USAID
FROM THE AMERICAN PEOPLE

DEVELOPING & TESTING POVERTY ASSESSMENT TOOLS

RESULTS FROM ACCURACY TESTS IN UGANDA

NOVEMBER 2005

This publication was produced for review by the United States Agency for International Development. It was prepared by the IRIS Center, at the University of Maryland.

DEVELOPING & TESTING POVERTY ASSESSMENT TOOLS

RESULTS FROM ACCURACY TESTS IN UGANDA

SUBMITTED BY

The IRIS Center, University Research Corporation International

SUBMITTED TO

Scott Kleinberg

CONTRACT NUMBER

GEG-I-02-02-00029-00
Task Order No. 2: Developing Poverty Assessment Tools

AUTHORS

At the time of this writing, Manfred Zeller, Ph.D., was a professor for Socioeconomics of Rural Development at the Georg-August University of Göttingen's Institute of Rural Development in Germany. He is currently Professor for Rural Development Theory and Policy at the University of Hohenheim, Germany.

Gabriela Alcaraz V. is a Ph.D. candidate at the Institute of Rural Development, Georg-August University of Göttingen, Germany.

SERIES INFORMATION

This series presents a number of technical reports about USAID's program to develop Poverty Assessment Tools, implemented by the IRIS Center at the University of Maryland. These reports are made available to microenterprise practitioners, donor agency representatives, researchers and various stakeholders interested in the development of accurate and user-friendly tools to assess the poverty level of poor microenterprise clients in developing and transition economies.

CONTACT INFORMATION

IRIS Center
University of Maryland
Department of Economics
2105 Morrill Hall
College Park, MD 20742
USA

E-mail: info@iris.econ.umd.edu
Phone: +1.301.405.3110
Fax: +1.301.405.3020
Web: www.iris.umd.edu

ABSTRACT

This report presents the results of poverty assessment tool accuracy tests conducted by IRIS in Uganda in 2004. The report first describes the design of the field research and the computation of the applicable poverty line, followed by an overview of the analytical methods chosen. A number of increasingly complex econometric methods are used to increase the accuracy of the estimation. This analysis results in the identification of groups of indicators that identify the poverty status of the 800 households included in the sample.

For more information on the project, please visit www.povertytools.org.

For more information on AMAP and related publications, please visit www.microLINKS.org.

For more information about the IRIS Center, visit *www.iris.umd.edu*.

CONTENTS

1 INTRODUCTION

USAID commissioned the IRIS Center to develop, test, and disseminate poverty assessment tools which meet U.S. Congressional requirements for accuracy and cost of implementation. Accuracy tests of poverty indicators have been implemented by IRIS in Bangladesh, Peru, Uganda, and Kazakhstan. Comprehensive information on the project is available at www.povertytools.org, and will not be summarized in this report.

This report presents the results of poverty assessment tool accuracy tests conducted in Uganda in 2004.[1] Chapter 1 (this Introduction) provides an overview of the design of the field research for the accuracy test, and the computation of the applicable poverty line. Chapter 2 provides an overview of the analysis. In Chapter 3, we present the results on selected poverty indicators from eight regression models (known as "Ordinary Least Squares," or OLS, regression models). Each of these models potentially represents a newly designed poverty assessment tool, calibrated for Uganda. The regression models are run in the econometric package SAS, using the maximum R2 improvement technique (MAXR) procedure that seeks to maximize the explained variance of the dependent variable (per capita daily expenditure) by a set of the best 5, 10, and 15 regressors (referred to in this report as BEST5, BEST10, and BEST15). Any set of five, ten, or fifteen poverty indicators can be considered a poverty assessment tool for purposes of identifying the poverty status of a household.

The first five regression models differ with respect to the set of poverty indicators allowed in the model, starting from a model with a full set of potential regressors and gradually restricting the set of regressors on the basis of implementation practicality. A sixth model is run as an example of a tool that considers only those poverty indicators that were rated as "highly verifiable" by Nkoola Institutional Development Associates (NIDA), the survey firm in Uganda.[2] A subsequent model compiles these indicators with five powerful subjective and two monetary indicators. Finally, the last model makes use of poverty indicators usually available in the World Bank's Living Standards Measurement (LSMS) surveys. Thus, the first seven models can be considered alternative best combinations of poverty indicators which were mainly derived from existing practitioner tools for poverty assessment, while the last model is a tool derived from poverty indicators usually available in LSMS surveys.

Chapter 4 presents results from an alternative estimation approach, the so-called "two-step" models. In addition to Ordinary Least Squares (OLS), we also test Quantile, Probit, and the so-called Linear Probability regression technique. Compared to the models presented in Chapter 3, the performance of models presented in Chapter 4 is overall much better. Finally, Chapter 5 summarizes the results.

1.1 FIELD SURVEY FOR ACCURACY TESTS IN UGANDA

NIDA carried out the survey and double entered the data using SPSS (Statistical Package for the Social Sciences) data entry software. (The questionnaires can be downloaded at *www.povertytools.org*.)

