permit to test whether the predicted poverty indicators match those found in the poverty profile at strata level, the lowest statistically robust level achievable in GLSS 4.

13. On the basis of the information collected in the latest Census a number of GLSS 4 localities have been reclassified from rural to urban - an urban location is one with 5,000 or more persons. However the urban/rural variable use in GLSS4 was defined on the basis of information from the 1984 Census. Therefore many EAs (clusters) in GLSS4 had been considered rural while they surely became urban by 1998/99 when GLSS 4 was conducted. This phenomenon is illustrated by figures in Table 1. Compared to the latest Census the urban localities (outside Accra) are underrepresented in GLSS4 while the rural ones are overrepresented. The problem is particularly important in Coastal and Forest ecological zones. For the current study 24 clusters have been redefined from rural to urban. The last column of Table 1 clearly shows that the new GLSS 4 distribution of clusters across strata is much more similar to the Census one and therefore, closer to the reality at the time of GLSS4 (1998/99).

[Insert Table 1 about here]

Administrative Layers

14. Ghana is currently in the process of an important decentralisation effort which formally started more than ten years ago. The Local Government Act of 1993 and the National Development Planning (Systems) Act of 1994 have defined the current local government structure. The structure consists of four tiers. The top tier is the Regional Coordinating Council, followed by the Metropolitan/Municipal/District Assemblies. The Town/Zonal/Urban/Area Councils and the Unit Committees are the bottom two tiers. However, the implementation of this administrative structure was held back by limited financial and human resources (Awoosah *et al.* 2004). In practice, only regions and districts have been formally defined. In our study, we use the official definitions for the regions and the districts, as well as an unofficial definition of the different type of councils. No attempt was made to define the last tier. Table 2 presents some descriptive statistics on the size of those different administrative levels.

[Insert Table 2 about here]

15. *Strata*: the GLSS 4 sample design was based on seven strata defined in terms in agro-climatic zones (coastal, forest and savannah) and urban/rural breakdown. Although that level is not an administrative level, poverty estimates were done at this fairly aggregated level mainly to establish the statistical validity of the poverty estimates. Those predicted figures can be compared with actual figures found in the latest Ghana Poverty Profile and statistical tests performed on the equality of those indicators.

16. *Region*: the national territory is divided into 10 regions which are further down divided into districts. No districts overlapped two or more regions.

17. *District*: the lowest administrative level for which a formal geographical definition is currently available is the 110 districts. The importance of the District Assemblies in the on-going decentralisation process makes district-level poverty figures fundamental. Those poverty figures, presented in this report, are the first value-added product coming out from the poverty map. In 2004, a district remapping has yield 28 new districts but unfortunately the information needed to perform the poverty map using this new district definition was not available on time for this study. Once an operational EA-based definition of the 138 districts become available, it would be easy to update the poverty map to reflect the new administrative reality.

18. C*ouncil*: although district-level poverty estimates would surely be useful, that level of politico-geographical breakdown could still be too aggregated to be used for more finely targeted interventions. Currently there is no properly mapped sub-district breakdown. Each District Assembly has created a series of sub-district councils, broadly defined – in words - in a series of Legislative Instruments (LI) from 1988, prior the formal establishment of the

current four-tier system. However, those councils do not have formally mapped boundaries. Based on those LI, a Ghana Statistical Service team from cartography and GIS departments has been able to establish the link between those "councils" and the Census 2000 EAs. Although the definition of those councils was not made official, we believe it would be a very decent approximation to an on-going data collection exercise being done by CERSGIS from the University of Ghana at Legon[2]. All together, we defined 1048 councils. These units would be small enough for most decision making while being large enough to enable a statistically robust poverty maps to be computed.

Results

19. In order to maximise accuracy we have estimated the model at the lowest geographical level for which the GLSS survey is representative. In the case of the fourth round of GLSS that level is the sampling strata: Accra, Urban Coastal, Urban Forest, Urban Savannah, Rural Coastal, Rural Forest and Rural Savannah. A household level expenditure model has been developed for each of these strata using explanatory variables which are common to both the GLSS and the Census. Those variables do not need to be causal as we are only interested in their predictive power. The results are presented stage by stage.

Stage 1: Aligning the data

20. The first task was to make sure the variables deemed common to both the census and the survey were really measuring the same characteristics. In the first instance, we compared the questions and modalities in both questionnaires to isolate potential variables. We then compared the means of those (dichotomized) variables and tested whether they were equal using a 95% confidence interval[3]. Restricting ourselves to those variables would ensure the predicted welfare figures would be consistent with survey-based poverty profile. As noted above that comparison exercise was done at strata level. The two-stage sample design of GLSS 4 was taken into account in the computation of the standard errors. The results are not presented here but are available on request.

Stage 2: Survey-based regressions

21. Table in annexe 2 presents the strata-specific regression results based on GLSS 4. The ultimate choice of the independent variables was based on a backward stepwise selection model. A check of the results confirmed that almost all the coefficients are of expected sign. As said earlier, those models are not for discussions. They are exclusively prediction model, not determinant of poverty models that can be analyzed in terms of causal relationships. In the models used for the poverty map we were only concerned by the predictive power of the regressions without regards, for example, for endogenous variables. At that stage, we attempt to control location effect by incorporating cluster average of some of the variables. We also ran a series of regressions using the base model residuals as dependant variables. Those

[2] The Centre for Remote Sensing and Geographic Information Services (CERSGIS) is working on a comprehensive project which involves exhaustive data collection and mapping, including the definition of the councils which would eventually be made official by the Government of Ghana. The project, called Establishing a Mapping and Monitoring System for Development Activities in Ghana (EMMSDAG), is co-sponsored by the Ministry of Finance and the European Union. Final results are not expected before a year.

[3] We also deleted or redefined dichotomic variables being less that 0.03 or larger than 0.97 to avoid serious multicollinearity problems in our econometric models.

results – not shown here – would be used in the last stage in order to correct for heteroskedasticity.

22. The R^2s of the different regressions vary from 0.27 to 0.60. Although they might appear to be on the low side, they are typical of survey-based cross-section regressions and can be favourably compared with results from other poverty maps. While those coefficients look "credible", it is important to note those models were purely predictive in the statistical sense and should not be view as determinant of welfare or poverty. The relatively low R^2s for some of the models are mainly due to four important factors. First, in many areas households are fairly homogeneous in terms of observable characteristics even if there consumption varies relatively more. That necessarily yields low R^2. Second, a large number of potential correlates are simply not observables using standard closed-questionnaire data collection methods. Third, many good predictors had been discarded at first stage since their distributions did not appears to be identical. And finally, many indicators do not take into account the quality of the correlates. Not taking into account the wide variation in quality of the different observable correlates makes many of those potential correlates useless in term of predictive power.

Stage 3: Welfare indicators[4]

23. Based on the results from the previous stage, we applied the estimated parameters[5] to the Census data to compute a series of poverty and inequality indicators: the headcount ratio (P_0), the poverty gap index (P_1), the poverty severity index (P_2), the Gini Index, the mean log deviation and the Theil index[6]. Table 3 presents estimated poverty figures for each stratum and compares them with actual figures from the latest survey-based poverty profiles. For each stratum and poverty indicators, the equality of GLSS 4-based and Census-based indicators cannot be rejected (at 95%)[7]. Apart the case of Urban Forest where the census-based headcount ration is 3.2 points higher, the gaps are always smaller that 1.5% and often minute. Although census-based poverty figures can only be compared with the ones provided by the GLSS survey at stratum level, equality of those poverty figures provided an excellent reliability test of the methodology used here.

[Insert Table 3 about here]

24. Using the same econometric results, table in annexe 3 presents poverty figures for each of the 10 regions and 110 districts, broken down into urban/rural areas. The standard errors are also presented and are – for most cases – relatively small which make the predicted poverty figures quite reliable. Those district-level estimates are the first ever monetary-based poverty figures available in Ghana. Overall those figures seem to make sense and anecdotal evidences support those results although some results might raise question at first look. In particular, in a few districts the urban population are found to be poorer than the rural population. However those districts tend to be isolated ones where the so-called urban

[4] The computation of the welfare indicator has been greatly eased thank to PovMap, a software especially written to implement the methodology used here. We used the February 2005 version developed by Qinghua Zhao (2005).

[5] Apart from regression models explaining household welfare level, we also estimated a model for the heteroskedasticity in the household component of the error. We also estimated the parametric distributions of both error terms. See the methodological annexe for further details.

[6] Because of space constraint, only the poverty figures are presented in this paper. The inequality figures would be found in a forthcoming GSS report. That report would be an extended version of the current paper.

[7] It is worth noting that the standard errors of the mean of the Census-based figures are systematically lower that the ones calculated from GLSS 4.

population are likely to live in "big" villages not having the infrastructure usually found in Ghanaian towns.

25. Council-level figures were also computed but space constraint does not permit their presentation. Those council-level results are available in an exhaustive companion report published by the Ghana Statistical Service. Nonetheless some analysis concerning the relevancy of those finely disaggregated estimates can be found next.

*How low **can** we go?*

26. Further examination of poverty estimates from the table in annexe 3 reveals that the standard errors - in relation to their associated indicators - seem to indicate our poverty estimates at district level are fairly precise. However, it is difficult to make an "objective" judgement on the precision of those estimates without some kind of benchmark. To do so, Figure 1 presents the headcount incidence coefficients of variation (inverted) of the district- and council-level estimates and compared them to the ones computed from the GLSS 4 survey. Hence, we use the precision of the GLSS4-based headcount incidence as our benchmark which is represented by the step curve. Those steps represented the different inverted coefficient of variation associated with the different stratum. The curves in Figure 1 clearly show that our district-level headcount incidence estimates does at least as well as GLSS4-based poverty estimates since the district-level curve lie on or below the GLSS4 one. Since council estimates are based on smaller samples, its curve shows that the council-level estimates are not as precise although they compared favourably with the GLSS4 figures. How low can we go? If one take the GLSS 4 benchmark as a good one, it is clear that both district- and council-level poverty estimates would be good guides to policy-makers.

[Insert Figure 1 about here]

*How low **should** we go?*

27. Although we just demonstrate that we can used the district and council headcount figures with some confidence about their precision level, it might be the case that those disaggregated figures does not yield much information. Within a rather homogenous region, it might be possible that the different districts are not statistically different from each others in terms of monetary poverty. The same question can be raised concerning the use of the council-level figures within a given district. To test whether additional information about the poverty level is gain when we disaggregated from regions to districts and from districts to councils, table 4 gives the proportion of districts (in terms of unit and of population) that are statistically poorer or richer than their associated regions. We also computed the relationship between districts and councils. Overall, some 36.6% of the different districts have a poverty headcount statistically smaller or higher than their own region. Similarly, 13.7% of councils are different poverty wise from their own district. In terms of population the overall figures are significantly higher at repetitively 45.2% and 13.2%. As expected, those figures show that urban areas are less homogeneous than the rural areas. At least in rural areas, it also clear that the smaller entity the more homogeneous they are. Urban areas are visibly more heterogeneous. Based on those results, it appears that using the results from councils on the top of those from districts should improve the targeting efficiency of any allocation of resources aiming at reducing poverty.

[Insert Table 4 about here]

Concluding Remarks

28. This paper has documented the construction of a regional-, district- and council-level poverty map for Ghana. The methodology developed by Elbers *et al.* (2003) has permitted to obtain the first ever reliable poverty estimates at the district and council levels. That map reports on 110 districts but it would be easy and straightforward to update it once we obtain the definition (in terms of EAs) of the recently redrawn districts. However, we acknowledge that the definition of the councils is our own and should not be view as official. Those finely disaggregated poverty figures are fully compatible with the latest Poverty Profile (GSS, 2000; Coulombe and McKay, 2003).

29. One of the main advantages of the methodology used here is the possibility of computing standard errors of the different poverty estimates and therefore has an idea of the reliability of those estimates. We viewed that using the precision level of the latest poverty profile as benchmark, both the district- and council- level are precise enough to be useful to planners, policy-makers and researchers.

30. However interesting those results, they would acquire their full potential if they are use. How? Amongst others, those results can be used to design budget allocation rules to be applied by the different administrative levels toward their subdivisions: the central government toward the districts, and the districts toward their councils. That map could become an important tool in support of the decentralization process currently undertaken in Ghana. Obviously such monetary-based target indicators could be used in conjunction with some alternative measures of poverty based on education, health or infrastructure indicators. In particular merging the poverty map with education and health maps would yields powerful targeting tools. Others uses of the poverty map would include the evaluation of locally targeted anti-poverty schemes (Social funds, Town/village development schemes), impact analysis etc. And finally, researchers could use it in a multitude of ways such as the study of relationship between poverty distribution and different socio-economic outcomes

References

Awoosah, E. T., J. C. Agamah, A. S. Kwarteng, BM Oppong and F. Owusu, 2004, Planning with the area council: experience with CBP in Ghana, *in* PLA Notes 49: Decentralisation and community-based planning, edited by Ian Goldman and Joanne Abbott, International Institute for Environment and Development (IIED)

Coulombe, Harold and Andrew McKay, 2003, Selective Poverty Reduction in a Slow Growth Environment: Ghana in the 1990s, mimeo

Elbers, Chris, Jean Olson Lanjouw, and Peter Lanjouw, 2002, "Welfare in Villages and Towns: Micro level Estimation of Poverty and Inequality." Policy Research Working Paper No. 2911, DECRG-World Bank, Washington DC

Elbers, Chris, Jean Olson Lanjouw, and Peter Lanjouw, 2003, "Micro-Level Estimation of Poverty and Inequality" *Econometrica*, 71(1), 355-364

Ghana Statistical Service (GSS), 2000, *Poverty Trends in Ghana in the 1990s*, Accra, October 2000

Mistiaen, Johan, Berk Ozler, Tiaray Razafimanantena and Jean Razafindravonona, 2002, Putting Welfare on the Map in Madagascar, Africa Region Working Paper Series, Number 34, The World Bank. Washington, D.C.

Zhao, Qinghua, 2005, User Manual for PovMap, mimeo, Development research Group, The World Bank, Washington, D.C.

Table 1: Distribution of households according to strata and ecological zone, GLSS 4 and Census 2000

	GLSS 4		Census 2000	Difference	
	1984 urban/rural definition	2000 urban/rural definition		1984 urban/rural definition	2000 urban/rural definition
GLSS 4 strata					
Accra	10.51	10.51	9.86	+0.65	+0.65
Urban Coastal	8.55	13.44	14.07	-5.52	-0.63
Urban Forest	13.11	18.08	17.98	-4.87	+0.10
Urban Savannah	4.49	6.68	5.26	-0.77	+1.42
Rural Coastal	15.46	10.84	11.58	+3.88	-0.74
Rural Forest	30.55	24.61	26.07	+4.51	-1.46
Rural Savannah	17.33	15.84	15.18	+2.15	+0.66
Ecological zone					
Accra	10.51	10.51	9.86	+0.65	+0.65
Coastal	24.02	24.28	25.65	+0.34	-1.37
Forest	43.66	42.69	44.05	-2.37	-1.36
Savannah	21.82	22.52	20.44	+1.38	+2.08
Total	100.0%	100.0%	100.0%	0.0%	0.0%

Sources: author's calculation based on GLSS4 and Census 2000

Table 2: Descriptive Statistics on the Ghanaian Administrative Structure

Administrative Unit	# of Units	Number of Households			Number of Individuals		
		Median	Minimum	Maximum	Median	Minimum	Maximum
Region	10	355,263	80,573	680,419	1,810,044	574,918	3,590,511
District	110	24,852	9,912	364,805	133,154	51,918	1,647,202
Council	1,048	2,055	41	48,334	12,258	263	272,208

Source: Author's calculation based on the Census 2000
Note: Although 263 individuals seem rather small for a council, only 8 councils (out of 1048) have less than 1000 people.

Table 3: Poverty Rates based on GLSS 4 (actual) and Census 2000 (predicted), by strata

	Headcount Incidence (P_0)		Poverty Gap Index (P_1)		Poverty Severity Index (P_2)	
	GLSS4 (Actual)	Census (Predicted)	GLSS4 (Actual)	Census (Predicted)	GLSS4 (Actual)	Census (Predicted)
Accra	0.038	0.052	0.008	0.012	0.002	0.004
	(0.017)	*(0.009)*	*(0.004)*	*(0.002)*	*(0.001)*	*(0.001)*
Urban Coastal	0.286	0.280	0.085	0.098	0.035	0.049
	(0.040)	*(0.020)*	*(0.016)*	*(0.009)*	*(0.008)*	*(0.006)*
Urban Forest	0.176	0.208	0.047	0.074	0.018	0.037
	(0.036)	*(0.013)*	*(0.011)*	*(0.007)*	*(0.005)*	*(0.004)*
Urban Savannah	0.518	0.510	0.162	0.183	0.067	0.088
	(0.078)	*(0.041)*	*(0.036)*	*(0.021)*	*(0.018)*	*(0.013)*
Rural Coastal	0.485	0.471	0.152	0.163	0.065	0.076
	(0.046)	*(0.025)*	*(0.023)*	*(0.013)*	*(0.012)*	*(0.008)*
Rural Forest	0.409	0.407	0.117	0.137	0.048	0.064
	(0.025)	*(0.021)*	*(0.012)*	*(0.010)*	*(0.007)*	*(0.006)*
Rural Savannah	0.695	0.690	0.324	0.331	0.181	0.197
	(0.054)	*(0.023)*	*(0.036)*	*(0.018)*	*(0.024)*	*(0.015)*

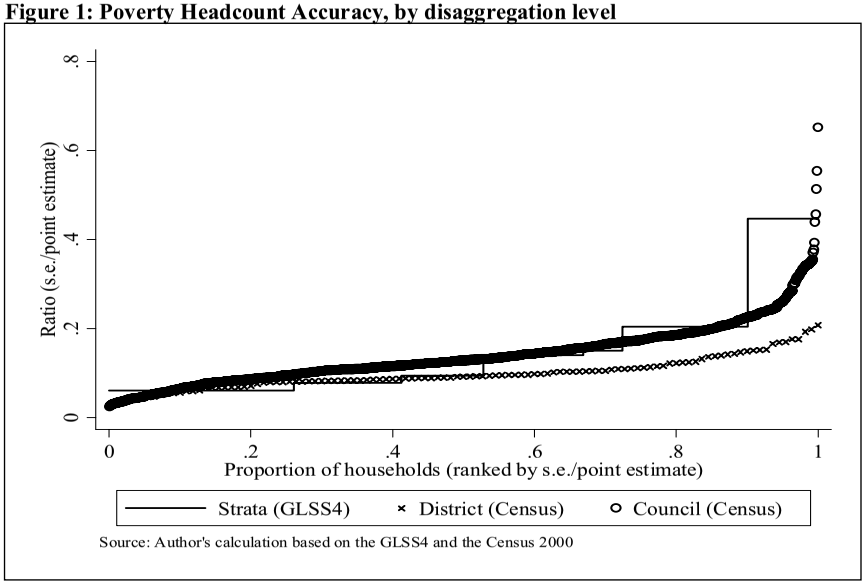
Sources: author's calculation based on GLSS4 and Census 2000
Notes: Robust standard errors in parentheses. The poverty indicators based on GLSS4 are slightly different from the ones already published by GSS since we used the new definition of the urban/rural breakdown (see table 1).

Table 4: Disaggregation and Change in Headcount Incidence, by Region

	% of geographic unit		% of the population	
	Districts different from their **Regions**	**Councils** different from their **Districts**	**Districts** different from their **Regions**	**Councils** different from their **Districts**
Urban	*43.4*	*19.8*	*66.0*	*15.0*
Western	27.3	18.2	16.5	26.5
Central	16.7	16.7	21.5	9.7
Greater Accra	100.0	14.8	100.0	7.3
Volta	33.3	10.0	39.3	8.9
Eastern	40.0	14.7	48.9	10.7
Ashanti	82.4	26.9	87.7	22.2
Brong Ahafo	46.2	33.3	59.8	28.1
Northern	38.5	10.5	20.0	6.6
Upper East	25.0	50.0	5.5	49.2
Upper West	0.0	0.0	0.0	0.0
Rural	*29.9*	*11.9*	*29.1*	*11.7*
Western	30.0	3.4	27.2	1.7
Central	83.3	27.3	85.8	23.1
Greater Accra	25.0	42.1	16.4	54.5
Volta	0.0	14.0	0.0	15.3
Eastern	26.7	9.0	18.8	9.0
Ashanti	11.8	6.9	8.1	7.5
Brong Ahafo	7.7	9.4	8.9	9.7
Northern	53.8	15.8	62.2	12.9
Upper East	66.7	6.9	66.7	5.0
Upper West	0.0	10.9	0.0	14.7
Total	*36.6*	*13.7*	*45.2*	*13.2*

Sources: author's calculation based on GLSS4 and Census 2000
Notes: following Mistiaen *et al*. (2002), those percentages represent difference in headcount incidence that are statistically different (at 95% confidence interval) using the standard errors of the point estimates for the lower level of disaggregation.

Figure 1: Poverty Headcount Accuracy, by disaggregation level



Source: Author's calculation based on the GLSS4 and the Census 2000

Annexe 1: Methodology

The basic idea behind the methodology developed by Elbers, Lanjouw and Lanjouw (2002, 2003) is unchallenging. At first a regression model of log of per capita expenditure is estimated using survey data, employing a set of explanatory variables which are common to both a survey and the census. Next, parameters from the regression are used to predict expenditure for every household in the census. And third, a series of welfare indicators are constructed for different geographical subgroups.

The term "welfare indicator" embrace a whole set of indicators based on household expenditures. This note put emphasis on poverty headcount (P_0) but the usual poverty and inequality indicators can be computed (Atkinson inequality measures, generalised Entropy class inequalities index, FGT poverty measures and Gini).

Although the idea is rather simple its proper implementation require complex computation if one want to take into account spatial autocorrelation and heteroskedasticity in the regression model. Furthermore, proper calculation of the different welfare indicators and its standard errors increase tremendously its complexities.

The discussion below is divided into three parts, one for each stage necessary in the construction of a poverty map. This discussion borrows from the original theoretical papers of Elbers, Lanjouw and Lanjouw as well as on Mistiaen *et al.* (2002).

First stage

In the first instance, we need to determine a set of explanatory variables from both databases that are meeting some criteria of comparability. In order to be able to reproduce a poverty map consistent with the associated poverty profile, it is important to restrict ourselves to variables that are fully comparable between the census and the survey. We start by checking the wording of the different questions as well as the proposed answer options. From the set of selected questions we then build a series of variables which would be tested for comparability. Although we might want to test the comparability of the whole distributions of each variable, in practice we restrain ourselves to test only the means. In order to maximise the predictability power of the second-stage models all analysis would be performed at the strata level, including the comparability of the different variables from which the definitive models would be determined.

The list of all potential variables and their equality of means test results are not presented in this note but can be obtained on request.

Second stage

We first model per capita household expenditure[8] using the limited sample survey. In order to maximise accuracy we estimate the model at the lowest geographical level for which the survey is representative. In the case of the fourth round of GLSS that level is the sampling strata: Accra, urban costal, urban forest, urban savannah, rural coastal, rural forest and rural savannah.

[8] In our study we used the welfare index constructed for the GLSS4 poverty profile. Although that welfare index is defined in terms of equivalent adults, the demonstration remains unchanged.

Let specify a household level expenditure (y_{ch}) model for household *h* in location *c,* $\mathbf{x_{ch}}$ is a set of explanatory variables, and u_{ch} is the residual:

$$\ln y_{ch} = \mathrm{E}[\ln y_{ch} \mid \mathbf{x_{ch}}] + u_{ch} \qquad (1)$$

The locations represent clusters as defined in the first stage of typical household sampling design. It usually also represents census enumeration areas, although it does not have to be. The explanatory variables need to be present in both the survey and the census, and need to be defined similarly. It also needs to have the same moments in order to properly measure the different welfare indicators. The set of potential variables had been defined in the first stage.

If we linearise the previous equation, we model the household's logarithmic per capita expenditure as

$$\ln y_{ch} = \mathbf{x_{ch}}'\boldsymbol{\beta} + u_{ch}. \qquad (2)$$

The vector of disturbances $\boldsymbol{u}$ is distributed $F(0,\Sigma)$. The model (2) is estimated by Generalised Least Square (GLS). To estimate this model we need first to estimate the error variance-covariance matrix Σ in order to take into account possible spatial autocorrelation (expenditure from households within a same cluster are surely correlated) and heteroskedasticity. To do so we first specify the error terms as

$$u_{ch} = \eta_c + \varepsilon_{ch} \qquad (3)$$

where η_c is the location effect and ε_{ch} is the individual component of the error term.