The composite, as well as the benchmark questionnaire, required adaptation to the country context. In the case of the composite questionnaire, for example, this meant including poverty indicators specifically for food consumption in Section E (see questions E151 through 159) or the type of agricultural equipment owned. Useful sources for the identification of country- and region-specific poverty indicators include reports by the Uganda Bureau of Statistics (UBOS) and the experience of NIDA's professional interviewers, who helped adapt the questionnaire. The adaptation of the benchmark questionnaire mainly involved the selection of major food items. For this, we referred to results from the most recent (2002) UBOS Household Expenditure Survey.

The adaptation and processing of the two questionnaires benefited from the expertise of NIDA personnel: Godfrey Kajobyo (Project Director); Catherine Anena, Peter Fuuna, and William Nanyenya Ntege (Senior Field Supervisors); as well as NIDA's interviewers. Dr. Charity Irungu,[3] a Kenyan national, advised and supported the adaptation of the questionnaires, training of the interviewers, pre-test of the questionnaire, final adaptation and translation of the questionnaires, and the data entry and initial cleaning.

The training of the interviewers began in mid-July, 2004. Each interviewer went through a three-day, in-house training session to become familiar with the questionnaire and data-collection procedures. In-house training was followed by two days of questionnaire pre-testing.

Data collection was conducted between August 4 and October 25, 2004. A team of 21 enumerators and three supervisors participated in the survey. The team was sub-divided into three main sub-teams: one for the central and western regions, a second team for the northern region, and a third team for the eastern region. Double entry of all data was completed by end of November, 2004.

In total, 25 of the 56 districts in all four regions of Uganda were visited by three teams who implemented the composite questionnaire survey with 800 households, followed two weeks later by the benchmark questionnaire. The survey teams were composed of interviewers that were originally from the survey areas and spoke the local languages well. After final adaptation incorporating the results from the pre-test, the questionnaires were translated into five main local languages: Luganda, Luo, Runyakitara, Lusoga, and Ateso.

1.2 SAMPLING

1.2.1 REQUIREMENTS FOR SAMPLING

In view of logistical and budget constraints, a sample size of 800 households was chosen. The sample was required to be nationally representative. It included both the urban and the rural areas, proportionate to their respective shares in the nation's population. About 12 percent of the population resides in areas categorized as urban by the Government of Uganda.

The highest administrative units in Uganda are regions (Central, Eastern, Northern, and Western). The four regions are further disaggregated into 56 districts. District is further sub-divided into counties and sub-counties, with an average 17 sub-counties per district. Each sub-county holds a number of parishes. A parish consists of several local councils that represent the lowest administrative level in Uganda. Within each sub-county, UBOS had identified "enumeration areas," which typically correspond to the designation of local councils. UBOS uses these enumeration areas for drawing nationally representative samples.

At the time of the survey, some districts in the North were not considered safe. To conduct field work in these areas, NIDA required that survey teams be accompanied by police or military personnel. However, such protection was impossible to organize and six districts had to be excluded from the sampling frame: Gulu, Kitgum, Kotido, Pader, Moroto, and Nakapiripirit.[4] For the same reason, two counties from each of three neighboring districts — Lira and Apac in the Northern region, and Katakwi in the Western region — were also excluded from the sampling frame. The excluded counties were Kwania and Oyam in the Apac district, Kapelebyong and Usuk in the Katakwi district, and Erute and Moroto in the Lira district. According to the 2002 provisional population census results (UBOS, 2002), Uganda has 24.75 million people. Our sampling frame included 21.76 million people, while 2.98 million (about 12.1% of the total population) reside in the excluded districts or counties. As the northern region is the poorest region in the country, the sampling frame and the derived sample can only be considered representative for a large part of Uganda and not for the entire country.

1.2.2 SELECTION OF COUNTIES AND ENUMERATION AREAS

A multi-stage cluster sampling approach was used to draw up a random sample of households. This sampling procedure draws successive samples at lower administrative units, a feature that is useful in Uganda since data on the size of the population are published only for the region, district, county, and sub-county levels. In order to minimize sampling errors, the first stage of the sampling was conducted at the sub-county level, as this is the lowest administrative level with centrally available and published population data. Because of logistical and budget constraints, and in order to obtain a reasonable geographical spread for the sample, it was determined to randomly select 25 sub-counties, where the probability of selecting a certain county was equal to its share of population in the country.

This type of sampling, called probability-proportionate-to-size (PPS), was repeated during the second stage. In each of the 25 sub-counties, an enumeration area (EA) was randomly selected, with a selection probability proportionate to the size of the enumeration area and the total population size in the sub-county. Data on enumeration areas was provided by UBOS based on the 2002 provisional population census results. The parish where the selected EA was located automatically became the sample parish. Out of the 25 EAs, 22 are rural and 3 are urban (according to the UBOS classification). The Government of Uganda designates the entire Kampala District, all municipalities, all district towns, all Town Councils and Town Boards, and some Trading centers as urban areas. The three urban EAs account for 12 percent of the sampled EAs, which corresponds to the 12.1 percent urban share of Uganda's total population. Annex A, Table 1 provides detailed information about the sample.

1.2.3 SELECTION OF INDIVIDUAL HOUSEHOLDS

To obtain a self-weighing sample, 32 households were randomly selected from each of the 25 enumeration areas. The selection of individual households in an EA was conducted by the supervisors and the enumerators prior to field data collection. Systematic random sampling was used, with a sampling frame based on Local Council 1 (LC1) household registers (obtained from executive committee members). Sampling was based on a sampling frame comprised of all households that were found in the village registers, which were either obtained from the village committee leadership or developed by the survey team with the help of the local council leadership. Recent changes in the number of households (for example, through in- or outmigration of households) were accounted for by updating the village registers, which entailed local council leadership about such demographic changes not yet included in the village register. Out of this sampling frame, thirty-two households were randomly selected; each household had the same probability of being chosen. In addition, four extra households were selected to replace any non-respondents from the original 32 households. In the event of a non-response — for example, due to the absence of household head during the survey — the first (or subsequent) household among the four extra households was used to replace the non-respondent household.

1.3 POVERTY LINE

The Microenterprise Results and Accountability Act of 2004 specifies to two alternative poverty lines in defining the "very poor." A household and all its members is counted as "very poor" if either

1. the household is "living on less than the equivalent of a dollar a day" — in other words, if the average daily consumption expenditure among household members is less than $1.08 per day at 1993 Purchasing Power Parity, the definition of "extreme poverty" under the Millennium Development Goals

or

2. the household is among the poorest 50 percent of the population living below the country's own national poverty line

The legislation suggests that Congress intends for the higher of these two alternative criteria in each country to determine the extreme poverty line, which divides "very poor" from "not-very poor" households.

The legislation thus identifies two alternative measures of extreme poverty, relating to two commonly used poverty lines:

- **National poverty line (A)** — the bottom 50 percent of those classified as poor by any national poverty line. In Uganda, the national poverty line is expressed in Ugandan shillings (Ug.Sh.), the local currency. However, due to the geographic diversity, there is no single national poverty line for the whole country. Instead, the national poverty line is disaggregated into regional poverty lines that reflect the level of consumer costs separately for eight regions of the country (Appleton et al., 1999).

- **International poverty line (B)** — $1 income per day per capita (equal to $1.08 per day in purchasing-power-parity dollars at 1993 prices).

Because the benchmark questionnaire (see *www.povertytools.org*) enumerates per-capita expenditures in current Ugandan shilling, we converted the national and international poverty line into Ugandan shilling values as of September 2004, the median month during which the benchmark survey took place. In the following discussion, we evaluate the level of these poverty lines in local currency for September 2004.[5]

1.3.1 INTERNATIONAL POVERTY LINE

The World Bank's World Development Indicator tables do not report poverty headcount indices using international poverty lines for Uganda for previous years, neither for the $1-per-day nor for the $2-per-day measure (World Bank, 2005). For Uganda, there are two different estimates of purchasing power parity (PPP). The PPP rate estimated by the World Bank is 410.15 Ug.Sh., equivalent to $1 at January 2003 prices (259.97 Ug.Sh. for 1993). Inflating this value by the inflation rate during the period from January 2003 to September 2004 yields a purchasing power parity rate of 482.22 Uganda shillings per $1.08.[6] Hester et al. (2002) uses a different methodology to estimate PPP rates. These rates are published in the Penn World Tables.[7] In 1992, according to the Penn World Data (Version 5.6), the purchasing power parity rate of $1 is 351.76 Ug.Sh.. Using the information on national Consumer Price Index indices for mid-1992 to September 2004, this amounts to 615.72 Ug.Sh. per $1 (or 664.98 Ug.Sh. per $1.08).[8]

As pointed out by Karshenas (2004) and Reddy and Pogge (2003), differences in PPP rates are an important cause of differing estimates of poverty across countries. To determine the PPP rate to use for Uganda, we used a recent World Bank research paper by Chen and Ravallion (2004), who chose the higher PPP rate from the Penn World Tables (Version 5.6) to conduct their analysis, instead of the lower rate calculated by the World Bank. Chen and Ravallion's choice led to a higher poverty rate for Uganda compared to the one obtained with the PPP rate of the World Bank's Data Development Group (see Ahmad, 2003). Hence, we also chose the PPP rate provided by Penn World Data (Version 5.6), with a $1-per-day poverty line equivalent to 664.98 Ugandan shilling for September 2004. Correspondingly, the $2-per-day poverty line ($2.15) for the month of October is 1323.78 Ug.Sh. per day.[9]

At the market exchange rate, $1 was worth about 1850 Ug.Sh. in September 2004. Using this rate, the $1 international poverty line would correspond to 36¢, whereas the $2 line would be equivalent to 72¢.

1.3.2 NATIONAL POVERTY LINE

UBOS has defined a poverty line, expressed in Ug.Sh., per adult equivalent and month. The computation of an adult equivalent in Uganda follows the recommendations by the World Health Organization (WHO), and is described in detail by Appleton et al. (1999, see Annex A, Table 2). The poverty line is computed separately for four regions in view of the regional differences in living costs. The Central, Eastern,

Western, and Northern regions are each divided further into an urban and a rural part (see Annex A, Table 3.2).