In practice we first estimate equation (2) by simple OLS and use the residuals as estimate of the overall disturbances, given by $\hat{\mu}_{ch}$. We then decomposed those residuals between uncorrelated household and location components:

$$\hat{u}_{ch} = \hat{\eta}_c + e_{ch} \qquad (4)$$

The location term ($\hat{\eta}_c$) is estimated as cluster means of the overall residuals and therefore the household component (e_{ch}) is simply deducted. The heteroskedasticity in the latest error component is modelled by the regressing its squared (e_{ch}^2) on a long list of all independent variables of model (2), their squared and interactions as well as the imputed welfare. A logistic model is used.

Both error computations are used to produce two matrices which are them sum to $\hat{\Sigma}$, the estimated variance-covariance matrix of the original model (2). That latest matrix permits to estimate the final set of coefficients of the main model (2).

Third stage

To complete the map we associate the estimated parameters from the second stage with the corresponding characteristics of each household found in the census to predict the log of per capita expenditure and the simulated disturbances.

Since the very complex disturbance structure has made the computation of the variance of the imputed welfare index intractable, bootstrapping techniques have been used to get a measure of the dispersion of that imputed welfare index. From the previous stage, a series of coefficients and disturbance terms have been drawn from their corresponding distributions. We then, for each household found in the census, simulate a value of welfare index ($\hat{y}_{ch}^r$) based on the predicted values and the disturbance terms:

$$\hat{y}_{ch}^r = \exp(\mathbf{x}_{ch}'\tilde{\beta}^r + \tilde{\eta}_c^r + \tilde{\varepsilon}_{ch}^r) \qquad (5)$$

That process is repeated 100 times, each time redrawing the full set of coefficients and disturbances terms. The means of the simulated welfare index become our point estimate and the standard deviation of our welfare index is the standard errors of these simulated estimates.

Annexe 2: Survey-Based Regression models

Accra	
# of observations	620
# of clusters	31
R^2 (without location means)	0.2444
R^2 (with location means)	0.2659
Variable	Coeff.
# of boys aged 7-14	-0.168 *(4.93)*
# of girls aged 7-14	-0.161 *(6.51)*
Head schooled (0/1)	0.199 *(2.83)*
Head is self-employed, non-agro (0/1)	0.141 *(3.25)*
Cement Roof (0/1)	0.143 *(2.25)*
Has flush toilet (0/1)	0.148 *(2.75)*
Use coal for cooking (0/1)	-0.254 *(7.00)*
Accra Metro Assembly no. 5 (0/1)	0.147 *(3.05)*
Garbage collection (EA average)	0.281 *(2.53)*
Use electricity (EA average)	0.751 *(2.30)*
Has flush toilet (EA average)	-0.405 *(2.40)*
Constant	14.107 *(46.96)*

Urban Coastal	
# of observations	799
# of clusters	40
R^2 (without location means)	0.3924
R^2 (with location means)	0.4179
Variable	Coeff.
# of boys aged 7-14	-0.083 *(2.66)*
# of girls aged 7-14	-0.100 *(2.70)*
Proportion of members that went to school	0.414 *(3.38)*
# of people that went to school	-0.107 *(4.62)*
Other Christian (0/1)	0.171 *(3.36)*
Protestant (0/1)	0.160 *(3.15)*
Head reads English and Ghanaian (0/1)	0.184 *(4.04)*
Use electricity (0/1)	0.189 *(4.00)*
Has flush toilet (0/1)	0.352 *(5.36)*
# of pc weekly hours worked	0.006 *(3.89)*
Eastern region (0/1)	-0.221 *(2.60)*
Central region (0/1)	-0.280 *(4.25)*
Western region (0/1)	-0.241 *(3.38)*
Shama 1 (0/1)	0.391 *(4.05)*
Hours worked (EA average)	0.014 *(3.87)*
Use water from wells (EA average)	0.630 *(4.81)*
Use pipe water (EA average)	0.514 *(5.61)*
Constant	13.164 *(75.80)*

Sources: author's calculation based on GLSS4

Annexe 2: Survey-Based Regression models (continued…)

Urban Forest	
# of observations	960
# of clusters	48
R^2 (without location means)	0.5749
R^2 (with location means)	0.5855
Variable	Coeff.
# of boys aged 7-14	-0.065 *(2.51)*
# of girls aged 7-14	-0.058 *(2.33)*
people that went to school	-0.079 *(6.09)*
Male head (0/1)	-0.104 *(2.77)*
Head age	-0.022 *(3.19)*
Head age squared	0.000 *(2.52)*
Head reads English (0/1)	0.143 *(2.86)*
Head reads English and Ghanaian (0/1)	0.217 *(4.86)*
Catholic (0/1)	0.221 *(3.42)*
Protestant (0/1)	0.086 *(2.48)*
Head is self-employed, non-agro (0/1)	0.291 *(4.54)*
Head does not worked (0/1)	0.169 *(2.25)*
Head is employed (0/1)	0.336 *(4.34)*
pc weekly hours worked in self agro	0.011 *(4.81)*
Use electricity (0/1)	0.245 *(3.64)*
Has flush toilet (0/1)	0.214 *(3.46)*
Use wood for cooking (0/1)	-0.312 *(5.05)*
Post office in EA (0/1)	-0.383 *(6.75)*
Phone in EA (0/1)	0.656 *(8.28)*
Volta Region (0/1)	0.185 *(4.27)*
Western Region (0/1)	0.231 *(3.92)*
Ashanti region (0/1)	0.305 *(8.42)*
Bronga Afaho region (0/1)	0.364 *(8.21)*
Kumasi Metro Assembly 1 (0/1)	0.387 *(4.83)*
Use coal for cooking (EA average)	0.380 *(3.25)*
Use electricity (EA average)	-0.173 *(2.53)*
Constant	14.245 *(68.51)*

Urban Savannah	
# of observations	300
# of clusters	15
R^2 (without location means)	0.5975
R^2 (with location means)	0.5975
Variable	Coeff.
Household size (in log)	-0.478 *(11.35)*
Mole (0/1)	-0.212 *(2.99)*
Islam (0/1)	0.179 *(2.48)*
Thatch roof (0/1)	-0.258 *(5.14)*
No toilet (0/1)	-0.224 *(2.70)*
Use coal for cooking (0/1)	0.154 *(2.55)*
Phone available in EA (0/1)	0.550 *(6.58)*
Upper East region (0/1)	0.149 *(2.63)*
Constant	14.212 *(167.42)*

Sources: author's calculation based on GLSS4

Annexe 2: Survey-Based Regression models (continued…)

Rural Coastal	
# of observations	699
# of clusters	35
R^2 (without location means)	0.5156
R^2 (with location means)	0.5300
Variable	Coeff.
household size (in log)	-0.494
	(9.68)
# of children aged 0-6	0.072
	(3.71)
Ga ethnic group (0/1)	0.272
	(3.91)
Head is unemployed (0/1)	0.437
	(4.92)
pc weekly hours worked	0.008
	(3.63)
Use electricity (0/1)	0.425
	(3.91)
No toilet (0/1)	-0.130
	(2.23)
Junior secondary school (0/1)	0.193
	(2.98)
Central region (0/1)	-0.331
	(4.73)
pschool (EA average)	0.887
	(3.04)
Cement wall (EA average)	-0.586
	(3.66)
Use coal for cooking (EA average)	1.102
	(3.71)
Constant	13.879
	(67.39)

Rural Forest	
# of observations	1680
# of clusters	84
R^2 (without location means)	0.2819
R^2 (with location means)	0.3011
Variable	Coeff.
# of boys aged 7-14	-0.137
	(7.04)
# of female adults aged 15-59	-0.147
	(6.61)
Head reads English (0/1)	0.085
	(2.84)
Head is self-employed, agro (0/1)	-0.167
	(3.49)
pc weekly hours worked - formal sector	0.013
	(4.83)
pc weekly hours worked in self agro	0.011
	(5.21)
Thatch Roof (0/1)	-0.096
	(1.96)
Cement wall (0/1)	0.110
	(2.63)
Use coal for cooking (0/1)	0.276
	(5.56)
Post office in EA (0/1)	0.214
	(3.02)
Western region (0/1)	0.292
	(4.34)
Central region (0/1)	0.397
	(5.18)
Ashanti region (0/1)	0.126
	(1.99)
Head reads English (EA average)	0.406
	(2.29)
No Toilet (EA average)	-0.582
	(5.02)
Constant	13.822
	(127.64)

Sources: author's calculation based on GLSS4

Annexe 2: Survey-Based Regression models (continued…)

Rural Savannah	
# of observations	950
# of clusters	47
R^2 (without location means)	0.2496
R^2 (with location means)	0.4400
Variable	Coeff.
# of girls aged 7-14	-0.116
	(4.02)
# of boys aged 7-14	-0.146
	(6.55)
Head is employed - formal sector (0/1)	0.361
	(3.16)
pc weekly hours worked	0.009
	(3.61)
Upper East Region (0/1)	-0.305
	(4.51)
pc weekly hours worked (EA average)	0.063
	(4.92)
# of rooms (EA average)	-0.191
	(6.41)
Hours worked (EA average)	-0.035
	(3.36)
Constant	14.255
	(70.79)

Sources: author's calculation based on GLSS4

Annexe 3: Regional and District Level Poverty Estimates, by Urban/Rural

District		Total				Urban				Rural			
ID	Name	Population	P0	P1	P2	Population	P0	P1	P2	Population	P0	P1	P2
1	**Western**	**1,919,212**	**0.325**	**0.106**	**0.049**	**692,717**	**0.288**	**0.102**	**0.051**	**1,226,495**	**0.346**	**0.109**	**0.048**
			0.041	*0.018*	*0.010*		*0.039*	*0.019*	*0.012*		*0.041*	*0.017*	*0.009*
101	Jomoro	110,972	0.491	0.176	0.085	32,685	0.412	0.149	0.075	78,287	0.525	0.188	0.090
			0.047	*0.025*	*0.015*		*0.068*	*0.035*	*0.022*		*0.038*	*0.020*	*0.012*
102	Nzema East	142,523	0.446	0.151	0.071	37,716	0.427	0.157	0.079	104,807	0.452	0.149	0.068
			0.046	*0.023*	*0.014*		*0.071*	*0.038*	*0.023*		*0.037*	*0.018*	*0.010*
103	Ahanta West	94,826	0.378	0.126	0.058	18,750	0.297	0.095	0.043	76,076	0.398	0.133	0.062
			0.044	*0.020*	*0.012*		*0.072*	*0.031*	*0.017*		*0.037*	*0.018*	*0.010*
104	Shama-Ahanta E	366,215	0.264	0.090	0.044	366,215	0.264	0.090	0.044	n/a	n/a	n/a	n/a
			0.033	*0.015*	*0.009*		*0.033*	*0.015*	*0.009*				
105	Mpohor-Wassa	122,752	0.292	0.089	0.039	15,664	0.187	0.063	0.030	107,088	0.307	0.093	0.040
			0.045	*0.017*	*0.009*		*0.048*	*0.020*	*0.011*		*0.044*	*0.017*	*0.009*
106	Wassa West	231,952	0.222	0.067	0.030	82,002	0.171	0.056	0.027	149,950	0.250	0.073	0.031
			0.038	*0.014*	*0.007*		*0.036*	*0.014*	*0.008*		*0.039*	*0.014*	*0.007*
107	Wassa Amenefi	234,155	0.324	0.101	0.045	30,996	0.357	0.136	0.072	203,159	0.319	0.096	0.041
			0.044	*0.018*	*0.009*		*0.050*	*0.027*	*0.017*		*0.043*	*0.016*	*0.008*
108	Aowin-Suaman	118,978	0.350	0.113	0.052	18,625	0.323	0.122	0.064	100,353	0.355	0.111	0.049
			0.051	*0.022*	*0.012*		*0.047*	*0.025*	*0.016*		*0.052*	*0.022*	*0.011*
109	Juabeso-Bia	244,456	0.346	0.111	0.051	16,940	0.589	0.261	0.151	227,516	0.328	0.100	0.044
			0.049	*0.021*	*0.011*		*0.087*	*0.058*	*0.041*		*0.046*	*0.018*	*0.009*
110	Sefwi Wiawso	149,247	0.345	0.113	0.053	34,669	0.384	0.150	0.080	114,578	0.333	0.102	0.044
			0.048	*0.021*	*0.012*		*0.055*	*0.030*	*0.020*		*0.046*	*0.018*	*0.009*
111	Bibiani	103,136	0.315	0.102	0.047	38,455	0.278	0.097	0.048	64,681	0.337	0.104	0.046
			0.048	*0.021*	*0.011*		*0.044*	*0.020*	*0.012*		*0.050*	*0.021*	*0.011*
2	**Central**	**1,581,482**	**0.448**	**0.161**	**0.078**	**587,953**	**0.421**	**0.163**	**0.087**	**993,529**	**0.465**	**0.159**	**0.073**
			0.041	*0.020*	*0.012*		*0.039*	*0.020*	*0.013*		*0.042*	*0.019*	*0.011*
201	Komenda	109,940	0.514	0.184	0.087	31,932	0.401	0.145	0.073	78,008	0.561	0.200	0.093
			0.035	*0.019*	*0.012*		*0.040*	*0.019*	*0.012*		*0.033*	*0.019*	*0.012*
202	Cape Coast	114,142	0.273	0.085	0.038	78,358	0.275	0.088	0.041	35,784	0.268	0.079	0.034
			0.039	*0.015*	*0.008*		*0.039*	*0.016*	*0.009*		*0.039*	*0.014*	*0.007*
203	Abura	89,933	0.516	0.181	0.085	26,109	0.495	0.193	0.101	63,824	0.525	0.176	0.078
			0.042	*0.022*	*0.013*		*0.043*	*0.024*	*0.015*		*0.042*	*0.021*	*0.012*

Annexe 3: Regional and District Level Poverty Estimates, by Urban/Rural (continued…)

District		Total				Urban				Rural			
ID	Name	Population	P0	P1	P2	Population	P0	P1	P2	Population	P0	P1	P2
204	Mfantsiman	152,965	0.473	0.168	0.081	76,107	0.424	0.155	0.078	76,858	0.521	0.181	0.083
			0.038	*0.019*	*0.011*		*0.040*	*0.020*	*0.012*		*0.036*	*0.018*	*0.011*
205	Gomoa	191,824	0.630	0.253	0.132	48,326	0.647	0.320	0.199	143,498	0.625	0.230	0.109
			0.038	*0.026*	*0.018*		*0.043*	*0.036*	*0.030*		*0.036*	*0.022*	*0.014*
206	Awutu	169,084	0.526	0.200	0.101	110,593	0.466	0.181	0.096	58,491	0.641	0.234	0.110
			0.042	*0.024*	*0.016*		*0.043*	*0.024*	*0.016*		*0.040*	*0.025*	*0.016*
207	Agona	158,358	0.471	0.168	0.080	102,562	0.363	0.126	0.062	55,796	0.669	0.244	0.114
			0.042	*0.022*	*0.013*		*0.040*	*0.018*	*0.011*		*0.044*	*0.029*	*0.018*
208	Asikuma	89,237	0.576	0.204	0.095	28,364	0.421	0.154	0.078	60,873	0.648	0.227	0.103
			0.053	*0.029*	*0.017*		*0.063*	*0.031*	*0.019*		*0.049*	*0.028*	*0.017*
209	Ajumako	91,976	0.541	0.188	0.087	16,246	0.426	0.153	0.077	75,730	0.566	0.196	0.089
			0.045	*0.023*	*0.013*		*0.045*	*0.022*	*0.013*		*0.045*	*0.023*	*0.013*
210	Assin	195,792	0.290	0.096	0.046	28,388	0.451	0.205	0.123	167,404	0.263	0.078	0.033
			0.056	*0.022*	*0.012*		*0.049*	*0.031*	*0.023*		*0.057*	*0.021*	*0.010*
211	Twifu	110,215	0.289	0.096	0.046	15,126	0.516	0.238	0.143	95,089	0.253	0.073	0.031
			0.057	*0.023*	*0.013*		*0.062*	*0.043*	*0.033*		*0.056*	*0.020*	*0.009*
212	Upper Denkyira	108,016	0.262	0.081	0.037	25,842	0.320	0.120	0.062	82,174	0.244	0.069	0.029
			0.054	*0.021*	*0.011*		*0.055*	*0.027*	*0.017*		*0.054*	*0.019*	*0.009*
3	**Greater Accra**	**2,889,122**	**0.126**	**0.037**	**0.016**	**2,533,079**	**0.099**	**0.028**	**0.012**	**356,043**	**0.316**	**0.101**	**0.045**
			0.018	*0.007*	*0.003*		*0.016*	*0.006*	*0.003*		*0.035*	*0.015*	*0.008*
301	Accra	1,647,202	0.052	0.012	0.004	1,647,202	0.052	0.012	0.004	n/a	n/a	n/a	n/a
			0.009	*0.002*	*0.001*		*0.009*	*0.002*	*0.001*				
302	Ga	549,049	0.237	0.076	0.035	400,960	0.215	0.069	0.033	148,089	0.297	0.094	0.042
			0.031	*0.014*	*0.008*		*0.032*	*0.014*	*0.008*		*0.030*	*0.013*	*0.007*
303	Tema	503,627	0.154	0.044	0.019	445,372	0.153	0.044	0.020	58,255	0.164	0.044	0.018
			0.027	*0.010*	*0.005*		*0.026*	*0.009*	*0.005*		*0.036*	*0.011*	*0.005*
304	Dangbe West	96,309	0.353	0.119	0.055	22,749	0.270	0.088	0.042	73,560	0.378	0.128	0.060
			0.043	*0.020*	*0.011*		*0.036*	*0.016*	*0.009*		*0.045*	*0.021*	*0.012*
305	Dangbe East	92,935	0.387	0.126	0.057	16,796	0.289	0.096	0.046	76,139	0.408	0.133	0.060
			0.058	*0.025*	*0.013*		*0.041*	*0.018*	*0.011*		*0.062*	*0.026*	*0.014*

Annexe 3: Regional and District Level Poverty Estimates, by Urban/Rural (continued…)

District		Total				Urban				Rural			
ID	Name	Population	P0	P1	P2	Population	P0	P1	P2	Population	P0	P1	P2
4	**Volta**	**1,629,523**	**0.495**	**0.185**	**0.093**	**436,925**	**0.431**	**0.170**	**0.090**	**1,192,598**	**0.519**	**0.190**	**0.093**
			0.040	*0.023*	*0.015*		*0.046*	*0.027*	*0.018*		*0.038*	*0.021*	*0.013*
401	South Tongu	64,613	0.493	0.178	0.087	7,213	0.388	0.138	0.068	57,400	0.506	0.183	0.089
			0.040	*0.022*	*0.014*		*0.067*	*0.033*	*0.020*		*0.036*	*0.021*	*0.013*
402	Keta	132,800	0.458	0.164	0.080	70,780	0.427	0.153	0.077	62,020	0.494	0.177	0.084
			0.047	*0.025*	*0.015*		*0.049*	*0.025*	*0.016*		*0.045*	*0.024*	*0.014*
403	Ketu	237,457	0.494	0.181	0.089	82,249	0.394	0.143	0.072	155,208	0.547	0.201	0.098
			0.044	*0.025*	*0.015*		*0.043*	*0.020*	*0.012*		*0.044*	*0.027*	*0.017*
404	Akatsi	93,397	0.538	0.191	0.092	19,528	0.501	0.196	0.104	73,869	0.548	0.190	0.089
			0.043	*0.025*	*0.015*		*0.045*	*0.026*	*0.018*		*0.043*	*0.024*	*0.014*
405	North Tongu	130,106	0.511	0.181	0.087	25,239	0.432	0.161	0.083	104,867	0.530	0.186	0.088
			0.044	*0.024*	*0.015*		*0.053*	*0.028*	*0.017*		*0.042*	*0.024*	*0.014*
406	Ho	233,277	0.443	0.167	0.085	79,514	0.298	0.119	0.066	153,763	0.518	0.192	0.095
			0.043	*0.023*	*0.015*		*0.047*	*0.025*	*0.016*		*0.040*	*0.023*	*0.014*
407	Hohoe	112,198	0.501	0.192	0.099	22,380	0.415	0.184	0.108	89,818	0.523	0.194	0.097
			0.042	*0.025*	*0.017*		*0.047*	*0.032*	*0.024*		*0.040*	*0.024*	*0.015*
408	Kpandu	152,453	0.414	0.143	0.068	34,804	0.300	0.113	0.060	117,649	0.447	0.152	0.071
			0.046	*0.022*	*0.012*		*0.049*	*0.024*	*0.015*		*0.045*	*0.021*	*0.012*
409	Jasikan	111,021	0.534	0.210	0.111	22,054	0.574	0.277	0.171	88,967	0.524	0.194	0.097
			0.042	*0.026*	*0.018*		*0.058*	*0.043*	*0.034*		*0.038*	*0.022*	*0.014*
410	Kadjebi	51,918	0.535	0.202	0.102	8,230	0.329	0.087	0.034	43,688	0.574	0.224	0.115
			0.046	*0.024*	*0.015*		*0.063*	*0.022*	*0.010*		*0.043*	*0.025*	*0.016*
411	Nkwanta	150,588	0.631	0.263	0.140	35,262	0.810	0.368	0.203	115,326	0.576	0.230	0.121
			0.046	*0.034*	*0.024*		*0.079*	*0.073*	*0.054*		*0.036*	*0.022*	*0.015*
412	Krachi	159,695	0.474	0.172	0.085	29,672	0.507	0.181	0.087	130,023	0.467	0.170	0.084
			0.041	*0.021*	*0.012*		*0.058*	*0.031*	*0.019*		*0.037*	*0.018*	*0.011*
5	**Eastern**	**2,103,376**	**0.389**	**0.135**	**0.065**	**724,314**	**0.287**	**0.100**	**0.050**	**1,379,062**	**0.443**	**0.153**	**0.073**
			0.035	*0.017*	*0.010*		*0.034*	*0.015*	*0.009*		*0.036*	*0.018*	*0.010*
501	Birim North	124,016	0.471	0.166	0.079	12,124	0.292	0.102	0.050	111,892	0.490	0.172	0.083
			0.042	*0.021*	*0.013*		*0.073*	*0.033*	*0.020*		*0.039*	*0.020*	*0.012*
502	Birim South	178,920	0.421	0.153	0.076	87,490	0.354	0.136	0.072	91,430	0.485	0.169	0.080
			0.035	*0.019*	*0.012*		*0.033*	*0.018*	*0.012*		*0.038*	*0.020*	*0.012*

Annexe 3: Regional and District Level Poverty Estimates, by Urban/Rural (continued…)

District		Total				Urban				Rural			
ID	Name	Population	P0	P1	P2	Population	P0	P1	P2	Population	P0	P1	P2
503	West Akim	154,107	0.432	0.153	0.074	49,225	0.290	0.104	0.052	104,882	0.498	0.176	0.084
			0.040	*0.021*	*0.012*		*0.037*	*0.018*	*0.011*		*0.042*	*0.022*	*0.013*
504	Kwaebibirem	179,246	0.379	0.132	0.064	69,419	0.319	0.121	0.064	109,827	0.417	0.139	0.064
			0.037	*0.018*	*0.011*		*0.034*	*0.018*	*0.011*		*0.039*	*0.018*	*0.010*
505	Suhum	165,651	0.450	0.161	0.080	35,989	0.343	0.145	0.084	129,662	0.480	0.166	0.078
			0.042	*0.022*	*0.013*		*0.035*	*0.019*	*0.013*		*0.044*	*0.023*	*0.013*
506	East Akim	190,279	0.395	0.138	0.068	70,492	0.396	0.152	0.080	119,787	0.394	0.130	0.060
			0.036	*0.018*	*0.011*		*0.038*	*0.020*	*0.013*		*0.036*	*0.016*	*0.009*
507	Fanteakwa	86,708	0.465	0.166	0.081	15,906	0.443	0.170	0.089	70,802	0.471	0.165	0.079
			0.041	*0.021*	*0.013*		*0.041*	*0.022*	*0.014*		*0.041*	*0.021*	*0.012*
508	New Juaben	135,324	0.158	0.045	0.020	112,647	0.138	0.039	0.017	22,677	0.259	0.077	0.034
			0.026	*0.010*	*0.005*		*0.023*	*0.008*	*0.004*		*0.040*	*0.016*	*0.008*
509	Akwapim South	115,049	0.290	0.088	0.039	52,553	0.181	0.052	0.023	62,496	0.382	0.118	0.052
			0.035	*0.015*	*0.008*		*0.029*	*0.010*	*0.005*		*0.041*	*0.018*	*0.010*
510	Akwapim North	105,538	0.314	0.099	0.045	31,995	0.271	0.087	0.040	73,543	0.333	0.105	0.047
			0.035	*0.016*	*0.009*		*0.035*	*0.016*	*0.009*		*0.035*	*0.015*	*0.009*
511	Yilo Krobo	85,724	0.242	0.068	0.029	15,319	0.190	0.056	0.025	70,405	0.253	0.071	0.029
			0.041	*0.015*	*0.007*		*0.033*	*0.013*	*0.007*		*0.043*	*0.015*	*0.007*
512	Manya Krobo	153,990	0.431	0.147	0.069	61,358	0.291	0.074	0.028	92,632	0.524	0.195	0.097
			0.044	*0.020*	*0.011*		*0.056*	*0.019*	*0.008*		*0.037*	*0.020*	*0.012*
513	Asugyaman	75,523	0.452	0.156	0.074	19,695	0.343	0.091	0.035	55,828	0.491	0.179	0.089
			0.050	*0.022*	*0.012*		*0.066*	*0.023*	*0.010*		*0.045*	*0.022*	*0.013*
514	Afram Plains	135,854	0.462	0.166	0.081	6,885	0.339	0.092	0.036	128,969	0.469	0.170	0.083
			0.041	*0.019*	*0.011*		*0.079*	*0.030*	*0.015*		*0.039*	*0.019*	*0.011*
515	Kwahu South	217,447	0.406	0.145	0.071	83,217	0.314	0.115	0.059	134,230	0.464	0.164	0.079
			0.037	*0.019*	*0.011*		*0.038*	*0.018*	*0.011*		*0.036*	*0.019*	*0.011*
6	**Ashanti**	**3,590,511**	**0.272**	**0.090**	**0.042**	**1,832,441**	**0.141**	**0.047**	**0.023**	**1,758,070**	**0.407**	**0.136**	**0.063**
			0.024	*0.011*	*0.006*		*0.013*	*0.006*	*0.003*		*0.035*	*0.016*	*0.009*
601	Atwima	237,600	0.343	0.112	0.051	49,219	0.170	0.055	0.026	188,381	0.389	0.127	0.058
			0.036	*0.016*	*0.009*		*0.029*	*0.012*	*0.007*		*0.037*	*0.017*	*0.009*
602	Amansie West	108,679	0.437	0.146	0.068	n/a	n/a	n/a	n/a	108,679	0.437	0.146	0.068
			0.039	*0.018*	*0.010*						*0.039*	*0.018*	*0.010*