Poverty declined during the 1990s in Uganda, due to good macroeconomic performance and broad-based economic growth (Deininger & Okidi, 2003). The poverty headcount indices dropped from 55.7 percent in 1992/93 to 33.8 percent in 1999/2000, as shown in Table 1.3.1. Economic growth continued after the turn of the century. Yet, several authors (Appleton & Ssewanjana, 2003; Ssewanjana et al. 2004; Kappel et al., 2005) conclude that the recent economic growth — between 1999/2000 and 2002/2003 — has favored mostly the wealthy in urban areas and led to rising inequality (Ssewanjana et al., 2004; Kappel et al., 2005). The data source for the poverty estimates is the Uganda National Household Survey (UNHS), which has been carried out by the Uganda Bureau of Statistics in various years.

Table 1.3.1 — Percentage of Ugandan population below the national poverty line, by year of national survey

Region	1992/93	1996/97	1999/00	2002/03
National	55.7	44.0	33.8	37.7
Urban	27.6	16.2	9.6	12.2
Rural	59.7	48.2	37.4	41.7

Source: Appleton et al. (1999) for year 1996/97, Kappel et al. (2005) for all other years.

Table 1.3.2 gives an overview of the calculated percentage of poor households in the IRIS sample for alternative definitions of poverty.

Table 1.3.2 — Percentage of poor households and people in September 2004, by poverty line

Type of poverty line	Poverty rate (% of households)	Poverty rate (% of people)
National poverty line (differentiated by eight regions and expressed per adult equivalent following WHO definition)	31.60	37.51
International: Below 1.08 PPP dollar per day per capita (664.98 Ug.Sh.)	32.36	38.84
International: Below 2.15 PPP dollar per day per capita (1323.80 Ug.Sh.)	67.51	76.26

Source: Own calculations from data of IRIS survey.

In the IRIS sample, 31.6 percent of households fall below the national poverty line and 32.36 percent below the $1-per-day line.[10] Hence, the applicable poverty line to define the "very poor" for Uganda is the international poverty line of $1. Here, the headcount index for the "very poor" (expressed as percent of people) is 32.36 percent. As poorer households tend to have higher household sizes, the poverty rate as a percentage of total population is higher than the one expressed as a percentage of households. In the IRIS sample, the poverty rate (using the national poverty line) of 37.51 percent of the population coincides well with the poverty rate of 37.7 percent measured with the UNHS 2002/03 survey.[11] This result suggests that poverty in Uganda has continued to increase during the past two years (from 2002 to 2004).[12]

To stay true to the language of the legislation, we will use the term "very poor" or "VP" for those households having an expenditure falling below $1 valued at purchasing power parity rate (i.e. $1.08 equivalent to 664.98 Ug.Sh. at purchasing power parity rate) and the term "not-very poor" or "NVP" for those having an expenditure equal to or above this cutoff value. Readers should bear in mind that *any* such binomial, either/or labels tend to distort the underlying reality, which is continuous — the standard of living of a household and its members just above the line is not that much different than that of a

household just below the line. It is also important to note that even among those above the international poverty line there exist a considerable share of households that are vulnerable to poverty such that, for example, losing one's employment or an illness of a family member may drive them into poverty.

1.4 POVERTY RATES AMONG CLIENTS OF FINANCIAL INSTITUTIONS

In the sample of 788 households used for the analysis presented in this report, only 40 households were current clients of financial institutions. This low frequency of clients does not allow for inferential analysis about the poverty status of clients differentiated by type of financial institution. Nevertheless, readers may find the descriptive results from the sample somewhat informative. Table 1.4.1 shows the percentage of households that have per-capita daily expenditures below alternative poverty lines. The results show that fewer households among clients fall below any of the three poverty lines. Fifteen percent of clients are very poor, compared to 32.36 percent in the total sample. Sixty-three percent of clients are poor if the $2 poverty concept is used.

Table 1.4.1 — Poverty status of client and non-client households, by type of financial institution

Type of institution/ client status	Very poor: Below $1/day poverty line	Below national poverty line	Below $2/day poverty line
Household is a client of a …			
Public or private bank (N=4)	0	0	25.0
Savings and credit cooperative N=4)	25.0	25.0	50.0
NGO (N=32)	12.5	15.63	68.75
All clients (N=40)	15.0	12.5	62.5
Non-client (N=748)	33.42	32.49	67.78
Total (N=788)	32.36	31.60	67.51

Source: Authors' computations from IRIS survey data

2 OVERVIEW OF REGRESSION ANALYSIS

2.1 INTRODUCTION

In Chapter 3, we analyze the accuracy of newly designed tools and develop eight regression models for generating tools.[13] These models consider all the poverty indicators that were compiled in the composite questionnaire, based on submissions of practitioner tools to IRIS in late 2003 that are reviewed by Zeller (2003) (see *www.povertytools.org*). In addition, indicators have been included based on recent poverty assessment studies published in academic literature. Thus, with the exception of Model 9 that uses LSMS type indicators only, the newly designed tools considered in Chapter 3 seek the best combinations of poverty indicators, drawn from existing practitioners' tools.