Annexe 3: Regional and District Level Poverty Estimates, by Urban/Rural (continued…)

District		Total				Urban				Rural			
ID	Name	Population	P0	P1	P2	Population	P0	P1	P2	Population	P0	P1	P2
603	Amansie East	224,830	0.380	0.121	0.055	27,253	0.259	0.086	0.041	197,577	0.397	0.126	0.056
			0.037	*0.016*	*0.009*		*0.030*	*0.014*	*0.008*		*0.038*	*0.017*	*0.009*
604	Adansi West	229,061	0.225	0.071	0.032	136,172	0.122	0.037	0.017	92,889	0.376	0.121	0.055
			0.024	*0.010*	*0.005*		*0.015*	*0.006*	*0.003*		*0.036*	*0.016*	*0.009*
605	Adansi East	129,249	0.446	0.151	0.071	9,616	0.248	0.078	0.036	119,633	0.462	0.157	0.074
			0.038	*0.018*	*0.010*		*0.036*	*0.015*	*0.009*		*0.039*	*0.018*	*0.010*
606	Ashanti Akim S	96,833	0.383	0.123	0.056	15,965	0.337	0.109	0.051	80,868	0.392	0.126	0.057
			0.042	*0.018*	*0.010*		*0.055*	*0.025*	*0.014*		*0.039*	*0.017*	*0.009*
607	Ashanti Akim N	125,817	0.341	0.116	0.056	70,055	0.316	0.114	0.058	55,762	0.371	0.118	0.053
			0.035	*0.016*	*0.009*		*0.028*	*0.014*	*0.009*		*0.043*	*0.018*	*0.009*
608	Ejisu/Juaben	123,761	0.328	0.105	0.048	32,881	0.265	0.088	0.042	90,880	0.351	0.111	0.050
			0.037	*0.016*	*0.008*		*0.033*	*0.015*	*0.009*		*0.038*	*0.016*	*0.008*
609	Bosomtwi	145,918	0.347	0.109	0.048	7,368	0.321	0.095	0.041	138,550	0.348	0.109	0.049
			0.039	*0.016*	*0.008*		*0.053*	*0.023*	*0.013*		*0.039*	*0.016*	*0.008*
610	Kumasi	1,162,408	0.077	0.022	0.009	1,162,408	0.077	0.022	0.009	n/a	n/a	n/a	n/a
			0.011	*0.004*	*0.002*		*0.011*	*0.004*	*0.002*				
611	Afigya/Kwabre	164,454	0.229	0.071	0.031	63,923	0.094	0.028	0.012	100,531	0.316	0.098	0.044
			0.029	*0.011*	*0.006*		*0.016*	*0.006*	*0.003*		*0.036*	*0.014*	*0.007*
612	Afigya Sekyere	118,775	0.403	0.138	0.066	42,041	0.402	0.144	0.072	76,734	0.404	0.135	0.063
			0.042	*0.020*	*0.011*		*0.040*	*0.020*	*0.012*		*0.043*	*0.020*	*0.011*
613	Sekyere East	156,969	0.430	0.153	0.075	52,738	0.343	0.116	0.056	104,231	0.474	0.172	0.084
			0.039	*0.020*	*0.012*		*0.040*	*0.018*	*0.011*		*0.038*	*0.020*	*0.013*
614	Sekyere West	142,126	0.418	0.158	0.081	54,827	0.336	0.139	0.077	87,299	0.469	0.169	0.083
			0.034	*0.020*	*0.013*		*0.025*	*0.017*	*0.013*		*0.040*	*0.021*	*0.013*
615	Ejura/Sekyedu	80,694	0.397	0.152	0.078	39,206	0.201	0.063	0.029	41,488	0.583	0.236	0.125
			0.038	*0.021*	*0.014*		*0.032*	*0.013*	*0.007*		*0.044*	*0.029*	*0.020*
616	Offinso	137,973	0.444	0.163	0.082	42,661	0.463	0.189	0.104	95,312	0.435	0.152	0.073
			0.037	*0.020*	*0.012*		*0.040*	*0.023*	*0.017*		*0.036*	*0.018*	*0.011*
617	Ahafo Ano South	133,508	0.434	0.147	0.069	12,313	0.543	0.236	0.133	121,195	0.423	0.138	0.063
			0.041	*0.021*	*0.013*		*0.065*	*0.051*	*0.038*		*0.039*	*0.018*	*0.010*
618	Ahafo Ano North	71,856	0.385	0.126	0.057	13,795	0.231	0.074	0.034	58,061	0.422	0.138	0.063
			0.042	*0.019*	*0.011*		*0.027*	*0.011*	*0.006*		*0.045*	*0.021*	*0.012*

Annexe 3: Regional and District Level Poverty Estimates, by Urban/Rural (continued…)

District		Total				Urban				Rural			
ID	Name	Population	P0	P1	P2	Population	P0	P1	P2	Population	P0	P1	P2
7	**Brong Ahafo**	**1,812,472**	**0.435**	**0.157**	**0.078**	**676,690**	**0.318**	**0.107**	**0.051**	**1,135,782**	**0.504**	**0.188**	**0.094**
			0.037	***0.019***	***0.012***		***0.040***	***0.019***	***0.012***		***0.036***	***0.019***	***0.012***
701	Asunafo	174,096	0.433	0.149	0.070	49,293	0.257	0.079	0.036	124,803	0.502	0.176	0.084
			0.042	*0.021*	*0.012*		*0.048*	*0.020*	*0.011*		*0.040*	*0.021*	*0.012*
702	Asutifi	83,979	0.457	0.160	0.076	12,903	0.296	0.087	0.038	71,076	0.487	0.173	0.083
			0.039	*0.019*	*0.011*		*0.054*	*0.022*	*0.012*		*0.036*	*0.019*	*0.011*
703	Tanoso	123,084	0.393	0.132	0.062	53,078	0.287	0.087	0.039	70,006	0.472	0.166	0.079
			0.045	*0.021*	*0.012*		*0.049*	*0.020*	*0.011*		*0.042*	*0.022*	*0.013*
704	Sunyani	178,531	0.276	0.091	0.042	131,867	0.203	0.062	0.028	46,664	0.483	0.170	0.082
			0.038	*0.017*	*0.010*		*0.036*	*0.015*	*0.008*		*0.045*	*0.024*	*0.014*
705	Dormaa	150,050	0.426	0.147	0.070	46,785	0.280	0.092	0.044	103,265	0.492	0.172	0.082
			0.041	*0.020*	*0.012*		*0.043*	*0.020*	*0.011*		*0.040*	*0.021*	*0.012*
706	Jaman	147,686	0.629	0.273	0.152	46,725	0.617	0.253	0.133	100,961	0.635	0.283	0.161
			0.038	*0.030*	*0.023*		*0.051*	*0.045*	*0.033*		*0.032*	*0.023*	*0.018*
707	Berekum	93,978	0.332	0.112	0.052	51,723	0.202	0.061	0.027	42,255	0.492	0.174	0.083
			0.041	*0.019*	*0.011*		*0.033*	*0.014*	*0.007*		*0.051*	*0.026*	*0.015*
708	Wenchi	166,354	0.468	0.168	0.082	49,570	0.408	0.130	0.059	116,784	0.494	0.184	0.092
			0.043	*0.021*	*0.012*		*0.055*	*0.023*	*0.013*		*0.038*	*0.020*	*0.012*
709	Techiman	175,170	0.347	0.125	0.061	97,812	0.191	0.064	0.031	77,358	0.544	0.201	0.099
			0.036	*0.019*	*0.011*		*0.029*	*0.013*	*0.008*		*0.045*	*0.026*	*0.016*
710	Nkoranza	128,626	0.515	0.194	0.097	37,398	0.559	0.216	0.109	91,228	0.497	0.184	0.092
			0.043	*0.026*	*0.017*		*0.054*	*0.038*	*0.027*		*0.038*	*0.020*	*0.012*
711	Kintampo	146,206	0.491	0.180	0.089	39,019	0.475	0.171	0.084	107,187	0.497	0.184	0.091
			0.041	*0.022*	*0.014*		*0.046*	*0.027*	*0.018*		*0.039*	*0.021*	*0.012*
712	Atebubu	162,634	0.464	0.166	0.081	53,477	0.427	0.128	0.054	109,157	0.482	0.185	0.095
			0.046	*0.022*	*0.013*		*0.060*	*0.027*	*0.014*		*0.038*	*0.020*	*0.012*
713	Sene	82,078	0.442	0.156	0.075	7,040	0.397	0.116	0.048	75,038	0.446	0.160	0.078
			0.042	*0.019*	*0.011*		*0.081*	*0.031*	*0.015*		*0.038*	*0.018*	*0.010*
8	**Northern**	**1,807,615**	**0.695**	**0.325**	**0.190**	**476,041**	**0.570**	**0.212**	**0.104**	**1,331,574**	**0.740**	**0.366**	**0.220**
			0.034	***0.027***	***0.021***		***0.049***	***0.028***	***0.018***		***0.028***	***0.027***	***0.023***
801	Bole	127,188	0.648	0.285	0.159	15,604	0.440	0.132	0.055	111,584	0.677	0.306	0.174
			0.038	*0.025*	*0.018*		*0.072*	*0.031*	*0.016*		*0.033*	*0.024*	*0.018*

Annexe 3: Regional and District Level Poverty Estimates, by Urban/Rural (continued…)

District		Total				Urban				Rural			
ID	Name	Population	P0	P1	P2	Population	P0	P1	P2	Population	P0	P1	P2
802	West Gonja	138,701	0.572	0.234	0.125	19,898	0.496	0.157	0.068	118,803	0.584	0.247	0.135
			0.041	*0.023*	*0.015*		*0.074*	*0.035*	*0.019*		*0.036*	*0.021*	*0.015*
803	Wast Gonja	174,566	0.551	0.220	0.116	23,881	0.463	0.142	0.061	150,685	0.565	0.232	0.125
			0.040	*0.023*	*0.015*		*0.068*	*0.030*	*0.016*		*0.036*	*0.021*	*0.015*
804	Nanumba	143,866	0.712	0.323	0.184	28,308	0.691	0.277	0.142	115,558	0.717	0.335	0.194
			0.039	*0.030*	*0.023*		*0.056*	*0.040*	*0.028*		*0.035*	*0.028*	*0.021*
805	Sabsugu-Tatale	79,036	0.684	0.298	0.164	16,720	0.760	0.322	0.171	62,316	0.663	0.291	0.162
			0.046	*0.032*	*0.022*		*0.059*	*0.047*	*0.033*		*0.043*	*0.028*	*0.019*
806	Chereponi-Saboba	93,471	0.752	0.351	0.202	6,144	0.866	0.410	0.231	87,327	0.744	0.347	0.199
			0.036	*0.033*	*0.026*		*0.066*	*0.073*	*0.057*		*0.034*	*0.031*	*0.024*
807	Yendi	128,387	0.718	0.333	0.192	43,889	0.629	0.236	0.117	84,498	0.764	0.383	0.231
			0.038	*0.031*	*0.024*		*0.056*	*0.035*	*0.022*		*0.028*	*0.029*	*0.024*
808	Gushiegu-Karaga	121,117	0.857	0.459	0.287	23,545	0.926	0.492	0.297	97,572	0.840	0.452	0.285
			0.037	*0.044*	*0.038*		*0.047*	*0.071*	*0.061*		*0.034*	*0.038*	*0.033*
809	Savelugu-Nanton	90,202	0.672	0.293	0.163	32,574	0.544	0.178	0.078	57,628	0.745	0.357	0.211
			0.045	*0.032*	*0.022*		*0.072*	*0.038*	*0.022*		*0.030*	*0.028*	*0.023*
810	Tamale	292,151	0.565	0.226	0.120	196,126	0.461	0.148	0.065	96,025	0.777	0.385	0.231
			0.048	*0.029*	*0.019*		*0.056*	*0.027*	*0.015*		*0.031*	*0.033*	*0.027*
811	Tolon-Kumbungu	131,791	0.835	0.453	0.289	20,532	0.660	0.238	0.112	111,259	0.868	0.492	0.322
			0.036	*0.042*	*0.037*		*0.076*	*0.047*	*0.029*		*0.029*	*0.041*	*0.039*
812	West Mamprusi	114,220	0.800	0.405	0.246	18,038	0.683	0.290	0.156	96,182	0.822	0.426	0.262
			0.035	*0.037*	*0.031*		*0.053*	*0.041*	*0.030*		*0.032*	*0.037*	*0.032*
813	East Mamprusi	172,919	0.861	0.470	0.299	30,782	0.739	0.314	0.167	142,137	0.888	0.504	0.327
			0.031	*0.041*	*0.037*		*0.050*	*0.042*	*0.031*		*0.027*	*0.041*	*0.038*
9	**Upper East**	**914,016**	**0.715**	**0.337**	**0.197**	**141,885**	**0.511**	**0.182**	**0.088**	**772,131**	**0.752**	**0.365**	**0.217**
			0.035	***0.03***	***0.023***		***0.049***	***0.026***	***0.016***		***0.033***	***0.031***	***0.025***
901	Builsa	75,246	0.575	0.224	0.115	n/a	n/a	n/a	n/a	75,246	0.575	0.224	0.115
			0.039	*0.024*	*0.016*						*0.039*	*0.024*	*0.016*
902	Kassena-Nankani	148,719	0.611	0.246	0.128	23,245	0.532	0.194	0.095	125,474	0.626	0.255	0.134
			0.050	*0.032*	*0.021*		*0.059*	*0.035*	*0.023*		*0.049*	*0.032*	*0.021*
903	Bongo	77,768	0.706	0.299	0.160	n/a	n/a	n/a	n/a	77,768	0.706	0.299	0.160
			0.051	*0.036*	*0.025*						*0.051*	*0.036*	*0.025*

Annexe 3: Regional and District Level Poverty Estimates, by Urban/Rural (continued…)

District		Total				Urban				Rural			
ID	Name	Population	P0	P1	P2	Population	P0	P1	P2	Population	P0	P1	P2
904	Bolgatanga	227,725	0.647	0.276	0.150	48,472	0.395	0.115	0.047	179,253	0.716	0.319	0.178
			0.043	*0.030*	*0.021*		*0.066*	*0.028*	*0.014*		*0.037*	*0.031*	*0.022*
905	Bawku West	80,109	0.832	0.419	0.251	7,747	0.847	0.408	0.234	72,362	0.830	0.420	0.253
			0.038	*0.043*	*0.036*		*0.071*	*0.075*	*0.059*		*0.035*	*0.040*	*0.033*
906	Bawku East	304,449	0.821	0.443	0.282	62,421	0.552	0.201	0.098	242,028	0.891	0.506	0.329
			0.031	*0.037*	*0.033*		*0.051*	*0.028*	*0.017*		*0.026*	*0.039*	*0.037*
10	**Upper West**	**574,918**	**0.758**	**0.385**	**0.236**	**100,458**	**0.379**	**0.112**	**0.047**	**474,460**	**0.839**	**0.443**	**0.276**
			0.058	***0.044***	***0.033***		***0.123***	***0.049***	***0.024***		***0.044***	***0.043***	***0.035***
1001	Wa	223,424	0.677	0.319	0.187	66,364	0.361	0.104	0.043	157,060	0.811	0.410	0.248
			0.071	*0.045*	*0.030*		*0.124*	*0.047*	*0.023*		*0.048*	*0.043*	*0.034*
1002	Nadawili	83,013	0.855	0.452	0.280	n/a	n/a	n/a	n/a	83,013	0.855	0.452	0.280
			0.040	*0.043*	*0.035*						*0.040*	*0.043*	*0.035*
1003	Sissala	84,707	0.801	0.432	0.275	8,839	0.385	0.116	0.050	75,868	0.850	0.469	0.301
			0.054	*0.046*	*0.037*		*0.130*	*0.054*	*0.028*		*0.046*	*0.045*	*0.038*
1004	Jirapa-Lambussie	96,602	0.754	0.377	0.229	13,296	0.402	0.121	0.051	83,306	0.810	0.418	0.257
			0.069	*0.053*	*0.039*		*0.132*	*0.054*	*0.027*		*0.059*	*0.053*	*0.041*
1005	Lawra	87,172	0.836	0.454	0.287	11,959	0.449	0.141	0.061	75,213	0.898	0.504	0.323
			0.046	*0.044*	*0.036*		*0.139*	*0.061*	*0.032*		*0.032*	*0.041*	*0.037*

Sources: author's calculation based on GLSS4 and Census 2000

14.16. Measuring Poverty in South Africa

Statistics South Africa

Measuring Poverty in South Africa

12 December 2005

Statistics South Africa

What Has Been Done

- Four different ways of measuring poverty
- Report

Statistics South Africa

Overview of the Report

- Chapter 2: Combining census and survey data to construct a poverty map of South Africa
- Chapter 3: Key baseline statistics for poverty measurement

Statistics South Africa

Overview of the Report

- Chapter 4: Earnings inequality in South Africa, 1995-1998
- Chapter 5: Income distribution in south Africa: a social accounting matrix approach

Statistics South Africa

Overview of the Report

<u>Combining census and survey data to construct a poverty map of South Africa</u>

- In collaboration with the World Bank
- Series of regression analyses to impute household expenditure
  - Monthly household expenditure
  - Explanatory aspects of poverty

Statistics South Africa

Overview of the Report

- Calculated expected poverty and inequality statistics at 3 geographic levels:
  - Province
  - District Council
  - Magisterial District
- Constructed poverty maps at these geographic levels

Statistics South Africa

Overview of the Report

Key baseline statistics for poverty measurement

– To measure the extent of under-development in different parts of SA

•Household infrastructure index

•Household circumstances index

Statistics South Africa

Overview of the Report

– Data sets
•Census '96
•Imputed expenditure

– Factor analysis – 2 indices

– 1st index – points to meeting of basis needs

– 2nd index – is related to empowerment issues

Statistics South Africa

Overview of the Report

Earnings inequality in South Africa

– Inequality of earned monetary income

– Gini coefficient

– Data sets
•October Household Surveys (OHS)– 1995 to 1998

Statistics South Africa

Overview of the Report

Income distribution in South Africa: a social accounting matrix approach

– Shows the relationship between income generation and consumption at a household level

– Sources of data for compiling the SAM
•Census, IES, OHS

– Makes a comparison with the SAMs previously calculated: 1978, 1988 and 1993

Statistics South Africa

After the Report

• Introduction of a bi-annual Labour Force Survey in 2000
– Re-engineering to move to quarterly surveys

• Income and Expenditure Survey – 2000 and 2005/2006 (currently running)

Statistics South Africa

After the Report

• Census 2001

• October Household Survey replaced by the General Household survey – 2002 (annual)

• Community survey (planned for 2006)

Statistics South Africa

The Demand for Information on Poverty

- Government
  - President's state of the nation address - 2005
  - The Minister of Finance Budget speech – 2005
- The Treasury – official poverty measure
  - To determine whether poverty alleviation measures were working

Statistics South Africa

The Demand for Information on Poverty

- The Presidency - National Income Dynamics Study
  - To determine the effect of government interventions
- Other stakeholders

Simple household poverty assessment models for Malawi

Proxy means tests from the 1997-98 Malawi Integrated Household Survey

by
Ellen Payongayong[1], Todd Benson[1], Akhter Ahmed[1], Shelton Kanyanda[2], Charles Chakanza[2], Peter Mwanza[3], Kondwani Chilopa[2], Nelson Banda[2], and Alick Mphonda[2]

[1] International Food Policy Research Institute, Washington, DC, USA
[2] National Statistical Office, Zomba, Malawi
[3] National Economic Council, Lilongwe, Malawi

September 2002

Table of contents

1. INTRODUCTION

Since 1994, poverty alleviation has been the guiding principle for policy formulation by the government of Malawi. This poverty focus of government policy was reinforced in 2002 with the adoption of the Malawi Poverty Reduction Strategy Paper (PRSP) (Government of Malawi, 2002). In order to develop sound, appropriate policies, a clear understanding of the characteristics of the poor, causes of their poverty, and where they are located is required. With this in mind, in 1996 the Poverty Monitoring System (PMS) was officially launched as an instrument of government for monitoring the progress of poverty oriented policies, programmes, and projects in terms of their impact, effectiveness, and efficiency. In 2000, the PMS, through two of its participating institutions, the National Statistical Office (NSO) and the National Economic Council, with the assistance of the International Food Policy Research Institute, completed the first comprehensive poverty analysis of Malawi based on the 1997-98 Malawi Integrated Household Survey (IHS). Malawi's poor were found to make up over 65 percent of the population (National Economic Council, 2000).

The vision of the PMS is dynamic in that it seeks to capture information as to how the poor and vulnerable groups are faring during times of rapid social and economic change. However, large household surveys, such as the IHS, require considerable resources and time to implement and analyze. Consequently, such surveys are problematic to use for quickly assessing household welfare conditions. Simple and efficient assessment methods are required for the timely monitoring of poverty in Malawi.

This report presents several simple quantitative models of household welfare developed from the 1997-98 IHS data and building on the poverty analysis of the same data set. Such models are often called *proxy means tests*, as they have most commonly been used in targeting social programs on a means-test basis. The models are used to estimate the means, that is, the income level, of households using proxy measures of income. The dependent variable used in the models in this report is a consumption-based household welfare indicator that can be used with basic-needs poverty lines for Malawi to categorize households as poor or nonpoor. The explanatory variables are made up of a handful of household characteristics that are important correlates of welfare status. The principal advantage of these models is that their data requirements are relatively modest, allowing for rapid estimates of the poverty status of study households in Malawi to be made at relatively low cost. (See Skoufias and Coady (2002) for a more detailed discussion on proxy welfare measures.)

There are several immediate applications of these models in Malawi:

- Monitoring of progress towards the poverty reduction benchmarks stipulated in the Malawi PRSP. Poverty and ultra-poverty headcount are key benchmarks, both of which can be estimated using the models with nationally representative data collected by the Malawi Core Welfare Indicators Questionnaire (CWIQ) survey, the principal data collection instrument to be used for annually assessing progress under the PRSP.
- Targeting of the poor for social safety net, social assistance, or other programs that seek to reduce the level of leakage of program benefits to the nonpoor. Such models are most commonly applied in this manner and are used directly in the field for such purposes.
- Assessing the impact of policies on the poor. Socio-economic researchers who are engaged in household level research to formulate or evaluate policies can use the models with the household data they have collected to disaggregate their study populations into poor and nonpoor. Assessments of the different levels of impact that a policy has on the poor and the nonpoor can then be made.

This paper first will provide a brief overview of how proxy means test models are developed. A description is then provided of the data used in developing the Malawi models, before presenting the models themselves. The prediction performance of the models is then examined. In the final two sections of the paper, first the results of a qualitative, wealth ranking exercise in several villages across Malawi are brought into the analysis to suggest additional variables that might be incorporated in future refinements. Secondly, the questionnaire for the up-coming 2002 Malawi CWIQ survey is examined to determine whether the models developed here can be used with the data to be collected.