2.2 COMPOSITE QUESTIONNAIRE

The structure of the composite questionnaire is as follows (see www.povertytools.org):

A. Identification of household (location, client status, etc.)

B. Household roster/demography, including individual as well as household-level indicators (derived from all practitioner tools)

C. Household expenditures by category (adapted from FINCA and ACCION tool)

D. Housing indicators (generic questions adapted from tools by AIM, ASA, CASHPOR, CIMS-OI, PRIZMA, and TUP), plus poverty indicators concerning minimum wages acceptable to respondents

E. Food consumption/Food Security Scales (adapted from tools by CGAP, Freedom from Hunger, and World Food Program Food Security and Hunger Questionnaire)

F. Asset based indicators (adapted from GRAMEEN Network and most other tools)

G. Social capital, voice, and vulnerability (adapted from recent advancements in social science research)

H. Estimates of objective and subjective poverty (adapted from recent advancements in social science research)

I. Information on client status of individual household members in programs and institutions supporting micro-finance or business development services (including information on size of loans and outstanding debt)

K. Monetary voluntary savings by individual household members (WOCCU)

NIDA's survey staff observed that the questions in section J are fairly sensitive, and special emphasis was given to test this section during the pretest. It was then decided to retain this section as the last section of

the composite questionnaire, and ask the interviewers to introduce this section to respondents with the assurance of absolute confidentiality.

2.3 SELECTION OF INDICATORS

In Chapter 3, we present results from eight models that were run with Ordinary Least Squares using SAS software. The models differ by the type of regressors used. While Model 1 includes 250 regressors, Model 9 uses the most restrictive list, with 131 potential poverty indicators.

As one can see from the results for Model 1, the set of best poverty indicators is composed by a multidimensional combination of housing, community, expenditure, education, and demographic characteristics.[14] In Model 1, there are only a few poverty indicators from other dimensions and sections of the composite questionnaire. In a gradual process starting with Model 2, we reduce the number of regressors so as to allow indicators from other dimensions and sections of the questionnaire to enter among the best set of poverty indicators.

The overriding principle is to narrow down the list of poverty indicators with respect to two criteria.

- **Difficulty of indicator** — Information on some indicators is easy to obtain, while for others it is not. Difficulty can be expressed in terms of time, money, and social costs expended for obtaining information. Social costs are especially important when addressing culturally sensitive questions. The difficulty of an indicator will therefore vary with the socio-economic and cultural context. It will also depend on the skill level and quality of interviewer training. Furthermore, difficulty is strongly affected by the educational level and intellectual skills of the respondent, and by the interview situation (whether in private at home, or among peers and/or strangers in public — where certain types of questions may incur high social costs for the respondent). For example, the value of total assets is very difficult and tedious to obtain, and therefore is not really suitable for an operational poverty assessment tool. Another example is question C2 in the composite questionnaire, regarding the value of food that is home-produced and consumed by the household in an average week, as well as several other expenditure indicators.

- **Verifiability of indicator** — Another useful characteristic of an indicator for its operational use is its ease of verification (in terms of time, monetary, and social costs). Here, we distinguish between objective and subjective indicators. Subjective indicators include any self-assessment (perception, feeling, attitudes) by the respondent (e.g., Section E9 onwards and Section H, regarding perceived adequacy of livelihood); or any assessment done by the interviewer (e.g., rating the poverty status of a household on a scale from 1 to 5, as in Section A). While some subjective indicators are among the more powerful poverty indicators, as will be shown later, they are hardly verifiable, as the scales used are subjective and not disclosed to others. Objective indicators use measurement scales that can be — at least in principle — verified by consistent standards of measurement metrics. Examples of objective indicators include the age of a person (in years), the size of the rooms (in square meters), or whether the roof is of natural fibers; these are directly measurable through conventional and universally comparable scales. Measurability using comparable scales is a prerequisite for direct verifiability. Objective indicators, however, may also vary in their degree of verifiability. An example of an objective but unverifiable indicator is the number of luxury foods eaten in the past 7 days, or the amount of money received from migrant relatives, or how many days a child was sick in the past 12 months. Common to this group of unverifiable objective indicators is the fact that actions or states occurred in the past.

Because the difficulty and verifiability of an indicator cannot be generalized across different socio-economic and cultural contexts, it is somewhat arbitrary to classify a particular indicator (or group of indicators) as being more or less difficult to ask, or more or less verifiable. Therefore, the selection of

progressively smaller subsets of regressors for defining Models 1 through 6 inevitably reflects some arbitrary assumptions and choices.

In Model 7 and 8, we use the subjective assessment of verifiability by NIDA as an alternative attempt to address the practicality issue. To get more information about the practicality of poverty indicators, the Poverty Assessment Tools project will undertake practicality tests carried out by microfinance (MF) and business development services (BDS) organizations.

Our sequence of regression models using progressively fewer poverty indicators (from Model 1 to Model 6) aims to generate a series of poverty assessment tools that gradually become less accurate (because they use a less-comprehensive set of regressors) but at the same time become more practical, less costly, and less prone to falsification by respondents or survey intermediaries.

For each model presented in Chapter 3, we present a set of BEST5, BEST10, and BEST15 poverty indicators. Each of these three sets can be considered a poverty assessment tool in itself, and we document for each tool several measures of accuracy. From an operational point of view, and if all other factors are held constant, a tool derived only from the five best indicators presents an easier, more practical poverty assessment tool than one that uses the BEST15 (or even more) poverty indicators.[15] This is quite obvious — fewer questions must be asked and analyzed with a BEST5 tool than with a BEST15 tool. However, the inclusion of fewer poverty indicators in the tool also tends to imply a lower degree of accuracy.

This highlights the important trade-off between the accuracy and practicality of a poverty assessment tool. Achieving the right balance requires careful consideration between the trade-offs between accuracy and practicality, and this will ultimately determine the choice and certification of certain poverty assessment tools.

2.4 SPECIFICATION OF REGRESSION MODELS

In all regressions, the sample size is 788 because of missing benchmark surveys or missing values in 12 of the sample households. The dependent variable is the natural logarithm of per-capita daily expenditures in Ugandan shillings.

Table 2.4.1 — Dependent variable per capita daily expenditures

Variable	N	Minimum	Maximum	Mean	Standard deviation
Per capita daily expenditures	788	42.65	11545.66	1293.77	1167.71
Ln expenditures per capita (natural logarithm)	788	3.75	9.35	6.86	0.76

In all regressions, an INCLUDE statement always includes the following seven regressors as control variables:

Table 2.4.2 — Description of the seven control variables

Variable	N	Minimum	Maximum	Mean	Standard deviation
Household size	788	1	18	5.83	2.97
Household size squared	788	1	324	42.81	42.67
Age of household head	788	18	102	43.09	15.43
Western region	788	0	1	0.32	0.47

Northern region	788	0	1	0.12	0.33
Eastern region	788	0	1	0.27	0.45
Urban location	788	0	1	0.12	0.33

The first three control variables take into account the influence of important demographic factors that, in previous research, have been found to be powerful variables in explaining per-capita expenditures at the household level. The four dummy variables seek to capture regional agro-ecological, cultural, and socioeconomic differences between different regions as well as differences between rural and urban locations. The inclusion of these four dummy variables ensures that the estimated regression coefficients are controlled for regional differences.

All variables that are defined in monetary values, such as for expenditures and assets, are converted into natural logarithms,[16] including the dependent variable per-capita daily expenditures. All ordinal variables (type of roof, for example, with lower values indicating inferior materials and higher values indicating superior materials) have been converted into dummy variables that reflect the different subtypes. For example, if the database has three types of roof (1=natural material; 2=metal; and 3=superior, such as tile), then dummy variables for two of the three different types of roof were formulated and tested in the statistical analysis for their potential of being a significant poverty indicator.

The different models were run as Ordinary Least Squares in SAS using the MAXR procedure that seeks to obtain a model with a high R-square. The R-square (R2) is the ratio of the variance in the dependent variable that is explained by the model and its regressors, divided by the overall observed variance of the dependent variable. The coefficient ranges between 0 and 1. R2 takes on the value of 1 when predicted values for the dependent variable for all households are the same as the observed values. An R2 coefficient of 0.6 implies that 60 percent of the observed variance in the dependent variable is explained by the model and its regressors.

A model with a high explanatory power is a prerequisite for good predictions of the dependent variable per-capita daily expenditures (and thereby poverty status). The maximum R2 improvement technique (MAXR) is a subcommand for regressions in SAS. The MAXR technique seeks to maximize explained variance (i.e., R2), and considers all combinations among pairs of regressors to move from one step to the next. In the first step, the MAXR method begins by finding the one-variable model that produces the highest R2. In the second step, another variable, the one that yields the greatest increase in R2, is added. Once the two-variable model is obtained, each of the variables in the model is compared to each of the variables not in the model. For each comparison of single pairs of variables, MAXR demonstrates whether removing one variable and replacing it with the other one increases R2. After comparing all possible switches, MAXR makes the switch that produces the largest increase in R2. Comparisons then begin again in the third step and so forth, and the process continues until MAXR finds that no switch can increase R2. This limit may not be reached at 15 variables, but may include many more regressors. Thus, the MAXR technique allows us to identify the best model in each category — with only one, only five (the BEST5 model), only ten (the BEST10 model), or only fifteen (the BEST15 model) regressors.

2.5 DIFFERENCES BETWEEN THE MODELS

The following eight model types were run as Ordinary Least Squares in SAS. From the composite questionnaire, many alternative poverty indicators were computed. For the regression analysis, all poverty indicators being monetary variables in Ug.Sh. were replaced by their natural logarithms, and all poverty indicators derived from original ordinal or nominal variables in the composite questionnaire were converted into dummy variables. For example, the ordinal variable education (see section B of the composite questionnaire) was converted into dummy variables reflecting the achievement of different

levels of education for individual household members. In total, 633 regressors were contained in the basic regression data file and tested as potential poverty indicator. Similar to the subsequent analysis of the eight models, the SAS MAXR procedure (as explained in Chapter 2.4) was used to select the best 250 potential regressors (in addition to the seven control variables) for the regression models 1 to 9. All of the dimensions of poverty (as well as all submitted poverty indicators from practitioner tools) from the total number of composite questionnaire sections were represented in the final regression data file containing the best 250 indicators, and hence in the generation of the tools. Special care was given to the generation and testing of gender-specific poverty indicators. Annex C separately lists the gender-specific indicators that were selected for the final regression analysis (i.e., for the subset of 250 regressors).