2. DEVELOPING PROXY MEANS TEST MODELS

The proxy means test approach to estimating the welfare level of households uses household characteristics which are correlated with the consumption level of the household, but which are considerably easier to collect and verify than is comprehensive consumption information for the household. The models are developed through the analysis of data from detailed household surveys with the aim of generating a simplification of the relationships seen in the household survey data between welfare level and household characteristics.

An Ordinary Least Squares (OLS) regression procedure is used to construct the model. The dependent variable in the regression is the welfare indicator of interest. This will usually be total consumption as computed from the information in the household survey. However, for specialized purposes, one could use other measures. For example, measures of nutritional intake or nutritional status could be used in a nutrition or food security application of this method.

Household surveys typically will include modules on demography, education, employment, income sources, consumption, living conditions, and a range of other information on the members of the household. The explanatory variables on the right-hand side of the model can include any of these data as independent variables in the model, as well as any other characteristics of the household available from other sources, e.g., location characteristics. The independent variables used in the model are selected based on several criteria:

- their correlation to the welfare level of households, most importantly,
- diversity in the type of household data used – a range of information which reflects different aspects of household welfare,
- the ease with which the information can be collected from households of interest, and
- the ease with which the information can be verified. This last criterion is of greater concern if the model will be used for program targeting purposes.

Different sets of independent variables are used with the welfare measure to determine which set best explains variation in the welfare measure. While economic theory should guide the choice of candidate variables to include in a model, stepwise regression procedures can be used to further reduce the number of variables in the model by eliminating those candidate variables that do not increase the model's explanatory power.

Finally, after a model is developed, its prediction performance can be determined. Applying the model to the household survey data generates a predicted welfare indicator for each survey household. Households can be classified as poor or nonpoor or be assigned to welfare quintiles based on the predicted value. These estimated classifications are then used in a cross-tabulation with the actual poverty status or welfare quintile of the households based on the actual welfare indicator derived for each household from the consumption data in the survey. Assessments of errors of inclusion (actually nonpoor households classified as poor using the model) and exclusion (actually poor households classified as nonpoor using the model) can be made using poverty status, while a quintile based cross-tabulation will provide more information on the severity of errors of estimation in the model.

Such approaches are becoming increasingly common around the world for poverty monitoring, for the targeting of social programs (Grosh and Baker, 1995; Ahmed and Bouis, 2002), and for the segmentation of markets for advertising and marketing purposes (Haupt, 2001). Moreover, similar models can be derived to assess the aggregate welfare of different groupings of persons other than within the household. Hoddinott and Morris (1999) provide an example of the application of this method using data from a community survey, rather than a household survey, to rank communities in Côte d'Ivoire according to their estimated aggregate well-being.

3. DATA

(a) 1997-98 Malawi Integrated Household Survey

The data source for the construction of the Malawi simple household poverty assessment models is the nationally representative 1997-1998 Malawi Integrated Household Survey. NSO administered the IHS questionnaire over a 12-month period, November 1997 to October 1998. The then 25 administrative districts of Malawi, plus the four major urban centers of the country (Lilongwe, Blantyre, Zomba and Mzuzu), constituted the 29 primary sampling strata.

The questionnaire was administered in two parts. The first was a large questionnaire that was administered to the respondent household in a single visit. This consisted of approximately a dozen modules on household composition, educational attainment, health and nutritional status, agriculture, home-produced and purchased consumption items, assets, and so on. The second part was a diary of expenditure. The diary was maintained over a minimum of 14 days by literate households or through

twice-weekly visits by the enumerator to the survey household to record any expenditure made since the previous visit.

Although administered to 12,960 households, the data set consisted of 10,698 households when the cleaned IHS data was released in early May 2000. However, as the diary of expenditure was not always consistently maintained, upon additional assessment, only 6,586 survey households were judged to have reliable expenditure and consumption information for use in the derivation of the consumption and expenditure based household welfare indicator. As the household welfare indicator is used as the dependent variable in this analysis, the more restricted 6,586 household IHS data set is used here.

(b) The household welfare indicator from the poverty analysis of the IHS

In our models, we employ as the dependent variable the natural log of total daily per capita consumption and expenditure reported by IHS survey households. This variable (unlogged) was used as the measure of welfare for the poverty analysis of the IHS and was made up of four components:

- total food consumption, whether purchased or from home consumption;
- total non-food non-durable goods expenditure, including gifts to others outside the household;
- estimated use-value of durable consumer goods; and
- rental value of housing for the household, actual or imputed

The welfare indicator is expressed in real terms with the value of nominal consumption normalized to an April 1998 base.[1] (National Economic Council, 2000, pp. 96-97).

We used the natural logarithm of the welfare indicator as the preferred dependent variable in deriving our models to improve the predictive power of the models. The distribution of the logged welfare indicator across the IHS sample households better approximates a normal distribution than does that of the simple, unlogged welfare indicator. Consequently, when using the logged welfare indicator in the OLS regression, gains in predictive power are achieved and significantly better estimates result. However, as most individuals do not commonly and comfortably use logarithms in their every day work, the direct utility of the models that could be achieved through using an unlogged welfare indicator is sacrificed for this improved predictive power. Particularly if used for targeting purposes, simple modifications to the logged models could be made, however, to allow users to employ them without regard to the underlying logarithm. Nevertheless, models using the unlogged dependent variable are presented in this paper as alternatives to the preferred models.

The proxy means test estimates of household welfare will be used with the poverty lines of Malawi to classify households as poor or nonpoor. The poverty line is that level of welfare which distinguishes poor households from nonpoor households, and is expressed in the same units as the household welfare indicator. Cost-of-basic-needs poverty lines for four areas of the country – Southern rural, Central rural, Northern rural, and Urban – were developed in the poverty analysis of the IHS. The objective core of the poverty line is the per capita recommended daily calorie requirement for the IHS households as established by nutrition researchers. The calorie requirements were used to establish the food component of the poverty line by determining what it costs for a poorer household in Malawi to acquire sufficient calories to meet their requirements. The total poverty line is computed by adding to this food poverty line a non-food component that is derived from the daily non-food consumption of those IHS households the value of whose total consumption and expenditure is in the neighborhood of the value of the food poverty line. The weighted mean poverty line for Malawi as a whole in March 1998 Kwacha is MK10.47 per person per day, or US $0.41 (*ibid.* pp. 95-110).

4. PREDICTION OF HOUSEHOLD WELFARE

(a) IHS variables for proxy means test

Using the criteria for independent variables stipulated above, 72 candidate household-level independent variables were extracted from the IHS data set for further analysis in constructing the proxy means test models for Malawi. These variables can be categorized by type as reflecting household demography, employment, education, agriculture, living conditions, assets, and income and consumption. Both continuous and categorical variables are used. In addition, 28 location dummy variables were also constructed based on districts, urban centers, and rural regions. All of the candidate independent variables are shown in Table 1.

The choice of independent variables is, of course, constrained by what is available in the 1997-98 Malawi IHS. Several variables that one might consider to be important proxies for household welfare and poverty status could not be included in our model construction due to their absence from the data set or to data quality considerations. Among these variables are various housing quality characteristics (wall and roof type, sanitation facilities), asset ownership (radio, sofa), and whether the household makes use of certain purchased items (toothpaste, hand soap, shoes). Such alternative variables to

[1] April 1998 is the middle month of the IHS survey period. Welfare indicators were temporally normalized to this month using separate monthly consumer price indices (CPI) for the three rural regions and four urban centers of Malawi. These CPIs were calculated by the NSO. In April 1998, US $1.00 = MK 25.40.

The four poverty lines were used to spatially normalize the welfare indicator. As each poverty line represents the cost of a comparable basket of basic goods in a poverty line area, a set of spatial price normalization indices for the four poverty line areas were derived to normalize the welfare indicators for spatial price differences.

those available for the models here are considered later in this report. If future household surveys include these sorts of data, the proxy means models presented here should be updated in the future.

(b) Model derivation

As noted, an OLS regression was used to develop the models based on a sub-set of the candidate variables. Stepwise regression procedures and judicious selection of independent variables through an iterative process were used to reduce the number of explanatory variables to a manageable number of variables with statistically significant coefficients. The final regressions to develop the models were run using the *svyreg* command in the Stata™ statistical software program to take the IHS sample design into account in computing the error terms of the regression coefficients.

The regressions were governed by several design considerations:

(i) Dependent variable – logarithm or unlogged

Models were developed and are presented using as the dependent variable both the natural logarithm and the unlogged value of the household welfare indicator from the poverty analysis of the 1997-98 Malawi IHS. As noted above, the distribution of the logged welfare indicator more closely approximates a linear distribution. Regressions based on this dependent variable have significantly higher explanatory power than those based on the unlogged welfare indicator. Consequently, the proxy means test models based on the logged welfare indicator are the preferred models.

However, in some contexts it will be difficult for users to make use of the preferred models because of the logarithm. Consequently, models based on the same explanatory variables, but using the unlogged household welfare indicator, are also presented below.

(ii) Separate models for rural and urban households.

The principal correlates of household welfare found in the 1997-98 Malawi IHS differ significantly between rural and urban households. This is demonstrated quite consistently in the poverty profile of the 1997-98 Malawi IHS (National Economic Council, 2000). Consequently, separate rural and urban models are presented in all cases.[2]

Efforts were made initially to develop universal Malawi proxy means test models, with relatively good results in terms of explanatory power. Using all of the candidate independent variables with district dummy variables in a stepwise regression, models with R^2s of 0.42 using the unlogged dependent variable and 0.46 with the logged welfare indicator were obtained. However, the models consisted of a large number of household variables (more than 35), many of which would be relevant only in rural or in urban areas, respectively. As this would make data collection problematic, it was decided to pursue separate rural and urban models.

(iii) District or regional location dummy variables in rural models.

Location dummy variables are presented in all models. In the urban models, these consist of the urban centers. However, for the rural areas, two separate models are presented based on whether district location dummies or regional dummies are employed. The preferred model is that using the district dummies.

However, the IHS household sample size in some districts is quite small. In these districts, the large error terms for the district dummy coefficient may call into question the validity of the coefficient.[3] There may be instances where researchers working in a subset of these and other districts would want to assess whether the generalizations they make are more reasonable when the model based on the regional dummy variable is used, rather than the district dummies. Consequently, rural models using the same household characteristics but based on regional dummies are presented below. However, it is important to note that the R^2 values for these models are considerably lower than for comparable models that use district dummies.

(c) Models

The preferred rural and urban models consist of those that employ the logarithm of the household welfare indicator as the dependent variable and district or urban center location dummy variables. These models are presented in Table 2. The R^2s provide an indication of the explanatory power of the models: rural – 0.37; urban – 0.60. Although the models leave unexplained a considerable portion of the welfare level of these households, these R^2s are comparable to those of proxy means test models developed elsewhere. For example, the best national model for Jamaica computed by Grosh and Baker (1995) had an R^2 of 0.41,while their urban and rural models both had R^2s of 0.36. The national model for Egypt of Ahmed and Bouis (2002) had an R^2 of 0.43.

(i) Rural preferred model

There are 14 variables in the rural model, excluding the district dummy variables. These include variables on household composition, education, employment, agriculture, purchase of a specific item, and asset ownership – a relatively diverse set of characteristics is used to estimate household welfare. The signs of the coefficients for each are reasonable and consistent with economic theory and with the

[2] The urban population is defined as those resident in the four urban centers of Malawi. Households resident in rural urban centers are grouped with other rural households to develop the rural models.

[3] District dummy variables are retained in the models regardless of whether or not the coefficients are significantly different from zero.

patterns revealed in the absolute and relative (quintile) poverty profiles of the IHS (National Economic Council, 2000, Poverty Monitoring System, 2001).

Descriptive statistics for these variables for the rural IHS households are presented in Table 3. Most of the 14 variables are self explanatory in their construction. Three deserve comment:

- 'sugar' - This variable is a dummy variable with a value of 1 if the IHS household reported purchasing sugar in their diary of expenditure. The diary of expenditure was kept for each household for a minimum of 14 days. This variable could be restated as "purchased sugar in the past two weeks."
- 'educhh' – The education level of household head variable is a categorical variable on the level of educational attainment: 1 – Standards 1 to 4; 2 – Standards 5 to 8; 3 – Forms 1 or 2; 4 – Forms 3 or 4; 5 – University or higher. Data is missing on this variable for just over 100 households, the only model variable with any missing data.
- 'tacres' – Total acreage cultivated was estimated by the household respondent. Field measurements were not done. The area unit is acres, not hectares.

Regarding their ease of collection and validation, the variables are varied in this regard. None of the variables are independently observable with certainty. All will likely require a household member to provide the information required. However, the accuracy with which the data is reported should be relatively good. The only variable for which the respondent in reporting might unintentionally err is total acreage cultivated, due to poor estimation of land area.

However, as the information cannot be corroborated through direct observation, validating the information will be more difficult. A possible scenario by which the information could be checked would be though the village headman. He or she would probably be aware of the condition of the household concerning household size, education level of the household head, cropping of tobacco, and ownership of bicycle, bed, and motor vehicle. While the headman could likely make reasonable assessments for most of the other variables, the fact that these variables are somewhat problematic to validate does make this model less useful for targeting of social programs than one consisting of easily verified characteristics.

(ii) Urban preferred model

There are nine household characteristic variables in the urban model. These include variables on household composition, education, employment, amenities, and asset ownership. Descriptive statistics for these variables for the urban IHS households are presented in Table 3, and their construction is either self-explanatory or has been explained earlier. This set is less diverse than that for the rural model, but much of the difference is accounted for by the lower importance of agriculture for household welfare in the urban centers. The signs of the coefficients on the variables in the urban model are consistent with expectations.

Much the same criticisms of the model variables with regard to ease of collection and validation made on the preferred rural model apply to the urban model here. Given the greater level of anonymity in urban centers, validating the household characteristics using traditional authorities will be more difficult with this model than for the rural model, in spite of the fewer variables involved.

One potentially significant difference between the rural and urban models is that the rural model includes the variable on the purchase of sugar, a variable whose level is expected to change quite quickly with changes in household welfare. This point is raised with regard to the potential use of the proxy means test models for welfare monitoring in Malawi over relatively short-time periods. The sugar variable is the only one in both models that might be expected to vary quite rapidly with welfare changes. The asset variables would also change with sustained changes in welfare status, but not so quickly as would sugar consumption. Particularly for the urban model, if one seeks to monitor month-on-month changes in household welfare, for example, models consisting of a wider range of considerably more volatile household characteristics would be more useful than those presented here.[4] We expect that these models will be most useful in monitoring household welfare over a time horizon of one year or longer.

(iii) Alternative models

Four alternative models are presented in Table 4: a rural model using district dummy variables and an unlogged dependent variable, two rural models that use regional dummy variables and logged and unlogged dependent variables, respectively, and an urban model which has the unlogged welfare indicator as dependent variable. Aside from the regional dummy variables in two of the rural models, these models use the same explanatory variables as the preferred rural and urban models presented in Table 2.

The justification for the inclusion of these models was presented earlier. As expected, the R^2s for all of the models in Table 4 are lower than those of the preferred models. However there may be situations in which the models with the unlogged dependent variables will be of use. The rural models with the regional dummy variables are likely to be of more limited value. The coefficients on both dummy variables in both models are not statistically significant.

---

[4] Our thanks to Dr. N. Minot for this observation.

Several variables similar to the sugar variable were investigated for the urban model, but none proved significant. Because of the higher use of the market for the consumption requirements of urban households, both poor and non-poor, variables on commonly purchased items are less likely to be correlated with welfare than in rural areas. However, a more diverse range of items should be considered when the models are updated in the future.

5. ASSESSING THE PREDICTIVE PERFORMANCE OF THE MODELS

Some idea of the explanatory power of the models can be gauged by the R^2s for each. However, as the models are being constructed to predict the welfare status of households, applying the models back on the IHS data set from which they were derived can provide a better assessment of their performance. We determine what proportion of the IHS households whose reported consumption classifies them as poor (nonpoor) are also categorized as poor (nonpoor) by the proxy means test models. Two sorts of predictive errors are possible. An *error of omission* occurs when a poor household is classified as nonpoor. In contrast, an *error of inclusion* occurs when a nonpoor household is classified by the model as poor. The level of these errors provides an indication of how reliable the models are. It is usually the case that altering the model to reduce one sort of error will increase the other.

We extend the analysis by considering the ultra-poor as well. The ultra-poor are those households whose level of consumption is less than 60 percent of the basic needs poverty line. The poverty analysis of the IHS showed 28 percent of Malawi's population to be ultra-poor. Considering the model performance in identifying the ultra-poor provides an indication of whether the models are useful in identifying the poorest of the poor.

(a) Errors of inclusion and exclusion

Table 5 presents the poverty and ultra-poverty errors of inclusion and exclusion for the two preferred models presented in Table 2 and for the four alternate models described in Table 4. Arguably, the two preferred models give the best results of all of the models presented for both predicting the poor and the ultra-poor. When considering the poor, the rural model has an error of inclusion of about 35 percent and an error of exclusion of about 25 percent. The poverty status of about 29 percent of all rural IHS households is misclassified. For the urban model, these values are 18, 23, and 20 percent, respectively. The alternative models with the unlogged dependent variable have somewhat lower errors of inclusion than do the preferred models, but these are at the expense of significantly higher errors of inclusion. In all cases, the overall proportion of households that are misclassified using the alternative models are higher than for the preferred models.[5]

Turning to the performance of the models in identifying the ultra-poor, the results are less encouraging. The rural model has a low error of inclusion of about 6 percent. However, this low value is linked to the fact that only 24 percent of the IHS rural households are classified as ultra-poor. In contrast, the error of exclusion for the ultra-poor using the preferred rural model is about 67 percent – if this model were used to identify the ultra-poor for a targeted program, many of the ultra-poor would be excluded. Overall, the ultra-poverty status of about 21 percent of all rural IHS households is misclassified. For the urban model, these values are 8, 44, and 17 percent, respectively. Both preferred models do not do a good job in identifying the poorest, predicting their welfare level to be higher than it actually is. However, all the alternative models provide no advantages over the preferred models, as both their errors of exclusion and inclusion are higher than those for the preferred models.

[5] As the proportion of the population considered to be poor vary between countries, the value of comparisons between countries of proxy means test model performance is difficult to judge. Nevertheless, the preferred model of Grosh and Baker (1995) for Jamaica had an error of inclusion of 34.2 and an error of exclusion of 41.0 for a population with a poverty headcount of 30.0 percent. For Egypt, Ahmed and Bouis (2002) report errors of 16.3 and 28.2 percent, respectively, for a population with a poverty headcount of 36.5 percent.

(b) Performance on the basis of welfare quintiles

An alternative way of examining the predictive performance of the models is by categorizing households into welfare quintiles based on their actual welfare rank among all households and on their welfare rank as predicted by the models and then constructing a cross-tabulation of the households by these quintiles. This is done for all six models in Table 6. These tables allow a more flexible assessment to be made of the models, as one is not constrained to use the poverty and ultra-poverty lines alone in judging their performance.

The patterns are as expected. The diagonal cells, indicating the percentage of households in the welfare quintile that the model correctly predicted, usually have the highest values. Relatively few predicted welfare levels are strong outliers. For the rural models, about 75 percent of all households in a welfare quintile are predicted to have a welfare level within the same quintile or one quintile lower or higher. For urban households, the welfare levels of 85 percent of the households in a quintile are predicted to within one quintile of their true quintile.

6. IDENTIFYING ALTERNATIVE PROXY WELFARE INDICATORS

(a) Description of fieldwork

In October and November 2001, qualitative fieldwork was carried out in communities around Malawi to assist in selecting variables from the 1997-98 Malawi IHS for use in the proxy means test and to determine whether there are some variables not in the IHS which are important correlates of poverty. The methods used to identify these variables were qualitative wealth ranking and focus group discussions. The qualitative wealth ranking involved the use of reasonably well-informed individuals in the community sorting the households in the community into categories based on their perceived welfare level. Following the sorting procedure, the informant was asked to describe the salient characteristics of the households in each welfare group (Reitbergen-McCracken & Narayan, 1998, pp 154-156). If information on them could be found in the IHS, the characteristics noted were used as candidate variables for the proxy means test models. The wealth ranking method was only used in rural communities.

The second method was simple focus group discussions where welfare categories of households within the communities were discussed in a relatively abstract manner. The living conditions of specific households in the community were not considered. However, after identifying the range of welfare categories into which households in the community could be grouped, the characteristics of households categorized in each welfare group were identified.

The research involved working with households and individuals in 11 villages in 11 districts in all three regions and in four neighborhoods in Mzuzu and Lilongwe cities.[6] The research sites were located in seven different agroecological zones.

(b) Proxy household welfare variables

A wide range of variables were collected as being appropriate to use in distinguishing the rich from the poor or the poorest from the better-off. Results from three rural communities and two urban neighborhoods are presented in Table 7 as an example of the sort of information gathered.

A compilation was made of all of the proxy welfare indicators collected in the fieldwork. This is shown in Table 9. Thereafter, a cross-tabulation of the compiled variables that are not in the 1997-98 Malawi IHS was done against the welfare quintile status of households from several rural communities as determined through the qualitative wealth ranking exercise. This was used to identify those variables that are quite strongly correlated with high or low welfare status and would be worth considering as proxy welfare indicators in the future. Two examples of such cross tabulations are presented in Table 8. Those variables that are not in the IHS, but are good proxy indicators are identified in Table 9.

7. APPLICABILITY OF THE MALAWI 2002 CWIQ SURVEY DATA TO THE MODELS

The Malawi Poverty Reduction Strategy Paper includes a poverty monitoring and evaluation plan (Government of Malawi, 2002, pp. 110ff.). The Core Welfare Indicators Questionnaire survey is specified as the principal instrument for short term (annual) monitoring of the outcomes related to the PRSP, including changes in the aggregate poverty status of Malawi's population. Consequently, it is of importance to the government of Malawi that household welfare levels can be estimated from the information contained in the CWIQ survey.

The first CWIQ survey is to be conducted starting in September 2002. We received one of the later drafts of the questionnaire as the survey was being prepared to go to the field (25 August draft). This was reviewed to determine whether the household variables in the rural and urban proxy means test models presented in this report could be implemented using the data to be gathered by the CWIQ. The results of this review of the questionnaire are presented in Table 10.

All of the variables are included in the 2002 Malawi CWIQ. Two of the variables in the models differ slightly from the variables collected in the CWIQ: The model uses a variable on sugar purchases, while the CWIQ asks about sugar use in the past two weeks. In addition, the land variables asked in the CWIQ are on ownership rather than on cultivation. Land cultivated was the variable used in constructing the rural proxy means test model. However, these are relatively minor problems, and should cause no hesitation in applying the CWIQ data to the models when it becomes available.

In summary, the authors of the CWIQ questionnaire have been quite comprehensive in including both the variables that are included in the proxy means test models and those variables that were identified by the fieldwork as important proxy welfare indicators. Barring implementation difficulties, when coupled with the proxy means test models, the CWIQ survey of 2002 should prove to be a useful data set for poverty assessment and, with later rounds of the same survey, for assessing progress towards the goal of poverty eradication in Malawi.

[6] The districts in which the fieldwork was carried out were Chikwawa, Mulanje, Phalombe, Machinga, Mangochi, Ntcheu, Ntchisi, Kasungu, Nkhotakota, Nkhata Bay, and Mzimba.

8. REFERENCES

Ahmed, Akhter U., and Howarth E. Bouis. 2002. *Weighing what's practical: Proxy means tests for targeting food subsidies in Egypt*. Food Consumption and Nutrition Division Discussion Paper no. 132. International Food Policy Research Institute, Washington, DC, USA.

Government of Malawi. 2002. *Malawi Poverty Reduction Strategy Paper*. final draft, mimeo. Ministry of Finance and Economic Planning, Lilongwe.

Grosh, Margaret E., and Judy L. Baker. 1995. *Proxy means tests for targeting social programs – Simulations and speculation*. Living Standards Measurement Study Working Paper no. 118. The World Bank, Washington, DC, USA.

Haupt, Paul. 2001. *The SAARF Universal Living Standards Measure.* South African Advertising Research Foundation, Pretoria, South Africa.

Hoddinott, John, and Saul S. Morris. 1999. *Pipeline Study: Côte d'Ivoire*. IFPRI-IFAD Technical Assistance Grant – Strengthening Household Food Security and Nutritional Aspects of IFAD Poverty Alleviation Projects: Developing Operational Methodologies for Project Design and Monitoring, Research Paper #3. International Food Policy Research Institute, Washington, DC, USA.

National Economic Council. 2000. *Profile of poverty in Malawi, 1998: Poverty analysis of the Malawi Integrated Household Survey, 1997-98*. mimeo.

Poverty Monitoring System, Government of Malawi. 2001. *A relative profile of poverty in Malawi, 1998: A quintile-based poverty analysis of the Malawi Integrated Household Survey, 1997-98*. mimeo.

Reitbergen-McCracken, J., and D. Narayan. 1998. Participation and Social Assessment – Tools and Techniques. The World Bank, Washington, DC, USA.

Skoufias, Emmanuel, and David Coady. 2002. *Are the welfare losses from imperfect targeting important?* Food Consumption and Nutrition Division Discussion Paper no. 125. International Food Policy Research Institute, Washington, DC, USA.

9. TABLES

Table 1: Candidate independent variables from 1997-98 Malawi IHS for proxy means test models.

Variable	Definition	Variable	Definition
Demography		*Assets*	
hhsize	Household size	bed	HH owns a bed
hhszsq	Household size squared	fridge	HH owns a fridge
agehhh	Age of head of household	bicycle	HH owns a bicycle
agehhh2	Squared age of head o f household	carmbike	HH owns a car or motor cycle
femhhh	Female headed household	canoe	HH owns a canoe or boat
married	Head of household married	plough	HH owns a plough
femmarr	Married female head	radio	HH owns a radio
ltage5	Number in hh less than age 5	eleccost	Mean monthly electricity bill
depndnt	Number of dependents (age 4 5 or 6 5)	expas set	HH owns expensive asset
dpndrat	Dependency ratio: dependents/hhsize	ownhouse	HH in owned house (not rented)
dpndrat2	Squared dependenc y ratio	*Income and consumption*	
Employment		trnsrec	HH received income transfer
employer	Head of household employer	gift	HH gave income transfer
selfempl	Head of household self employed	instcred	Credit from institutional source in past year
unemploy	Head of household unemployed	salt	Purchased salt -reported IHS diary expend.
employee	Head of household employee	sugar	Purchased sugar -IHS diary
forfish	HH engaged in forestry/fishing	soap	Purchased soap -IHS diary
manufact	HH engaged in manufacturing	pjelly	Purchased petroleum jelly -IHS diar y
service	HH engaged in service provision	*Location*	
prof	HH member w/ prof, admin, or cler ical occup.	dist1	Nsanje district
farmer	HH member with agricultural occupation	dist2	Chikwawa district
servind	HH member in service industry	dist3	Mwanza district
govtemp	HH member employed by government	dist4	Blantyre Rural district
hhhseekw	HH head seeking work	dist5	Blantyre City
numempl	No.of salaried HH members	dist6	Zomba Rural district
Education		dist7	Zomba Municipality
educhh	Education level of household head	dist8	Thyolo district
maxed	Max educational level of any HH adult	dist9	Mulanje district
hhhlit	Literate head of household	dist10	Phalombe district
secndry	Child in secondary school	dist11	Machinga district
pvtsch	Child in private school	dist12	Mangochi district
misssch	Child in mission school	dist13	Chiradzulu district
edcost	Total educational costs over 12 mo.for HH	dist14	Ntcheu district
Agriculture		dist15	Dedza district
pcland	Per capita acreage cultivated	dist16	Salima district
tacres	Total acreage cultivated	dist17	Mchinji district
hybmaize	HH grows hybrid maize	dist18	Ksungu district
cassava	HH grows cassava	dist21	Dowa district
milsorg	HH grows millet or sorghum	dist22	Nkhotakota district
cotton	HH grows cotton	dsit30	Mzimba district
tobacco	HH grows tobacco	dist30	Mzuzu City
fertfdcr	HH used fertilizer on food crop	dsit31	Nkhata-Bay district
fertcscr	HH used fertilizer on cash crop	dist31	Rumphi district
vinpfdcr	Total value of inputs used by HH on food crops	dsit32	Kronga district
vinpcscr	Total value of inputs used by HH on cash crops	dist32	Chitipa district
ncattle	No.cattle owned	rururb2	rural Central region
ngoats	No.goats owned	rururb3	rural Northern region
nsheep	No.sheep owned	*Lilongwe Rural district, Lilongwe City, and rural Southern region are left unspecified to avoid collinearity.*	
npigs	No.pigs owned		
nchickns	No.chickens owne d		
lvstock	No.of livestock pe r adult equivalent		
Living conditions			
Houseown	HH owns house in which it lives		
riverhh2	HH gets water from riv er or lake		
Owntap	HH gets water fro m own tap		
Colfirwd	HH cooks over collected firewood		
Liteelec	HH get lighting from electricity or gas		
Facaccss	Mean time (hr) to health centre, bus stage, ADMARC, bank, & PO		
Immuniz	Mean proportion of 8 mmunizations kids in hh receive d		

Table 2: Preferred Malawi proxy means test models

Preferred rural Malawi proxy means test model

*Dependent variable:*ln HH welfare indicator

	coef.	t-statistic
HH cooks over collected firewood	-07	(358)*
Household size	-8	(1912)*
Household size squared	05	(1193)*
Education level of household head	6	(6. 65)*
No.of salaried HH members	98	(411)*
HH owns a bicycle	053	(693)*
HH owns a car or motor cycle	693	(86)*
HH owns a fridge	691	(569)*
Purchased sugar-diary	052	(6)*
Total acreage cultivated	9	(67)*
HH grows tobacco	06	(43)*
HH owns a bed	93	(118)*
No.cattle owned	03	(310)*
HH grows hybrid maize	6	(38)*
Nsanje	07	-19
Chikwawa	061	-12
Mwanza	9	-09
antyre Rural	9	-153
mba Rural	068	-110
Thyolo	-96	(18)*
Mulanje	-9	-0
Phalombe	-06	(9)*
Machinga	69	-01
Mangochi	-05	-04
Chiradzulu	-08	-05
Ntcheu	-06	(7)*
Dedza	00	-4
Salima	041	-19
Mchinji	-62	-01
Kasungu	08	-0
Dowa	038	-62
Nkhotakota	09	-00
Mzimba	-7	-64
Nkhata-Bay	02	-92
Rumphi	-9	-2
Karonga	93	(194)*
Chitipa	-046	(9)*
Constant	8	(18)*
Observations	*5539*	
R-squared	*0.37*	

Robust t statistics in parentheses
** significant at 10%; ** significant at 5%; *** significant at 1%*

Preferred urban Malawi proxy means test model

*Dependent variable:*ln HH welfare indicator

	coef.	t-statistic
HH owns a fridg e	618	(62)*
Household siz e	-06	(96)*
Household siz e squared	06	(52)*
Age of head of household	6	(2)*
Education leve l of household head	051	(4)* *
No.of salaried H H members	61	(18)*
HH owns a car o r motor cycle	4	(630)* *
HH get light ing from electricity or gas	9	(50)*
HH owns a bed	97	(411)*
antyre City	-07	-02
mba Municipal ity	-047	(345)*
Mzuzu City	-9	-166
Constant	247	(168)*
Observations	*872*	
R-squared	*0.60*	

Table 3: Descriptive statistics for IHS household variables in Malawi proxy means test models.

Rural model

Variable	Description	bs	Mean	Std.Dev.	Max	Min	Mean poorest quintile	Mean wealthiest. quintile
hhsize	Household size	5,644	447	20 18		1	5314	3510
hhszsq	Household size squared	5,644	27 2554	32		1	326	10
educhh	Education level of household head	5,539	139	10 9	5	0	98	10
numempl	No.of salaried HH member s	5,644	21	092	6	0	047	02
bicycle	HH owns a bicycle	5,644	047	06 1		0	2	050
carmbike	HH owns a car or motor cycle	5,644	0	0 95	1	0	0	0 3
fridge	HH owns a fridge	5,644	6	7 1		0	0	2
tacres	Total acreage cultivated	5,644	242	2 35		0	18	22
tobacco	HH grows tobacco	5,644	08	390	1	0	048	9
bed	HH owns a bed	5,644	25	056 1		0	07	046
sugar	Purchased sugar-diary	5,644	66	068 1		0	69	6
colfirwd	HH cooks over collected firewood	5,644	65	04 1	1	0	92	7
ncattle	No.cattle owned	5,644	046	26 48		0	034	62
hybmaize	HH grows hybrid maize	5,644	042	04	1	0	20	07

Urban model

Variable	Description	bs	Mean	Std.Dev.	Max	Min	Mean poorest quintile	Mean wealthiest. quintile
hhsize	Household size	92	436	38 16		1	532	3697
hhszsq	Household size squared	92	235	652 26		1	34495	1919
agehhh	Age of head of household	92	36199	11190	8 7	18	387	3697
educhh	Education level of household head	8	9 130	5	5	0	161	382
numempl	No.of salaried HH member s	92	61	71 5		0	640	16
carmbike	HH owns a car or motor cycle	92	0	8 1		0	0	27
fridge	HH owns a fridge	92	09	38 1		0	01	68
bed	HH owns a bed	92	6	09 1		0	697	98
liteelec	HH gets lighting from electricity or gas	92	066	02	1	0	6	05

Table 4: Other Malawi proxy means test models

Alternate rural Malawi model – 1		
Dependent. variable:HH welfare indicator (unlogged)		
	coef.	t-statistic
HH cooks over collected firewood	-358	(3[illegible]*
Household size	-364	(18)*
Household size squared	091	(8[illegible]*
Education level of household head	1319	(649)*
No.of salaried HH members	1615	(37)*
HH owns a bicycle	157	(460)*
HH owns a car or motor cycle	1372	(548)*
HH owns a fridge	14513	(38)*
Purchased sugar-diary	176	(615)*
Total acreage cultivated	057	(512)*
HH grows tobacco	12	(333)*
HH owns a bed	361	(952)*
No.cattle owned	091	(12)*
HH grows hybrid maize	068	(16)*
Nsanje	622	(23)*
Chikwawa	8	(28)*
Mwanza	010	-64
Blantyre Rural	228	-159
Zomba Rural	365	-158
Thyolo	-920	(9)*
Mulanje	617	-[illegible]03
Phalombe	-112	-150
Machinga	958	-[illegible]9
Mangochi	10	-[illegible]03
Chiradzulu	-637	-[illegible]03
Ntcheu	-1624	-146
Dedza	-[illegible]6	-[illegible]7
Salima	267	(2)*
Mchinji	-[illegible]8	-[illegible]03
Ksungu	168	-[illegible]2
Dowa	19	-[illegible]07
Nkhotakota	-[illegible]033	-[illegible]02
Mzimba	-1416	-131
Nkhata-Bay	183	-158
Rumphi	-[illegible]37	-[illegible]05
Kronga	462	(48)*
Chitipa	-5499	(59)*
Constant	19915	(1191)*
Observations	*5539*	
R-squared	*0.27*	

Alternate rural Malawi model – 2 (regional dummy variables)		
Dependent variable:ln HH welfare indicator		
	coef.	t-statistic
HH cooks over collected firewood	-[illegible]098	(366)*
Household size	-[illegible]8	(8[illegible]*
Household size squared	[illegible]05	(1316)*
Education level of household head	[illegible]6	(659)*
No.of salaried HH members	[illegible]0	(8[illegible]*
HH owns a bicycle	[illegible]09	(58)*
HH owns a car or motor cycle	[illegible]655	(90)*
HH owns a fridge	[illegible]69	(567)*
Purchased sugar-diary	[illegible]041	(468)*
Total acreage cultivated	[illegible]01	(468)*
HH grows tobacco	[illegible]00	(80)*
HH owns a bed	[illegible]2	(2[illegible]*
No.cattle owned	[illegible]04	(99)*
HH grows hybrid maize	[illegible]7	(28)*
Central rural	[illegible]06	-[illegible]2
Northern rural	-[illegible]05	-[illegible]2
Constant	[illegible]86	(3561)*
Observations	*5539*	
R-squared	*0.33*	

Alternate rural Malawi model – 3		
Dependent. variable:HH welfare indicator (unlogged)		
	coef.	t-statistic
HH cooks over collected firewood	-[illegible]362	(399)*
Household size	-[illegible]374	(130)*
Household size squared	[illegible]099	(913)*
Education level of household head	[illegible]138	(2[illegible]*
No.of salaried HH members	[illegible]148	(86)*
HH owns a bicycle	[illegible]181	(473)*
HH owns a car or motor cycle	[illegible]13335	(561)*
HH owns a fridge	[illegible]14498	(38[illegible]*
Purchased sugar-diary	[illegible]158	(50[illegible]*
Total acreage cultivated	[illegible]044	(42)*
HH grows tobacco	[illegible]16	(192)*
HH owns a bed	[illegible]318	(518)*
No.cattle owned	[illegible]0	(18)*
HH grows hybrid maize	[illegible]76	-16
Central rural	-[illegible]09	-[illegible]04
Northern rural	-[illegible]12	-[illegible]07
Constant	[illegible]955	(1399)*
Observations	*5539*	
R-squared	*0.24*	

Robust t statistics in parentheses
** significant at 10%; ** significant at 5%; *** significant at 1%*

Table 4: (continued)

Alternative urban Malawi model		
Dependent variable:HH welfare indicator (unlogged)		
	coef.	t-statistic
HH owns a fridge	1[illegible]22	(454)*
Household size	-5511	(5[illegible]*
Household size squared	[illegible]0	(4.16)*
Age of head of household	[illegible]2	(3[illegible]*
Education level of household head	[illegible]214	(395)*
No.of salaried HH members	[illegible]263	-[illegible]2
HH owns a car or motor cycle	[illegible]8637	(4.1[illegible]*
HH get lighting from electricity or gas	[illegible]397	([illegible]*
HH owns a bed	[illegible]293	(193)*
Blantyre City	[illegible]169	-[illegible]8
Zomba Municipality	-[illegible]186	-[illegible]6
Mzuzu City	-[illegible]261	-[illegible]8
Constant	9462	([illegible]*
Observations	*872*	
R-squared	*0.41*	

Robust t statistics in parentheses
** significant at 10%; ** significant at 5%; *** significant at 1%*

Table 5: Poverty and ultra-poverty status predictive performance of proxy means models (percent)

	Rural models				Urban models	
Model	**Preferred rural**	**Alt. 1**	**Alt. 2**	**Alt. 3**	**Preferred urban**	**Alt. 1**
Dependent variable	*logged*	*not logged*	*logged*	*not logged*	*logged*	*not logged*
location dummy variables	*district*	*district*	*region*	*region*	*urban center*	*urban center*
Poverty - error of inclusion[1]	352	1[illegible]7	39	18 7	16	[illegible]8
error of exclusion[2]	[illegible]27	457	[illegible]3	48 4	[illegible]2	410
pct.households misclassified	[illegible]2. 4	3[illegible]2	316	349	[illegible]0	[illegible]23
Model-based poverty headcount	5[illegible]7	3[illegible]7	5[illegible]7	367	[illegible]48	34 7
Unweighted IHS household poverty headcount	548	54 8	548	548	518	518
Weighted IHS poverty headcount[3]	[illegible]60	[illegible]66	6 [illegible]6	[illegible]66	[illegible]56	50 6
Ultra-poverty - error of inclusion[1]	63	65	66	6 5	[illegible]7	98
error of exclusion[2]	668	699	[illegible]33	[illegible]7 4	436	444
pct.households misclassified	[illegible]0 9	[illegible]29	[illegible]2 [illegible]2		[illegible]17	[illegible]18
Model-based ultra-poverty headcount	[illegible]12	[illegible]12	11. 5	111	[illegible]8	2 17
Unweighted IHS household ultra-poverty headcount	[illegible]23	[illegible]2 3	[illegible]23	[illegible]23	[illegible]82	[illegible]82
Weighted IHS ultra-poverty headcount[3]	[illegible]23	[illegible]23	2 [illegible]93	[illegible]23	[illegible]28	[illegible]3 8

[1] The error of inclusion is the percent of nonpoor (non-ultra-poor) households that are categorized as poor (ultra-poor) by the model.

[2] The error of exclusion is the percent of poor (ultra-poor) households that are categorized as nonpoor (non-ultra-poor) by the model.

[3] The weighted national poverty headcount calculated using the 16[illegible]9 household IHS data set was 653 pe rcent.However, the 6,58 household data set used here gives a national poverty headcount of 596 percent, as the 4,112 h ousehold which do not feature in the smaller data set are poorer than those retained.

Table 6: Welfare quintile cross-tabulations of observed vs. predicted (percent of households).

The following cross-tabulations show the percent distribution of households across quintiles (1=lowest) of the predicted welfare indicator by quintiles of the actual welfare indicator values for households for each model. The boxed figures across the diagonals identify the percentage of households that are correctly predicted to be in their actual welfare quintile.

As an example of how to interpret these tables, the 50 in the upper left number in the first table below indicates that 50 percent of IHS rural sample households whose actual welfare indicator puts them into the lowest welfare quintile are correctly placed into this quintile using the predicted welfare indicator from the model. However, the figure to the immediate right indicates that 27 percent of IHS households who are actually in the lowest welfare quintile were predicted by the model to be in the second lowest quintile. Finally, the number immediately below indicates that 25 percent of households who are actually in the second lowest quintile were predicted by the model to be in the lowest quintile.

Rural models

Preferred rural model – logged welfare indicator, district location dummy variables

	Predicted				
Actual	1	2	3	4	5
1	50	27	14	7	2
2	25	38	23	17	7
3	15	21	28	23	14
4	8	16	24	28	25
5	2	8	13	25	51

Alternate 1 rural model – unlogged welfare indicator, district location dummy variables

	Predicted				
Actual	1	2	3	4	5
1	49	27	15	7	2
2	26	36	24	16	8
3	15	25	28	23	14
4	7	17	24	28	24
5	2	9	12	25	51

Alternate 2 rural model – logged welfare indicator, region location dummy variables

	Predicted				
Actual	1	2	3	4	5
1	46	27	15	8	3
2	24	27	24	19	6
3	16	20	23	26	15
4	10	17	23	26	25
5	3	9	14	23	51

Alternate 3 rural model – unlogged welfare indicator, region location dummy variables

	Predicted				
Actual	1	2	3	4	5
1	46	27	15	9	3
2	24	26	24	18	7
3	17	21	24	23	15
4	9	17	22	27	25
5	4	9	14	23	51

Table 6 (continued)

Urban models

Preferred urban model – logged welfare indicator, urban center location dummy variables

	Predicted				
Actual	1	2	3	4	5
1	59	31	10	3	0
2	23	38	29	14	1
3	11	27	34	28	9
4	5	7	20	36	30
5	1	2	8	19	60

Alternate 1 urban model – unlogged welfare indicator, urban center location dummy variables

	Predicted				
Actual	1	2	3	4	5
1	55	29	15	2	1
2	30	37	30	13	2
3	11	22	33	27	10
4	6	8	20	36	29
5	1	4	6	22	58

Table 7: Poverty indicators obtained from wealth ranking exercise and group interviews from selected communities.

District:	Machinga	Mzimba	Ntchisi	Lilongwe City	Mzuzu City	
Village:	Mwamadi	Chalamira Shonga	Chikho	Kauma	Masasa	Comments
Housing						
Roofing material			x	x	x	Iron roofing for nonpoor;grass-th atched roofing for poor
Type of floor			x		x	Cement floors for nonpoor;earth for poor
Type of wall material			x	x		Brnt brick or cement for nonpoor ;mud and poles for poor
Toilet facilities					x	Poor quality or absence of toilet facilities for poor
Ventilation				x		Poor may have no windows at all
Size of housing / quality of housing	x	x		x		Space seemed to be more of a constraint, such that 'space per capita'is also an indicator of status
Owns housing to let				x	x	People who have houses for rent are considered better off.
Piped water					x	
Electricity					x	
Insect infestation / Unkempt housing		x		x		Characteristic of poorer households
Other						Received comments that housing may not matter as long as household has a lot of money.
Demographic						
Widowed (female)	x	x				These three factors are among those cited as descriptions of the heads of poor households.
Handicapped / Ill	x	x				
Elderly	x					
HH takes care of orphans	x			x		We were told that the problem of caring for orphans cuts across all income classes.It adds a burden to all classes.
Other assets						
Livestock	x	x	x			Owning cattle is mark of a nonpoo r household.Ownership of medium-sized livestock associated with 'middle' groups of households, and of small-livestock or absence with poor.
Sofa set		x	x	x	x	
Cupboard or display cabinet				x	x	
TV / refrigerator / phones /				x	x	
Radio				x	x	Two respondents said household may own radio but cannot purchase batteries for it - usage a factor.
Bicycle	x					In Machinga, bicycles for hire seemed to be a big business.
Sewing machine	x					
Cars / motor vehicles		x			x	
Cooking utensils / Crockery	x			x	x	
Beddings / Beds / Mattresses		x		x		Nonpoor have beds and mattresses;poor use mats.Nonpoor have sufficient linen;poor use *chiterjes* or fertilizer sacks.
Own oxcart to let			x			
Employment / Source of income						
Salaried employment or ownership of a business enterprise'	x		x	x	x	Steady sources of income associated with nonpoor, especially salaried employment.Type of bus iness activity matters.Small-scale reselling of vegetables associated with poor;ownership of grocery stores with nonpoor.
Transfers of food/nonfood	x					Nonpoor can afford to give assistance to others;poor may depend on assistance for sustenance.
Ganyu labor (see Agriculture)						
Consumption						
Self-sufficiency in food throughout year		x	x		x	Poor are not self-sufficient in food year round.
Size of food purchases					x	Poor buy staples in small quantities like so-called 'walkman' (small container the size of a walkman used to sell maize); nonpoor able to buy in large quantities.
Type of food bought					x	Nonpoor households eat meat and fish
Quality of clothing		x		x	x	Better quality clothing or more ch anges of clothing for nonpoor.Availability of second -hand clothing cited as one reason for staying in urban areas.La ck of clothing/changes of clothing also affects school attendance - children from poorer households may not go to school if they dont have clothes.

(continued)

Table 7: (continued)

District:	Machinga	Mzimba	Ntchisi	Lilongwe City	Mzuzu City	
Village:	Mwamadi	Chalamira Shonga	Chikho	Kauma	Masasa	Comments
Agriculture						
Farm size		x	x		x	The larger the less poor.No cutof fs mentioned.
Use of fertilizer	x		x	x	x	Nonpoor households can afford to buy fertilizer, middle-groups purchase fertilizer on credit, poor cannot afford to buy fertilizer or they use manure.Note that need for fertilizer is area-specific.Some areas have rich soils and dont need fertilizer.
Transport of produce				x		Nonpoor households able to transport their produce
Ganyu labor		x	x	x		Nonpoor households can afford to hire ganyu labor or tenants. Poorer households offer their ganyu labor.A household completely reliant on ganyu for its income is considered poor.
Type of crops grown			x			Households that grow tobacco and other cash crops considered better off as dont rely on own-production of food crops for subsistence.Households that plant mainly food crops and consume this for their own-subsistence are poorer.
Sale of food crops	x					Small-scale resale of vegetables considered activity of poor.
Schooling						
Ability to send children to private school		x			x	
Lack of clothing affects schooling attendance				x		
Other						
Assistance at funerals		x	x			Nonpoor able to assist at funerals
Ability to get loans						Associated with nonpoor.Poor ca nnot offer collateral for loans.
Laziness	x	x				Associated with poor.In Mzimba, 'hardworking' was often added when describing the nonpoor.Presence of 'drunkards' associated with non-economically active and therefore poor.
Begging / Prostitution		x		x		Economic activities associated with the poor
Number of wives	x					

Note: Xs indicate variables brought up by respondents without prompting.

Table 8: Cross tabulation of welfare quintiles (derived by wealth ranking) with potential household welfare proxy indicators, by percent of households in quintile.

	Do all household members have shoes?			Do household members use toothpaste?		
Quintile	Yes	No		Yes	No	Occasionally
Poorest	41	59		15	82	3
2	62	38		18	76	6
3	76	24		24	73	3
4	56	44		38	56	6
Wealthiest	85	15		42	55	3
Total	64	36		27	68	4

Table 9: Potentially useful proxy welfare indicators for Malawian households.