Next, we describe the differences between the models (see also Figure 2.5.1).

- **Model 1** — Model 1 includes all 257 regressors considered in the regression analysis using SAS software (including seven control variables contained in every model).

- **Model 2** — In Model 2, we drop all expenditure-related variables except "daily per-capita clothing expenditures" and "other expenditures" (e.g. for social events, leisure etc.) in the past 12 months (see section C of the composite questionnaire). These variables were the two best expenditure categories among the 13 tested using the MAXR technique.[17] The third-best were education expenditures. A reduction of the number of expenditure variables is a first step towards a more operational set of poverty indicators. As pointed out, self-reported expenditures by respondents, irrespective of whether the recall period for expenditures was one week, one month, or one year, are impossible to verify directly.

  In general, questions contained in section C (question C1 to C12) require relatively intensive interviewer training, as they are prone to high measurement error. The interviewer needs to facilitate the interview by asking prompting questions on major elements of the different expenditure categories. Take, for example, home-produced food, a particularly difficult expenditure category. Questions related to home-produced food are especially difficult for interviewers unfamiliar with traditional (or metric) measures used for crop yields in agriculture and food subsistence production (see question C2). Furthermore, the interviewer needs to provide special assistance to respondents with no or low school education for even simple calculations, such as adding up expenses, especially since some elements of a certain expenditure category are recalled by the respondent on a monthly basis, and others are best remembered on a weekly basis (1 bag of potatoes per month, but a basket of rice per week). While these questions did not pose any significant difficulties for experienced NIDA interviewers, they may pose difficulties for less-experienced interviewers. In total, Model 2 has 244 regressors that were retained from Model 1 (see Annex B).

- **Model 3** — The set of regressors for Model 3 is similar to Model 2. The only difference is the exclusion of the "total value of household assets" variable as a regressor. This variable is the natural logarithm of the total value of all assets possessed by the household. The total asset value is a powerful poverty indicator, and its exclusion allows other variables for single assets (or subgroups) to enter among the best regressors. The variable has been calculated from the value of all assets (from section D and F of the composite questionnaire). This variable is considered costly, and therefore a less practical poverty indicator, since it would require asking many of the questions from section D and F.

  In spite of the relevance of the "total value of household assets" variable in the other three IRIS test countries, the SAS MAXR procedure did not identify it as one of the best regressors. This result led to the omission of Model 3, since the only difference between Model 3 and 4 is the inclusion of this variable. The insignificance of the "total value of household assets" variable in

the case of Uganda can be explained by measurement error. After the survey, a debriefing with NIDA interviewers revealed that the survey team reported that respondents had great difficulty in estimating values for major assets, especially agricultural land and houses. In comparison with the other three countries (i.e. Bangladesh, Kazakhstan, and Peru), the rural economy in Uganda — where most of the population live — is less commercialized. Many types of assets are home-produced and rarely purchased or sold. This lack of market transactions makes it more difficult for respondents to estimate the current value of assets, implying a more imprecise measurement of the variable being calculated as the sum of values for all assets possessed by the household.

- **Model 4** — The set of regressors for Model 4 is similar to Model 2. The only difference is the exclusion of the variables "daily per-capita clothing expenditures" and "other expenditures" (social events, leisure) in the past 12 months. As these were the most powerful poverty indicators among the expenditure group, their exclusion allows other poverty indicators to enter into the set of best regressors.

- **Model 5** — Model 5 is similar to Model 4, but all subjective poverty indicators are excluded. Such indicators include all ordinal rankings done by the interviewer (such as those at the beginning of the interview, in Section A, or the assessment of the structure of the house), as well as all ordinal rankings concerning the feelings or self-assessment of the respondent (for example, the ladder questions in Section H). While these subjective indicators can be powerful poverty indicators they cannot be verified, at least not directly. Thus, such indicators allow for strategic answers by the respondent, depending on her expectations of the interview outcomes. For example, if a respondent feels that appearing poorer will increase her chances of being accepted into a program or getting a loan, then she may strategically alter her response.[18] The subjective poverty indicators that were excluded in Model 5 (compared to Model 4) are presented in Annex B. Model 5 includes 181 variables.

- **Model 6** — Model 6 is similar to Model 5, but it excludes all monetary variables from the remaining subset of regressors. With this approach, we now base the model solely on demographic, housing, and community characteristics, as well as on the number and type of assets possessed. There are 147 variables in Model 6.

- **Model 7** — Compared to the previous models, Model 7 is more restrictive with respect to the criteria of verifiability, incorporating 172 indicators that NIDA rated (see Annex D) as "somewhat easy or easily verifiable."[19] The model contains many poverty indicators that are used in the housing index, variables on asset ownership, as well as other observable indicators.