Household characteristic	Not in IHS	Proven good indicator, not in IHS
No.HH members doing ganyu wo rk	X	X
Anyone in HH ever migrate for work?	X	X
No.rooms in the house	X	X
Type of toilet facility	X	X
No.radios owned	X	X
No.blankets/sheets owned	X	X
No.mattresses owned	X	X
Do all HH members have shoes	X	X
Do HH members use toothpaste	X	X
Does HH have a sofa	X	X
Does HH engage in resale of vegetables	X	X
Does HH use detergent for bathing	X	X
Material of roof	X	
Material of wall	X	
Material of floor	X	
Type of window	X	
Type of door	X	
Where food is served	X	
What HH members sleep on - beds, mattresses, mats	X	
No.mats owned	X	
Household size		
No.HH members present past 2 weeks		
Gender of head		
Age of head		
No.HH members under 5 years		
Schooling of household head		
Marital status		
Education of salaried member (highest)		
No.HH members 6-18 years act ually in school		
No.HH members in private schoo l		
Avg.immunization level of under-5 s		
HH tenancy status		
HH has electricity		
Source of drinking water		
Source of cooking fuel		
Source of lighting		
No.bicycles owned		
No.ploughs owned		
No.fridges owned		
No.cars/motorbikes owned		
No.livestock owned -large/medi um/small		
No.beds owned		
Size of land cultivated (acres);la nd owned		
Does HH grow hybrid maize		
Does HH grow tobacco		
Does HH use fertilizers on food crops		
Did HH receive noncash transfer past month		
Did HH receive food transfer past month		
Did HH give noncash transfer past month		
Did HH give food past month		
Did HH buy sugar past month		
Did HH buy salt past month		
Did HH buy petroleum jelly past month		
Did HH buy lotion past month		
Did HH buy soap past month		

Table 10: 2002 Core Welfare Indicators Questionnaire survey – Are the variables for the proxy means test models present in the questionnaire?

Variable	Data will be collected in the 2002 CWIQ?
Rural model	
Household size	yes
Household size squared	yes
Education level of household head	yes
No.of salaried HH members	yes
HH owns a bicycle	yes
HH owns a car or motor cycle	yes
HH owns a fridge	yes
Total acreage cultivated	qualified yes
HH grows tobacco	yes
HH owns a bed	yes
Purchased sugar-diary	qualified yes
HH cooks over collected firewood	yes
No.cattle owned	yes
HH grows hybrid maize	yes
Urban model	
Household size	yes
Household size squared	yes
Age of head of household	yes
Education level of household head	yes
No.of salaried HH members	yes
HH owns a car or motor cycle	yes
HH owns a fridge	yes
HH owns a bed	yes
HH gets lighting from electricity or gas	yes

14.18. Data Entry by Scanning

Dataentry by Scanning

Experience with the software Eyes&Hands (Documents) FORMS ver. 5-2

A presentation held in Maputo, Mosambique December 13, 2005

Per Schøning
Senior Adviser
Division for Development Cooperation
Statistics Norway
per@ssb.no

Statistisk sentralbyrå

Background:

- Scanning of forms was introduced in Statistics Norway as early as the Population and Housing Census 19
- Statistics Norway has experience with the software Eyes Hands (Documents) FORMS since 1994.
- Statistics Norway cooperated with Uganda Bureau of Statistics on agricultural statistics development in 2 and Eyes Hands was tested under Uganda conditions.
- Statistics Norway introduced Eyes Hands for a survey project in INSE Malawi in 2

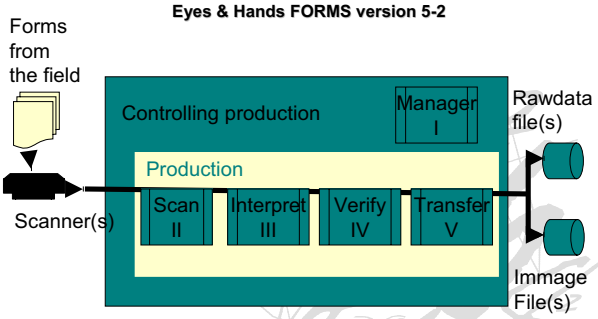
Statistisk sentralbyrå

Eyes & Hands (Documents) FORMS version 5-2

- A software for Optical Character Reading (OR)
- Produced by ReadSoft AB Box SE-191 Sollentuna, Sweden
- The software has five modules:
  1. Manager
  2. Scan
  3. Interpret
  4. Verify
  5. Transfer
- The software brings filled in information from questionnaires via a scanner through an automatic interpertation module and a manual additional verification process and a final data transfer process to the output product.
- The output is a rawdata file and optionally also a database of imagefiles of scanned forms
- By good questionnaire design and treatment, good training of enumerators/respondents and by optimal programming of the EH software, the manual verification process can be limited to a minimum and a fast and accurate data capture is obtained.

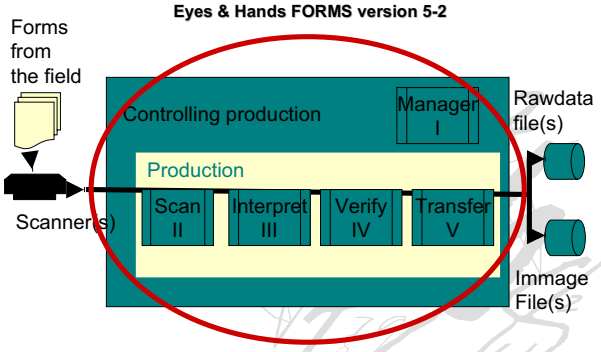
Statistisk sentralbyrå

Eyes & Hands FORMS version 5-2



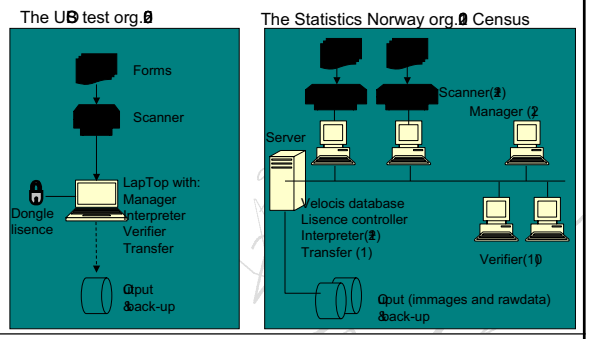
Statistisk sentralbyrå

Eyes & Hands FORMS version 5-2



Statistisk sentralbyrå

Eyes & Hands FORMS version 5-2
Examples of organization



Statistisk sentralbyrå

Eyes & Hands (Documents) FORMS version 5-2
Specifications from the Norwegian Household Census 2001

- 17 mill forms for scanning (2 mill forms total)
- Scanned, interpreted and verified in 42 work days
- 2 high capacity scanners (30 sheets/hour) used most of the time
- Eyes & Hands lisences used:
  - 2 (3) scan lisences
  - 3 interpertation lisences
  - 10 verification lisences

- Number of staff used:
  - 9 persons in 2 shifts for opening envelopes, unfoiling the forms and for bringing forms to and from the work areas
  - 4 to 5 persons in 2 shifts to operate the scanner and store the scanned forms
  - 10 persons in 1 shift for verification
  - Interpreting was done unattended during the night
  - Data transfer was done every morning by the supervisor

Statistisk sentralbyrå
Statistics Norway

Eyes & Hands (Documents) FORMS version 5-2
Specifications from the agric. pilot in UBOS 2005

Low capacity scanner: Fujitsu 4120
Capacity 1 000 sheets per day
Approx.price: USD 1 200

Eyes & Hands FORMS stand allone Standard version:
Capacity: 3 600 sheets per hour
Approximate price: USD 18 000

Annual support/update agreement:
18% of software price (optional after the first year)

Size & organisation of the UBOS pilot agricultural survey project:

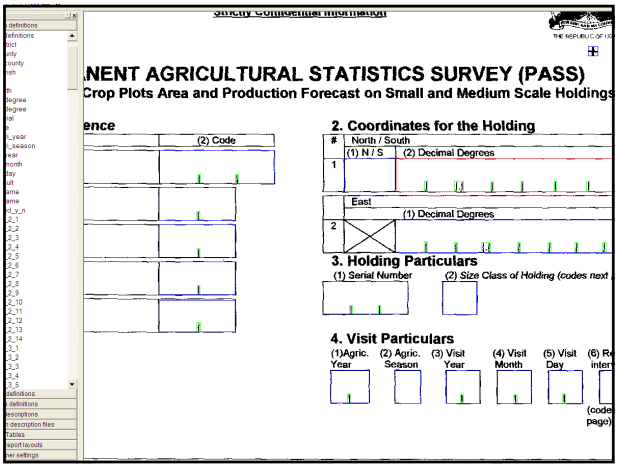
- 3 600 (15x48+5) agricultural holdings sampeled in a total of 5 districts
- 12 480 sheets for scanning per agricultural season x 2 seasons
- 3 data-entry staff working on the scanning and logistics (second season 2)
- 1 consultant working on further processing and tabulation (SPSS)

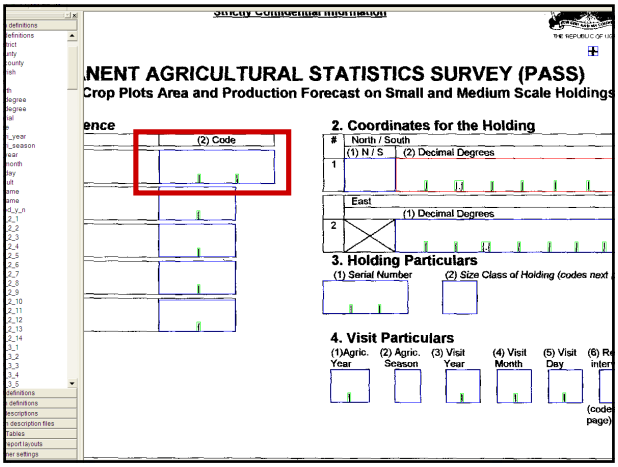
Statistisk sentralbyrå
Statistics Norway

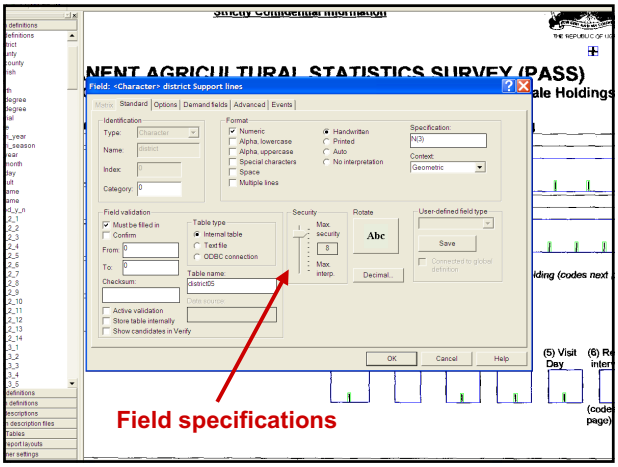
Eyes & Hands FORMS Product Characteristics

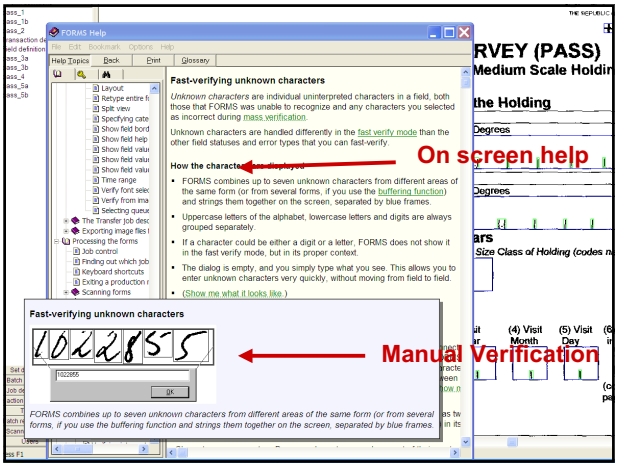
FORMS license type	Capacity		
	Machin printed interpretation character/sec	Handwritten interpretation character/sec	Forms per hour (pages)
Lite	15	7	600
Standard	30	15	3 600
Professional	50	25	7 200
Interpret 30	30	15	3 600
Interpret 300	300	150	30 000

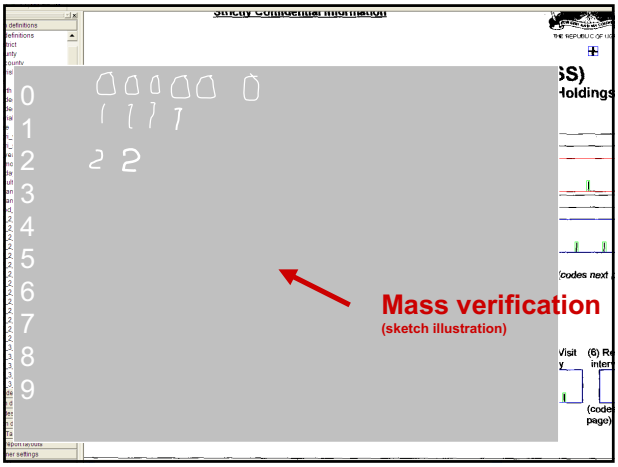
Statistisk sentralbyrå
Statistics Norway

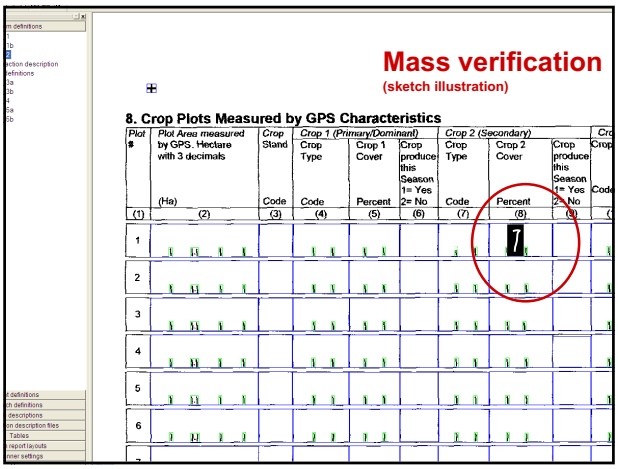


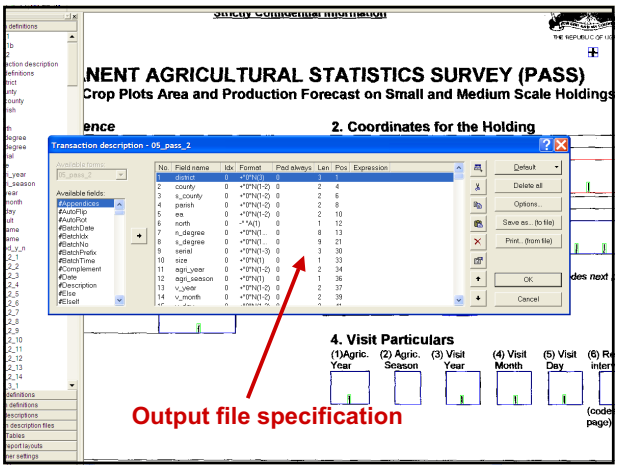












Output file formats:

Many options for export formats from ESS:

- fixed positions files
- semicolon separated
- immage files (pdf, TIFF etc)

For import and further processing in SPSS, STATA, SAS, EXCEL etc

```
1000066S0031556210141021LAWRENCE      LUBGA          410
1000066S0991561340021MBAMAD           SENTAMU        18
1000066S0031556102021MARIA            NALWANGA       18
```

```
1000066S0031556210LAWRENCELUBGA410
1000066S0991561340MBAMADSENTAMU18
1000066S0031556102MARIANALWANGA18
```

Statistisk sentralbyrå

Discussions/argumentation met when scanning was suggested for piloting in UBOS:

- We know what we have, not what we get"ie.general scepticism to new technology
- Red dust will destroy the scanner very fast"
- Vulnerability to breakdowns (no local service and support)
- Bad quality of locally printed forms (paper and accuracy)
- Dirt and rough treatment of forms during the field work expected to hamper the scanning
- Bad hand writing by field workers
- No clear incentive for local staff to learn and take responsibility for the new technology/method:
  - Bad policy to replace manual labour with technology
  - Difficult to build sustainable local technical capacity and ownership"

Statistisk sentralbyrå

Why it finally was concluded to pilot scanning in UBOS:

- Expected very fast data entry and processing with a small and easily manageable organization - "we have to be costefficient and deliver fast"
- Top management wanted to use "st ate of the art" technology
- Manual dataentry (CsPro) from the UBOS pilot agricultural census took far more time (pre entry checking + dataentry + vali dation) and staff than expected, and this staff was organized outside direct control of the agricultural project
- Testing new technology/methods was a core part of the actual project
- Recommendations from Statistics Norway for testing scanning on a modest scale (one software licence and a cheap low capacity scanner)

Statistisk sentralbyrå
Statistics Norway

Preparing the UBOS pilot for scanning

- Scanning require single sheets into the scanner ie.The forms should be single sheets also in the field (or booklets from the field "gilliotined" just before scanning).The pilot in UBOS used single sheets stored in C4 envelopes in the field
- Special design of forms taking full advantage of Optical Character Reading (OCR) and the chosen scanner's capacity for double sided scanning of A4 (or A3 sheets)
- Establishing the organization (staff, loaclities, storage and technical equipment)
- Programming (scanning immage, interpertation controls and output/transfer format)
- Capacity building of local staff, including changing staffs attitude from the previous resource demanding *pre entry manual data cleaning* to "on-screen" *post entry data cleaning* and validation

Statistisk sentralbyrå
Statistics Norway

Experience from scanning with E&H FORMS in the Uganda pilot:

- Forms can be constructed with any available software
- Standard copier paper is good enough for scanning
- Local printing company was used, but some problems occured with the printing quality
- The more auxilary information that is preprogrammed (look-up tables, valid range and mass-verification etc), the higher score of automatic interpretation
- The better handwriting and correct filling-ins, the higher dataentry efficiency (ie.Follow up on the field workers)
- Garbage in gives garbage out and call-back to the fieldworkers/holders was too expensive (ie.Follow up on the field workers)
- Practical experience (from Norway) indicates that up to 97 percent of characters are automatically interpreted
- E&H is sensitive for uncorrect positioning of the orientation marks on the printed pages, dirty and filthy forms and for low quality copies of questionnaires made in the field.
- Leading zeros filled in numeric fields caused some unneccessary problems and slowed down interpretation speed
- More testing should have been done before we used scanning for the first agricultural season pilot
- Some technical start-up problems when local staff tried to reinstall the software
- Many start-up problems and much testing before we achieved optimal interpretation of high quality
- Remote technical support works (by e-mail and telephone, in the Uganda case to the Norwegian supplier)
- Dust problem can be reduced to acceptable level by cleaning and covering the scanner
- Biggest problem was the change of organization and "involvement" of the IT- and dataentry staff

Statistisk sentralbyrå
Statistics Norway

Findings summary:

- Time used for dataentry and number of entry-staff is considerabely reduced by using scanning techniques compared to traditional data entry
- The bottleneck of the scanning-part is normally not the scanner or the interpretation capacity, but the manual verification
- The capacity of scanning technology was not fully exploited in the UBOS pilot

Challenges:

- Build up local technical capacity and continuity on the use of E&H
- Ensure efficient logistics - inflow of forms
- Take optimal advantage of E&H program facilities to increase share (and quality) of the automatic interpretation
- Build up good routines/programs for data processing also *after* scanning ie.use of imagefiles for on-screen post checking/data cleaning, storage and tabulation.

Conclution:

- Scanning should be an interesting option for reduction of the time used from data collection to data publishing - also under African conditions
- Eyes & Hands (Documents) is a feasible choice for scanning software
- Much practical testing&failing should be foreseen before the software can be efficiently used for a survey/census
- Think "ORS" rather than "ORT"

Statistisk sentralbyrå
Statistics Norway

14.19. Streamlined traditional manual keying - Mozambique

Instituto Nacional de Estatística

Streamlined traditional manual keying

Workshop on Light Core Surveys
December 12th to 14th, 2005
Maputo

Instituto Nacional de Estatística

An outline of the coding and initial quality control process

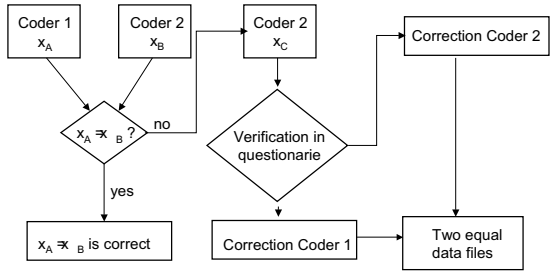
The coding of open ended questions and the control of skip patterns were done in three levels:

1. First control group: Codes the open ended questions and check the skip patterns.
2. Second control group: Reviews the codes in the open ended questions and the skip patterns.
3. Supervisor control: The last stage is that the supervisor reviews the controls done.

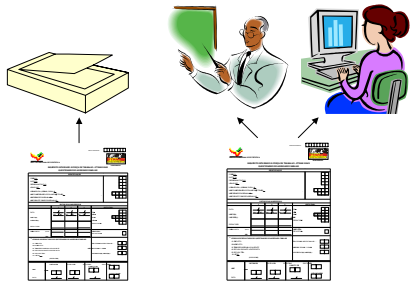
After this the questionnaires taken to the data entry.

Instituto Nacional de Estatística

The data entry process



Instituto Nacional de Estatística



Instituto Nacional de Estatística

The time needed

- Traditional data entry is generally considered to be time-consuming.This is however, not necessarily always the case.
- The data entry for IFTRAB ended one month after the data collection.
- Sufficient capacity to process the questionnaires as they came in to the central office during the field work.

Instituto Nacional de Estatística

The choise between scanning and data entry

- Time consuming to make and test a scannable questionnarie.
- Faster to make a data entry program than a scanning program (?)
- More manpower and management needed for data entry.
- Costs were considered to be ><=
- The suevey could go in the field earlier

Instituto Nacional de Estatística

Correcting errors and cleaning data

- CSPro used for data entry and making data checks.
- Error control at macro level is done on an aggregated level
- It is only important to correct errors that change our estimates.

Instituto Nacional de Estatística

Summing up:

- Manual coding and skip pattern control
- Controls on unit level mainly in data entry
- Data entered two times and compared
- Review frequency distributions and key tables
- Data file ready for use.

14.20. Rwanda Experience and Practices in Monitoring National PRSP

Rwanda Experience and Practices in Monitoring National PRSP

Augustin Twagirumukiza[1]
Dr. Geoffrey Greenwell[2]
Dr. Ir. Louis Munyakazi[3]

[1]Head of Computer Processing Division
National Institute of Statistics of Rwanda
[2] OPM Consultant
[3] Director General Rwanda Bureau of Statistics
Workshop on Light Monitoring Surveys Maputo Dec 12-14 2005

Objective

Applying best practices from the DHS and MICS to Household Surveys

Rwanda EICV (Household Survey) fully adapted the Data processing system as used by the DHS and MICS.

Fully extensible system adaptable to other surveys.

Saves on training, reinforces best practices in data Processing.

Could be replicable in other countries

Rationalization of resources

Unicef/Macro training is applicable for household surveys and vice versa. The World Bank could build on this synergy.

Added some critical edits:

Inserting blank records for household members excluded From a section (not applicable universe)

Dynamic range checks

Odd ending expenditures flagged during data entry (those not ending in 0 or 5)

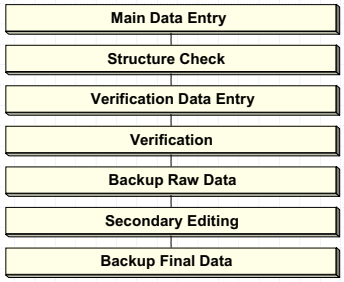
Data Processing Design

- Data processing is organized around clusters
- There is one set of data files for each cluster
- Allows us to
  - Process data in parallel with data collection
  - Allows feedback to the field
  - Keeps size of data files manageable
- Data file names include cluster number

Data Processing Design

- Data processing is split into two phases
  - Primary
  - Secondary
- Goal of primary phase
  - Clean, edited data
- Goal of secondary phase
  - Analysis files and tables

Primary Data Processing Flow



Primary Data Processing

- Main data entry
  - First time data is entered
- Structure check
  - Checks structure of data files
- Verification data entry
  - Second time data is entered
- Verification
  - Two data files are compared; differences resolved

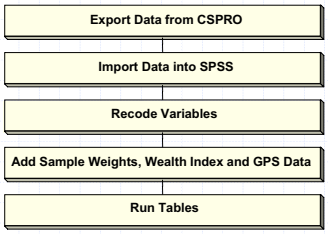
Primary Data Processing

- Raw data backup
  - Verified data are backed up to a separate directory
- Secondary editing
  - Complex inconsistencies are investigated
- Final data backup
  - Edited data are backed up to a separate directory

Control Sheet

- Keeps track of data processing
- Can be entered both on paper and on the supervisor's computer
- One row for each cluster
- Enter
  - Dates each task completed
  - Number of data entry operators
- Use total rows to track progress

Secondary Data Processing Flow



Secondary Data Processing

- Exporting data from CSPRO
  - Create SPSS data file and syntax file from CSPRO data file and dictionary
- Importing data to SPSS
  - Executing syntax file created by CSPRO
- Recoding variables
  - Creating new variables and recoding old variables

Secondary Data Processing

- Adding sample weights
  - Sample weights are added from weights spreadsheet
- Adding wealth index
  - Wealth index calculated then added to files
- Adding GPS data
  - Geographic location data added to files
- Tabulation
  - Tables are generated from the analysis files

Data Processing Personnel

- Questionnaire administrators
- Data entry operators
- Secondary editors
- Data processing supervisor

Questionnaire Administrators

- Receive clusters from the field
- Check that all questionnaires are present
- Check that questionnaires are ready to enter
- Keep track of location of all clusters

- Should follow interviewer training

Data Entry Operators

- Enter main data
- Enter verification data
- Resolve differences between files

- Must follow interviewer training
- Must be familiar with the questionnaires

Secondary Editors

- Investigate complex inconsistencies
- Tell supervisor if and how to resolve inconsistencies
- Review editing guidelines

- Must be present during interviewer training
- Need excellent understanding of questionnaire and goals of survey

Data Processing Supervisor

- Resolves data entry problems
- Maintains programs
- Oversees entire data processing system

- Must have excellent grasp of questionnaire
- Must have programming skills in SPSS and CSPRO

Questionnaire Administrator Training

- Review list of checks in data processing chapter
- Give QA several clusters and check work
- Establish questionnaire storage procedures

Data Entry Training

- Begin when you have one cluster for each data entry operator
- Allows you to
  - Train data entry operators
  - Debug programs
- Practice verification at the same time
- When you have finished
  - Fix entry programs
  - Delete data files

Secondary Editor Training

- Wait until you have 3-4 error listings
- Give secondary editors
  - Copy of editing guidelines
  - A cluster's error listing
  - The cluster's questionnaires
- Review work with secondary editor
- Try to schedule a day when data entry operators aren't working

Data Processing Equipment

- Data entry machines
  - Pentium, Win 98+, 16 Mb RAM, 500 Mb hard drive & floppy drive
- Supervisor's machine
  - Pentium, Win 98+, 64 Mb RAM, 500 Mb hard drive, floppy drive, secondary storage device
- Uninterrupted power supplies
- Surge Protectors

Data Processing Equipment

- A printer
- Paper
- Toner cartridges/printer ribbons
- Flash disk/Diskettes
- Green pens

Data Processing Rooms

- Data Entry
  - Room for computer and an entire questionnaire
- Editing
  - Quiet space for editors to work
- Questionnaires
  - Must contain means (e.g., shelving) to organize questionnaires by cluster

Data Entry Directory Structure

```
CSPRO
  DATA
  DICTS
  ENTRY
  VERI
```

Data Entry Directories

- Data
  - Main data entry files
- Dicts
  - CSPRO dictionary
- Entry
  - Data entry programs
- Veri
  - Verification data entry files

Supervisor Directory Structure

```
CSPRO
    BACKUP
    GPS
    EXPORT
    FINAL
    RAW
    SUPER
SPSS
WEIGHTS
```

Supervisor CSPRO Directories

- Backup
  - Backup of verified data
- GPS (if applicable)
  - GPS data entry program
- Export
  - Programs to transfer data
- Final
  - Backup of edited data
- Raw
  - Data from data entry machines

Supervisor Directories

- Super
  - Supervisor's programs
- SPSS
  - SPSS programs
- Weights
  - Weights spreadsheet

Summary

- By using this strict checking in household surveys allows every body:
- The reliable statistical information.
- A good measurement of PRSP.
- To ensure nation's ability to report the best information on MDGs .

14.21. Tracking Resource&Policy Impact

Tracking Resource and Policy Impact – a general approach and as applied in Malawi

Presented at Workshop on Light Core Surveys, Maputo, Mozambique, 12-14 December 2
by Bjørn Kr Wold, Statistics Norway, Kongensgt 6 N-0033 Oslo, Norway, bkw@ssb.no

Wold, Bjørn Kr, Rauan, Shannessen, Steen:Tracking Resource and Policy Impact.Incorporating Millennium Development Goals Indicators and Poverty Reduction Strategy Papers monitoring across sectors. Statistics Norway, Oslo - http//wwwssbno/english/subjects/00/00/app_20n/

Wold, Bjørn Kr et al:Tracking Resource and Policy Impact in Malawi.Incorporating Malawi Poverty Reduction Strategy Paper Indicators, Millennium Development Goals Poverty Monitoring Across Sectors. Stat.Norway National Statistical Office, Lilongwe, Malawi - http//www. ssbno/emner/00/00/app_02

Statistisk sentralbyrå
Statistics Norway
1

Why Tracking Policy Impact ?
Last decade growing focus on efficiency and impact of resource allocation

- UN has agreed upon 8 Millennium Goals, 15 Targets and 48-66 indicators.Annual Reporting led by UNDP
- HIPC agreement with IMF and World Bank requiring Poverty Reduction Strategy Papers (PRSP) including a monitoring component.
- Bilateral donors addressing output, outcome and impact monitoring
- Developing countries increasing focus on resource efficiency

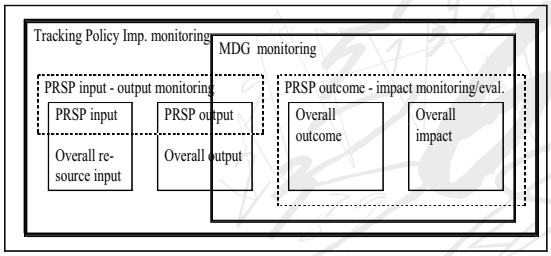
Statistisk sentralbyrå
Statistics Norway
2

Why Tracking Policy Impact ?
Some additional issues

- Current trends
  - From public to private markets and agents, globalization, poverty reduction
  - Decentralization
  - Donor support for democracy and good governance
  - From project and program to **sector support**
- Trends to come
  - Policy design towards Millennium Development Goals
  - Causal analysis of poverty
  - Public control of more or less competitive markets
  - Public control AND participation within competitive markets

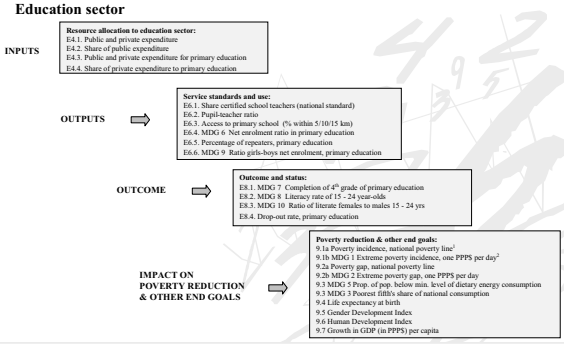
Statistisk sentralbyrå
Statistics Norway
3

Why Tracking Policy Impact ? Methodological trends



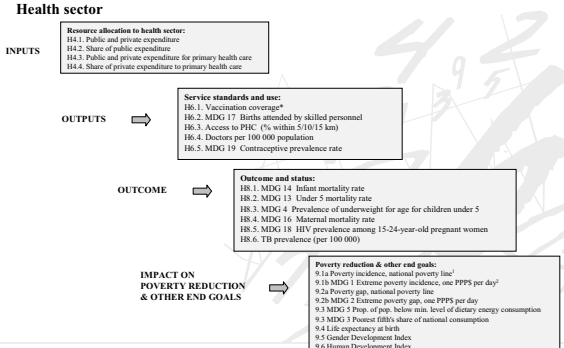
Statistisk sentralbyrå
Statistics Norway
4

Tracking Policy Impact in Education Sector



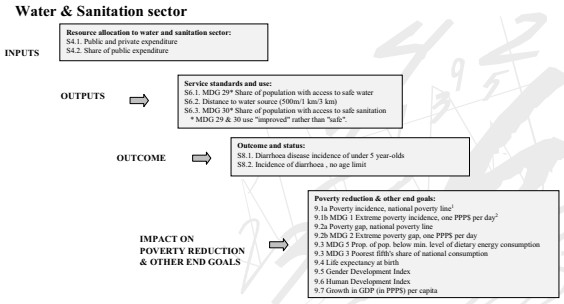
Statistisk sentralbyrå
Statistics Norway
5

Tracking Policy Impact in Health Sector

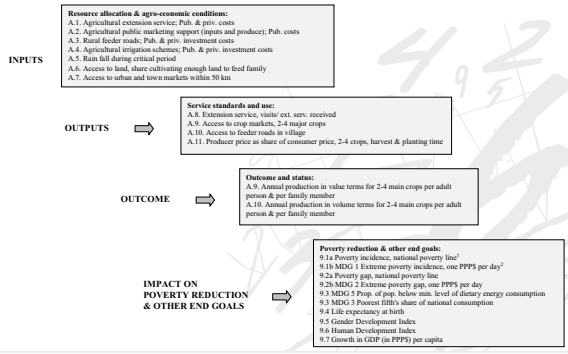


Statistisk sentralbyrå
Statistics Norway
6

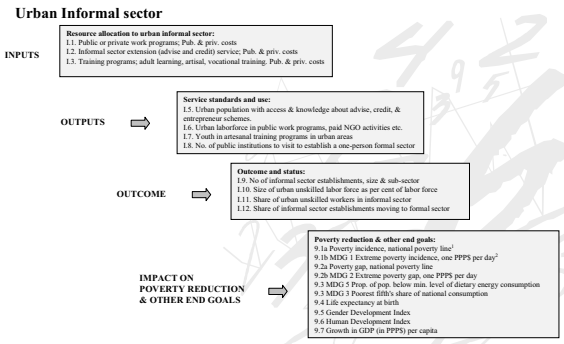
Tracking Policy Impact in Water & Sanitation Sector

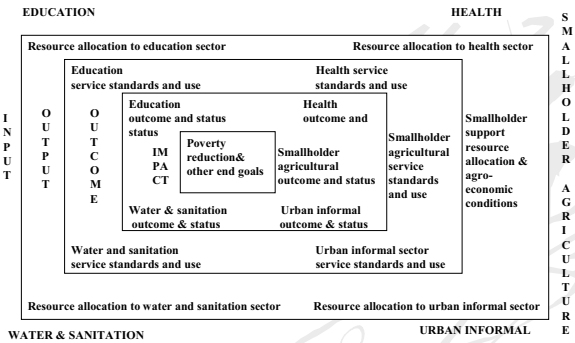


Tracking Policy Impact in Smallholder Agriculture



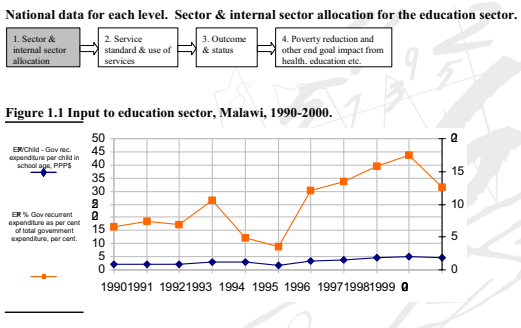
Tracking Policy Impact in Urban Informal Sector





International Data Sources

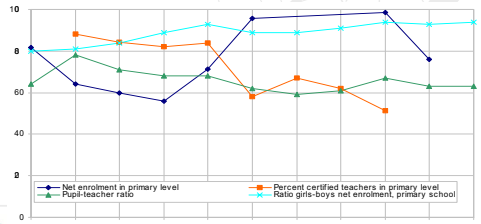
- Sector Data Sources
  - WHO from Ministry of Health, adjust figures
  - FAO mainly from Ministry of Agriculture, also from the NSI, adjust figures
  - UIS from Ministry of Education, also from the NSI, adjust figures
  - UN Population Division from Census, demographic models, demographic surveys, adjust to own model/ predictions.
  - UN Statistical Division, do not adjust but return with ?
  - OECD from membercountries, do no adjust.
  - IMF from Central Banks, NSIs.Adjust and collect.
- World Bank.From NSIs, adjust, may collect.Use UIS and FAO data.Produce WDI in open database, subscribers database, CD-Roms.



National data for each level. Service standard & use of services for the education sector.

1. Sector & internal sector allocation → 2. Service standard & use of services → 3. Outcome & status → 4. Poverty reduction and other end goal impact from health, education etc.

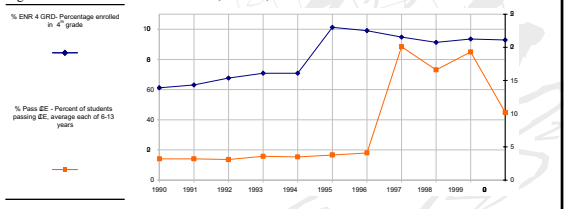
Figure 1.2 Output in education sector, Malawi, 1990-2000.



National data for each level. Sector outcome and status for the education sector

1. Sector & internal sector allocation → 2. Service standard & use of services → 3. Outcome & status → 4. Poverty reduction and other end goal impact from health, education etc.

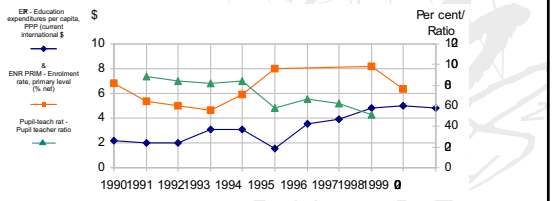
Figure 1.3 Outcome in education sector, Malawi, 1990-2000.



National data two by two levels. Are changes in education sector resource allocation followed by changes in service standard?

1. Sector & internal sector allocation → 2. Service standard & use of services → 3. Outcome & status → 4. Poverty reduction and other end goal impact from health, education etc.

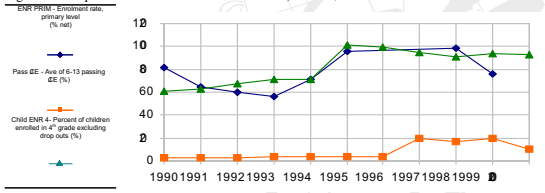
Figure 2.1 Input-Output in the Education Sector, Malawi, 1990-2000.



National data two by two levels. Are changes in education sector service standard and use of services followed by changes in outcome and education status?

1. Sector & internal sector allocation → 2. Service standard & use of services → 3. Outcome & status → 4. Poverty reduction and other end goal impact from health, education etc.

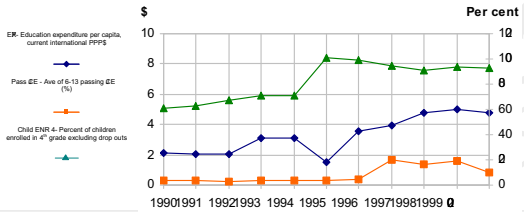
Figure 2.2 Output-Outcome in the Education Sector, Malawi, 1990-2000.



National data two by two levels. Are changes in education sector resource allocation followed by changes in final educational status?

1. Sector & internal sector allocation → 2. Service standard & use of services → 3. Outcome & status → 4. Poverty reduction and other end goal impact from health, education etc.

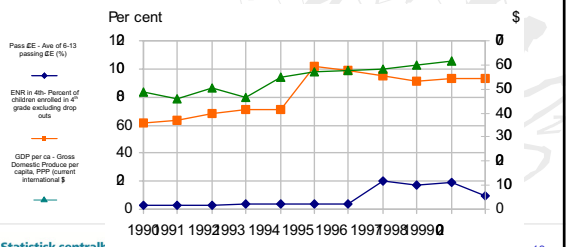
Figure 2.3 Input-outcome in the Education Sector, Malawi, 1990-2000.



National data two by two levels. Are changes in education sector outcome and educational status followed by changes of end goal achievements?

1. Sector & internal sector allocation → 2. Service standard & use of services → 3. Outcome & status → 4. Poverty reduction and other end goal impact from health, education etc.

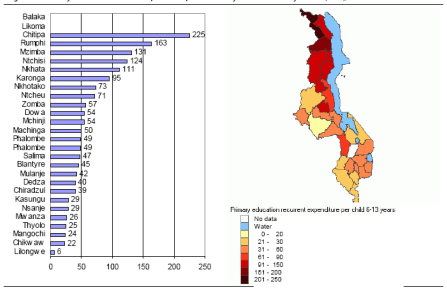
Figure 2.4 Outcome-Impact in the Education Sector, Malawi, 1990-2000.



5.3.1. Sector & internal sector allocation

1. Sector & internal sector allocation → 2. Service standard & use of services → 3. Outcome & status → 4. Poverty reduction and other end goal impact from health, education etc.

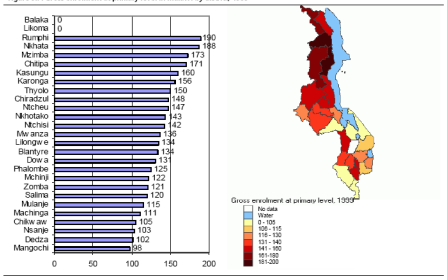
Figure 5.6. Primary education recurrent expenditure per child 6-13 years in Malawi by district (PPP$), 2000



5.3.2. Service standard & use of services

1. Sector & internal sector allocation → 2. Service standard & use of services → 3. Outcome & status → 4. Poverty reduction and other end goal impact from health, education etc.

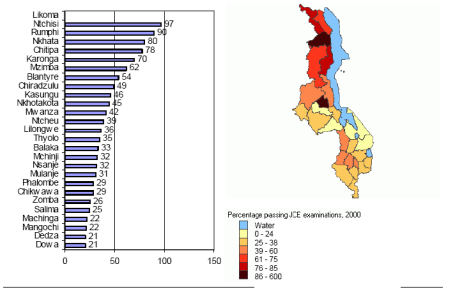
Figure 5.7. Gross enrolment at primary level in Malawi by district, 1999



5.3.3. Outcome status

1. Sector & internal sector allocation → 2. Service standard & use of services → 3. Outcome & status → 4. Poverty reduction and other end goal impact from health, education etc.

Figure 5.8. Percentage passing JCE examinations in Malawi by district, 2000



5.5. The final monitoring step, end goals, human welfare and poverty reduction at district level

1. Sector & internal sector allocation → 2. Service standard & use of services → 3. Outcome & status → 4. Poverty reduction and other end goal impact from health, education etc.

Figure 5.12. Poverty headcount by district in Malawi by district, 1998

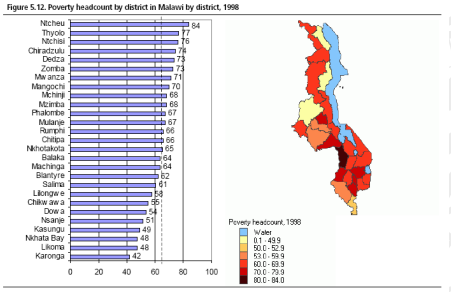
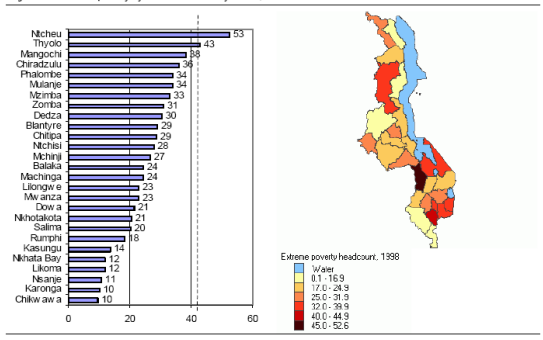


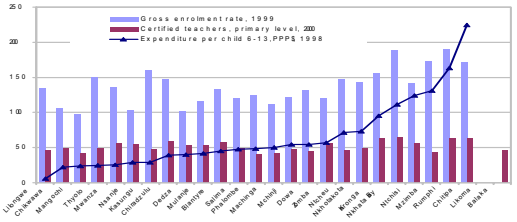
Figure 5.13. Extreme poverty by district in Malawi by district, 1998



District data two by two levels. Are district differences in education sector resource allocation followed by district differences in service standard?

1. Sector & internal sector allocation → 2. Service standard & use of services → 3. Outcome & status → 4. Poverty reduction and other end goal impact from health, education etc.

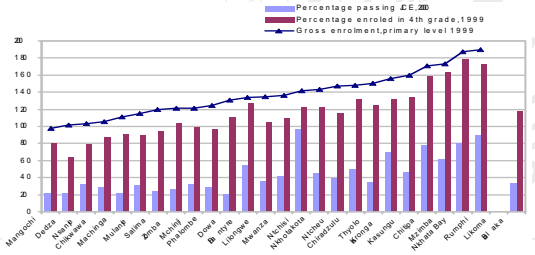
Figure 4.1 Input-output in the education sector, sorted according to expenditures per child.



District data two by two levels. Are district differences in education sector service standard and use of services followed by district differences in outcome and education status?

1. Sector & internal → 2. Service standard & → 3. Outcome → 4. Poverty reduction and

Figure 4.2 Output-outcome the education sector, sorted by gross enrolment.

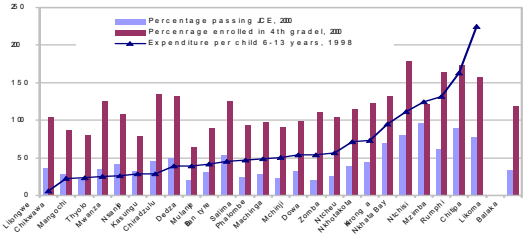


Statistisk sentralbyrå

District data two by two levels. Are district differences in education sector resource allocation followed by district differences in final educational status?

1. Sector & internal sector allocation → 2. Service standard & use of services → 3. Outcome & status → 4. Poverty reduction and other end goal impact from health, education etc.

Figure 4.3 Input-outcome in the education sector, sorted by expenditures.

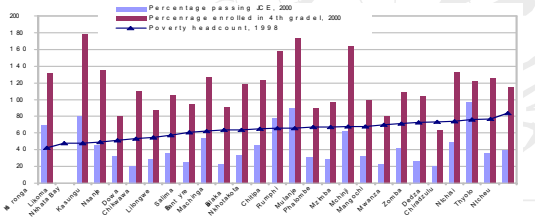


Statistisk sentralbyrå

District data two by two levels. Are district differences in education sector outcome and educational status followed by district differences of end goal achievements?

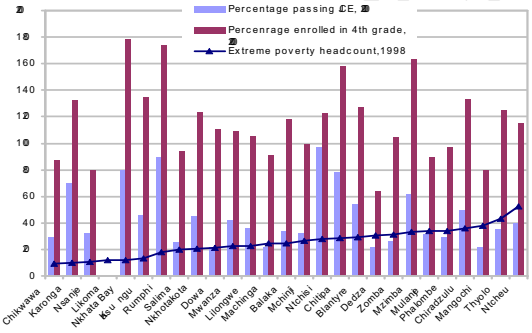
1. Sector & internal sector allocation → 2. Service standard & use of services → 3. Outcome & status → 4. Poverty reduction and other end goal impact from health, education etc.

Figure 4.4 Outcome - impact in the education sector, sorted by poverty headcount, 2000/ 1998.



Statistisk sentralbyrå

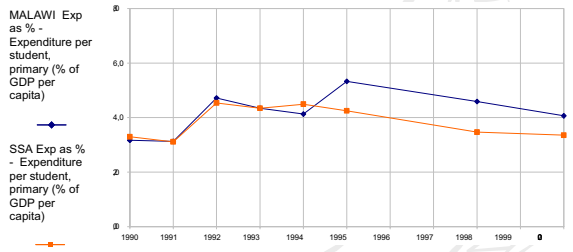
Figure 4.5 Outcome - impact in the educ. sector, sorted by extreme pov. headcount, 2000/1998.



Statistisk sentralbyrå

Regional data for each level. Input to the education sector.

Figure 5.1. Expenditure per student, Malawi and Sub Sahara Africa 1990-2000.



Statistisk sentralbyrå

Regional data for each level. Output in the education sector.

Figure 5.2 Gross enrollment, Malawi and Sub Sahara Africa, 1990-2000.

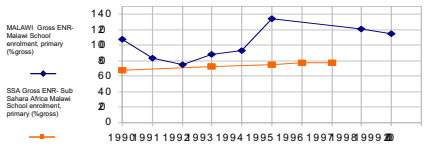
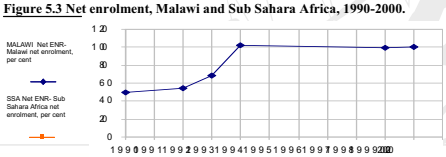


Figure 5.3 Net enrolment, Malawi and Sub Sahara Africa, 1990-2000.



Statistisk sentralbyrå

Regional data for each level. Outcome in the education sector.

Figure 5.4 Pupil-teacher ratio, Malawi and Sub Sahara Africa 1990-2000.

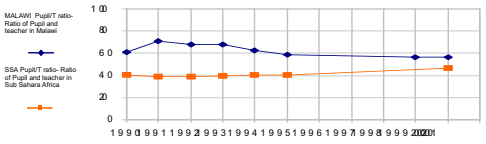
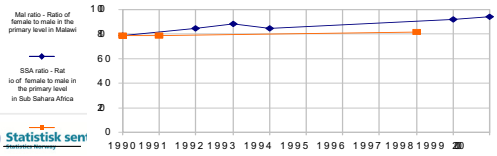


Figure 5.5 Ratio of female to male in the primary level, Malawi & Sub Sahara Africa, 1990-2000



Statistisk sen

31

WORKSHOP ON 'LIGHT' CORE MONITORING SURVEYS

Hotel Cardoso, Maputo, Mozambique, 12th - 14th December 2005

1. **What is Keninfo?**
   It is a database system that helps to organize and present social economic indicators. The database consists of the following elements:
   - Indicators
   - Time Periods
   - Geographic Areas
   - Subpopulations
     - Gender
     - Urban / Rural
     - Age groups
     - Total
   - Units
   - Sources
   - Data values

   The system provides access to indicators organized by sectors, goals, themes and other data management schemes. It is a user friendly software application integrated with Microsoft Office for easy presentation of data in tables, graghs and maps.

2. **History of Keninfo**
   Started by UNICEF as Childinfo 3.5 to monitor children progress and then adopted as indicators monitoring tool. After customizing the Kenyan version of Childinfo, we came up with KenInfo 1.0. Currently, it has been upgraded to DevInfo 4.0 and also upgraded the Kenyan version to KENINFO 4.0. Plans are underway to release a web and desk top version of KenInfo 5.0 in line with the release of DevInfo 5.0 after February 2006.

3. **Where are we?**
   The KenInfo 1.0 database was launched on May 2004. KENINFO 4.0 will help monitor Millennium Development Goals as well as the Economic Recovery Strategy for Wealth and Employment Creation goals. Currently, the database is being updated and being made more MDG compliant.

   KENINFO 4.0 has **50 indicators** cutting across these sectors;
   - Communications
   - Demography
   - Economy
   - Education
   - Environment
   - Health
   - Nutrition
   - Protection

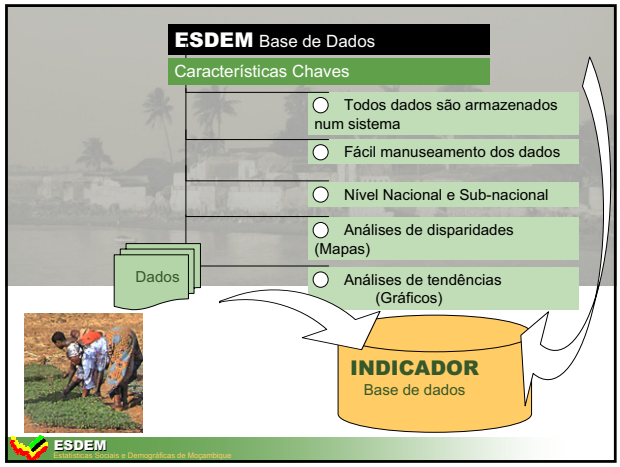
4. **Who controls what goes in**
   This has been guided to a large extent by the desire to track the MDGs and the ERSWC at the country level. The development of the database has been guided by a Technical Working Group comprising various UN agencies (UNICEF, UNDP, UNIFEM, UNFPA), line ministries (MOH, MOEST, MPND and Ministry of Water) and chaired by the CBS.

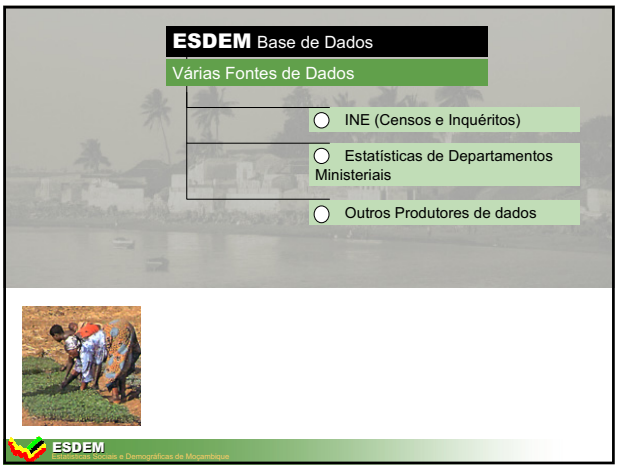
5. **How do we disseminate**
   This has been done through training workshops

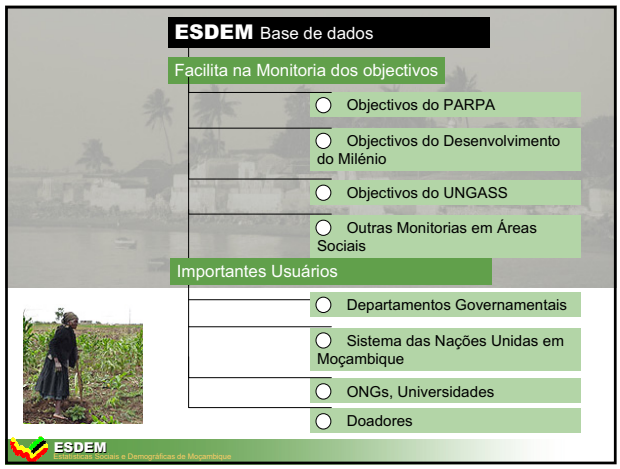
List of Indicators

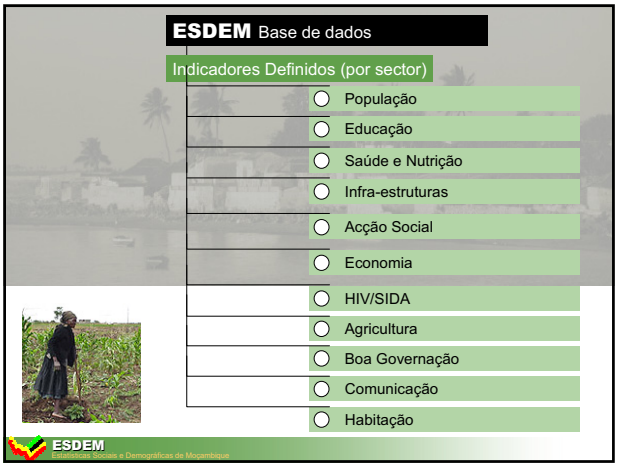
Sector	Class	Indicator
Communication	Mass Media	Daily average newspaper circulation
Communication	Mass Media	Population who own radio sets
Communication	Mass Media	Population who own television sets
Demography	Population	Crude birth rate
Demography	Population	Crude death rate
Demography	Population	Population growth rate
Demography	Population	Population size
Demography	Population	Population size as % of total population
Economy	Agriculture	Agriculture GDP (sector shares)
Economy	Agriculture	Agriculture growth (real GDP)
Economy	Income	Number of visitors arrivals
Economy	Income	Poverty index, poverty incidences
Economy	Production	Domestic debt/GDP
Economy	Production	External debt service as a percentage of exports of goods and services
Economy	Production	GDP per capita
Economy	Production	GNP per capital average annual growth rate
Economy	Production	Inflation rate, annual
Economy	Production	Real GDP growth
Economy	Production	Real output growth in manufacturing production
Economy	Production	Total expenditure/GDP
Economy	Production	Unemployment rate, 15-64 years
Economy	Production	Wages and salaries/GDP
Education	Early Childhood Care Development	ECD gross enrolment ratio
Education	ECCD	ECD gross enrolment
Education	Literacy	Literacy rates
Education	Primary Education	Primary school completion rate
Education	Primary Education	Primary school drop out rate
Education	Primary Education	Primary school gross enrolment females as a % of males
Education	Primary Education	Primary school net attendance rate
Education	Primary Education	Primary school net enrolment ratio
Education	Primary Education	Primary school transition rates
Education	Primary Education	Pupil per trained teacher ratio
Education	Primary Education	Pupil teacher ratio
Education	Secondary Education	Secondary school gross enrolment
Education	Secondary Education	Secondary school gross enrolment, females as a % of males
Environment	Sanitation	Access to safe means of excreta disposal
Environment	Water	Access to safe water
Health	Child Survival	Infant mortality rate (IMR)
Health	Diseases	Households with insecticide treated net (ITN) mosquito nets
Health	HIV/AIDS	Condom use at last high risk sex
Health	HIV/AIDS	Population tested for HIV/AIDS test
Health	HIV/AIDS	Women aged 15-49 years who know where to get an HIV/AIDS test
Health	Immunisation	DPT3 immunisation
Health	Immunisation	Measles immunisation
Health	Prevention	Domestic violence, physical/sexual violence, ever
Health	Prevention	Domestic violence, sexual violence, ever
Health	Safe Motherhood	Contraceptive prevalence rate
Nutrition	Iodine Deficiency Disorders	Households consuming iodised salt
Protection	Children	Proportion of children whose births are reported registered
Protection	Women	Female genital mutilation

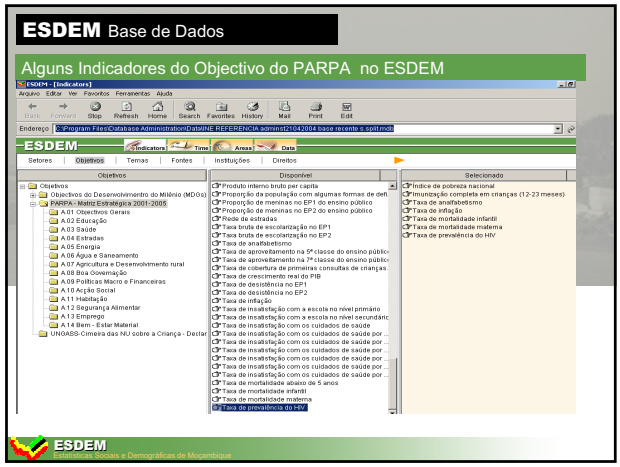
14.23. DevInfo-Mozambique





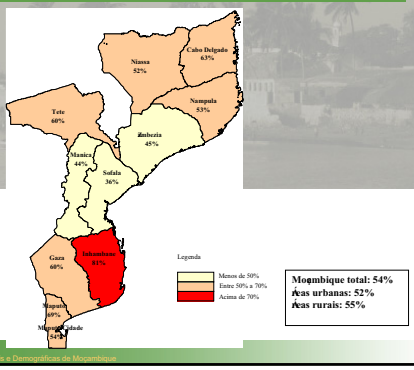






ESDEM Base de Dados

Análise de Disparidades - Linha de pobreza 2003

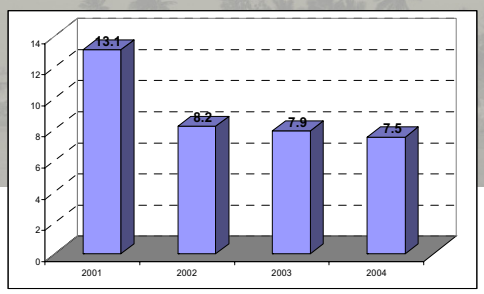


Fonte: INE/MPF

ESDEM Estatísticas Sociais e Demográficas de Moçambique

ESDEM Base de Dados

Análise de Disparidades – Taxa de Crescimento real do PIB



Fonte:MOZ_INE, Contas Nacionais

ESDEM Estatísticas Sociais e Demográficas de Moçambique

ESDEM Base de Dados

Análise de Disparidades – Taxa de Prevalência do HIV (15-49 anos)

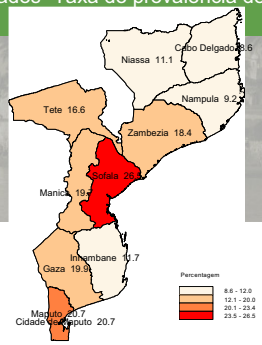
Nome da Area	2002	2004
Mocambique	**13.6**	**16.2**
Niassa	11.1	11.1
Cabo Delgado	7.5	8.6
Nampula	8.1	9.2
Zambezia	12.5	18.4
Tete	14.2	16.6
Manica	19	19.7
Sofala	26.5	26.5
Inhambane	8.6	11.7
Gaza	16.4	19.9
Maputo	17.4	20.7
Cidade de Maputo	17.3	20.7

Fonte:MISAU-RONDA 2004, 2005

ESDEM Estatísticas Sociais e Demográficas de Moçambique

ESDEM Base de Dados

Análise de Disparidades- Taxa de prevalência de HIV -2004



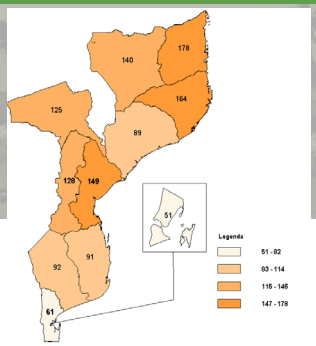
Fonte:MISAU-RONDA 2004, 2005

Taxa de Prevalência Nacional entre Adultos em 2004: 16.2%

ESDEM Estatísticas Sociais e Demográficas de Moçambique

ESDEM Base de Dados

Análise de Disparidades- Taxa de Mortalidade Infantil, 2003

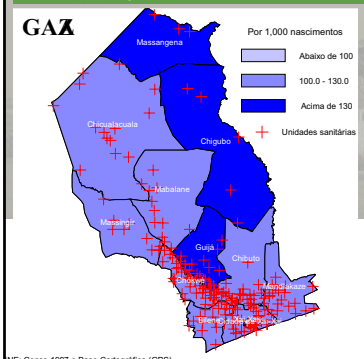


Fonte: IDS-2003

ESDEM Estatísticas Sociais e Demográficas de Moçambique

ESDEM Base de Dados

Análise de Disparidades - Taxa de mortalidade infantil 1997

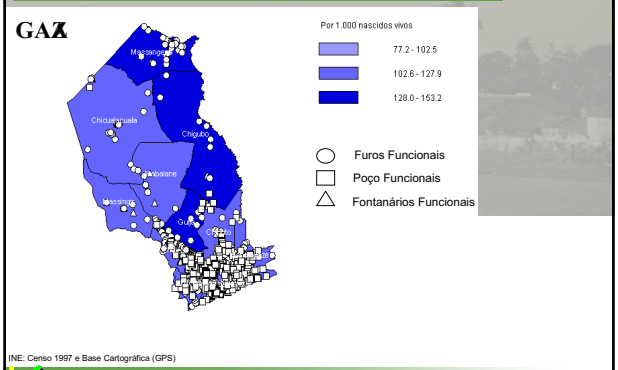


INE: Censo 1997 e Base Cartográfica (GPS)

ESDEM Estatísticas Sociais e Demográficas de Moçambique

ESDEM Base de Dados

Análise de Disparidades - Taxa de mortalidade infantil 1997

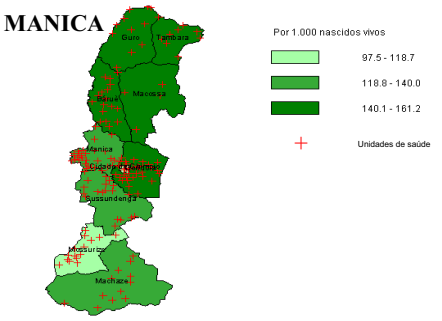


INE: Censo 1997 e Base Cartográfica (GPS)

ESDEM

ESDEM Base de Dados

Análise de Disparidades - Taxa de mortalidade infantil 1997

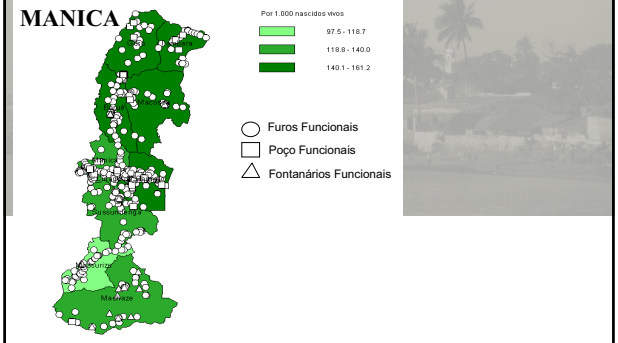


INE: Censo 1997 e Base Cartográfica (GPS)

ESDEM

ESDEM Base de Dados

Análise de Disparidades - Taxa de mortalidade infantil 1997

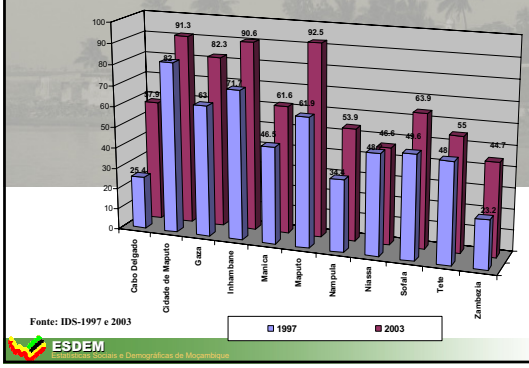


INE: Censo 1997 e Base Cartográfica (GPS)

ESDEM

ESDEM Base de Dados

Análise de Disparidades - Imunização Completa em Crianças (12-23 meses)

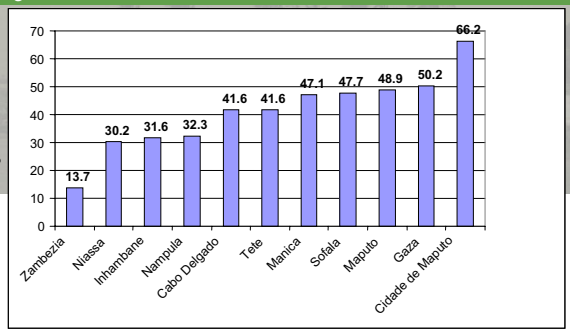


Fonte: IDS-1997 e 2003

ESDEM

ESDEM Base de Dados

Análise de disparidades – População com acesso a água segura, 2003

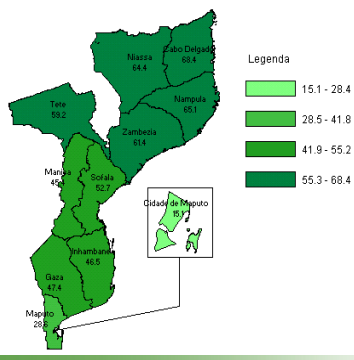


Fonte: IAF-2002/03

ESDEM

ESDEM Base de Dados

Análise de Disparidades - Taxa de analfabetismo, 2003

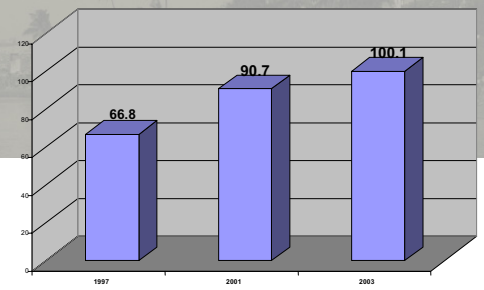


Fonte: IAF 2002/03

ESDEM

ESDEM Base de Dados

Análise de Disparidades – Taxa bruta de escolarização no ensino primário



Fonte: INE, Censo 97, QUIBB 2001 e IAF 2002/03

ESDEM
Estudos das Sociais e Demográficas de Moçambique

Muito Obrigado

ESDEM
Estudos das Sociais e Demográficas de Moçambique

14.24. Programmatic Approaches to Planning and Implementing Statistical Programs

Programmatic Approaches to Planning and Implementing Statistical Programs

Louise Fox
Lead Economist
The World Bank

The Results- Based Environment

- PRSP (PARP) is results-based Government program
  - End of program targets should be impact on beneficiaries: households, firms, farms
- Requires programmatic approach to data collection based on the needs of users:
  - Government ministries and agencies formulating and monitoring policy
  - Other users assessing Government programs and policies

Implications for Statistical Agencies

- Users need reliable, relevant, accessible data on behavior of households, farms, firms regularly
- Users need to be aware of costs of indicators
- Statistical agencies and PRSP agencies should agree on monitoring plan, including:
  - What data would be collected with what frequency
  - Cost (budget) and financing
  - Accessibility arrangements
- Donors committed to supporting this effort

Tidligere utgitt på emneområdet

Previously issued on the subject

Rapporter (RAPP)

2004/20: Wold, Bjørn K., Stein Opdahl, Estrellita Rauan, Randi Johannessen and Ingvar T. Olsen: Tracking Resource and Policy Impact. Incorporating Millennium Development Goals & Indicators and Poverty Reduction Strategy Paper monitoring across sectors

Documents

2004/17: Wold, Bjørn K., Dag Roll-Hansen, Astrid Mathiassen and Stein Opdahl: A Sustainable Household Survey Based Poverty Monitoring System A Poverty Monitoring System Based upon Household Survey Estimation of Total Consumption. A Preliminary Paper Asking for Cooperation

2002/8: Wold, Bjørn K., Olsen, Ingvar and Opdahl, Stein: Basic Social Policy Data. Basic Data to Monitor status & Intended Policy Effects with Focus on Social Sectors incorporating Millennium Development Goals

Discussion Papers

415: Mathiassen, Astrid: A Statistical Model for Simple, Fast and Reliable Measurement of Poverty

Økonomiske analyser (ØA)

2/2005: Wold, Bjørn K.: Fra ressursinnsats til velferdsvirkninger

Recent publications in the series Documents

2004/10 A-K. Mevik: Uncertainty in the Norwegian Business Tendency Survey.

2004/11 A.G. Hustoft, J. Linnerud and H.V. Sæbø: Quality and metadata in Statistics Norway.

2004/12 E. Engelien, R. Klæboe and Margrete Steinnes: Neighbourhood sonoscapes. Context sensitive noise impact mapping .

2004/13 P. V. Hansen: Regional electricity spot price responses in Norway.

2004/14 A.G. Hustoft and J. Linnerud: Development of a variables documentation system in Statistics Norway. International Statistical Conference "Investment in the future," Prague, Czech Republic, 6-7 September 2004.

2004/15 J.L. Hass: Compilation of data on expenditure in Environmental protection by businesses. Report to the European Commission DG for Environment.

2004/16 A. Raknerud, J. Rønningen og T. Skjerpen: Documentation of the capital database. A database with data for tagible fixed assets and economic data at the firm level.

2004/17 B.K. Wold D. Roll-Hansen A. Mathiassen and S. Opdahl: A Sustainable Household Survey Based Poverty Monitoring System. A Poverty Monitoring System Based upon Household Survey Estimation of Total Consumption. A Preliminary Paper Asking for Cooperation

2004/18 T. Karlsen, D. Quang Pham and T. Skjerpen: Seasonal adjustment and smoothing of manufacturing investments series from the quartely Norwegian national accounts

2005/1 V. Skirbekk: The Impact of a Lower School Leaving Age and a Later Retirement on the Financing of the Norwegian Public Pension System.

2005/2 H. Utne: The Population and Housing Censushandbook 2001.

2005/3 J. L.Hass and R. Straumann: Environmental Protection Expenditure: Methodological work for the Oil and Gas Extraction Industry. Report to Eurostat.

2005/4 L. Hobbelstad Simpson: National Accounts Supply and Use Tables (SUT) in Constant Prices SNA-NT "SUTCONSTANT"

2005/5 L. Hobbelstad Simpson: National Accounts Supply and Use Tables (SUT) in Current Prices. SNA-NT "SUTSTARTER"

2005/6 S. Todsen: SNA-NT User's Guide for Supply and Use Tables in Current and Constant Prices.

2005/7 E. Ugreninov, T.M. Normann and A. Andersen: Intermediate Quality Report EU-SILC 2003 Statistics Norway.

2005/8 H.V. Sæbø: Metadata strategy in Statistics Norway. Eurostat Metadata Working Group Luxembourg, 6-7 June 2005.

2005/9 J.L. Hass, K.Ø. Sørensen, K. Erlandsen and T. Smith: Norwegian Economic and Environment Accounts (NOREEA). Project Report 2002.

2005/10 A. Benedictow and T. Harding: Modeling Norwegian balances of financial capital.

2005/11 A.L. Mathiassen,J.B Musoke, P.Opio and P. Schøning: Energy and Poverty A feasibility study on statistics on access and use of energy in Uganda.

2005/12 E. Nyju, R. Strauman, Ø. Skullerud, J. Hass and B. K. Frøyen: Statistics on pre-treatment of waste. Pilot study - Norway 2004. Report to Eurostat

2005/13 H. Skullerud, Ø. Skullerud and S. Homstvedt: Pilot study: Treatment of Hazardous Waste. Final report to Eurostat.

2005/14 H. Skiri, B. Strand, M. Talka and H. Brunborg: Selected Documents on the modernisation of the Civil Registration System in Albania Vol. II.

2006/1 O. Andersen og M. Macura: Evaluation of theproject "Modernisation of the Civil Registration System in Albania"

2006/2 T. Åvistland: The problem with a risk premium in a non-stochastic CGE model.

2006/3 Å. Cappelen, R. Choudhury and T. Harding: A small macroeconomic model for Malawi.

2006/4 J. Ramm og A. Sundvoll: Translating and Testing the European Health Status Module in Norway, 2005.

2006/5 A.G. Hustoft og J. Linnerud: Statistical Metadata in Statistics Norway. 10s.

2006/6 H.V. Sæbø Systematic Quality Work in Official Statistics - Theory and Practice

2006/7 H. Skullerud: Methane emissions from Norwegian landfills Revised calculations for waste landfilled 1945-2004. 15s.

2006/8 R. Choudhury: User's Guide for a Macroeconomic Model for Malawi. 77s.