- **Model 8** — Model 8 is similar to Model 7, but includes the two best expenditure variables (clothing and education expenditures) that received a fairly good score (4) for difficulty to ask (see Annex D). While the group "other expenditures" was the second-best in terms of accuracy, its practicability is judged to be low as the use of this expenditure group requires first asking about the prior 12 expenditure groups. Therefore, we replaced it by the next best expenditure group: education expenditures. In addition to the two expenditure variables, the following five powerful subjective variables are included:[20]

    Did you or any other adult in your household stop eating meals for an entire day because you didn't have enough money to buy food?

    Position on the ladder of a household with a monthly income of 70,000 Ug.Sh.

    Number of days in past 7 days with any of 4 superior foods eaten (max. 28)

    Amount household needs per month to live

Change in number of steps (ladder) from 7 years ago to today (if minus=worsened)

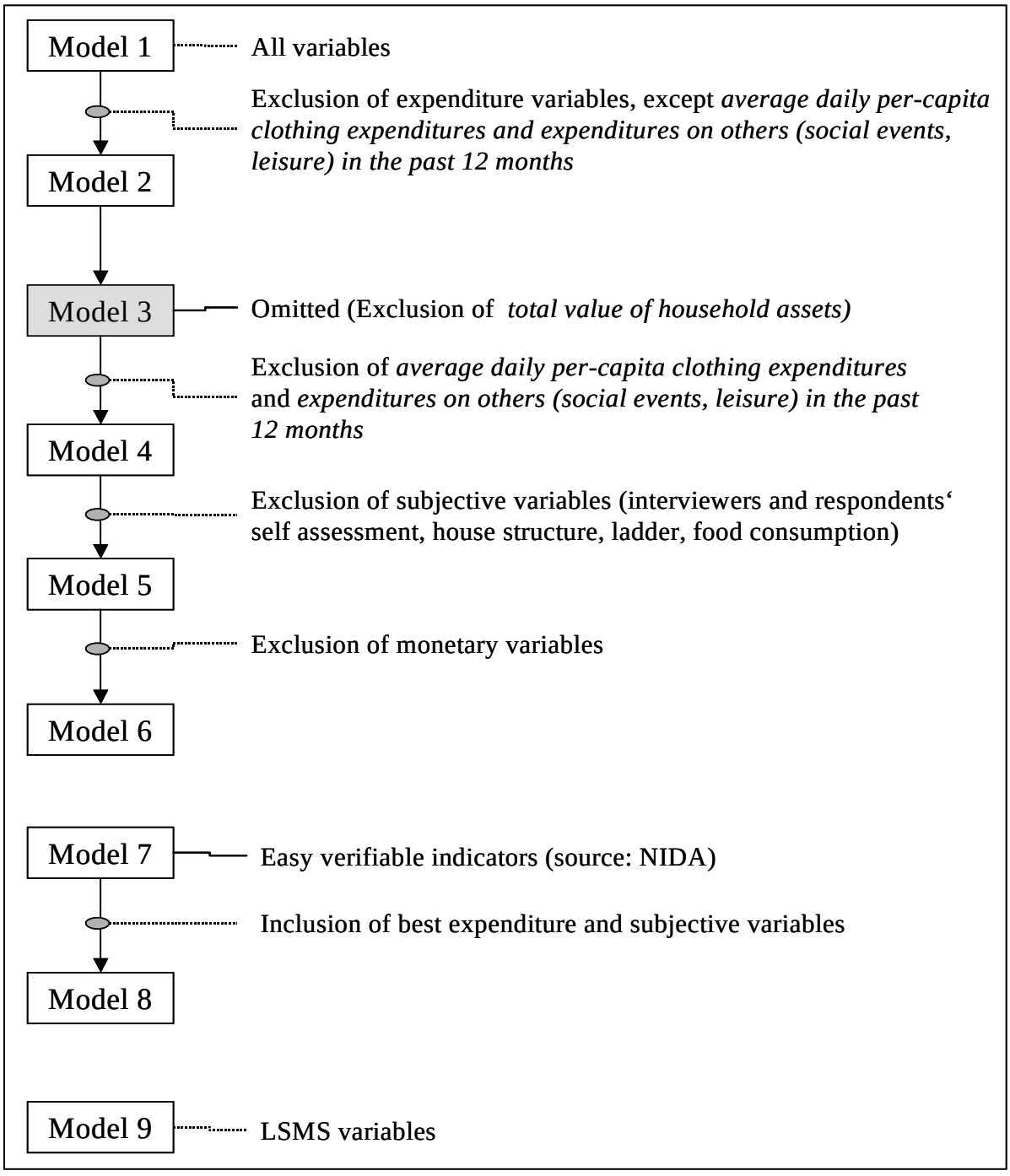
Model 8 is an example of a combination of indicators that are deemed easily verifiable by survey experts in Uganda (some of the indicators are directly observable) with powerful subjective and objective indicators that are not directly verifiable but fairly easy to ask (see Annex D). Moreover, this model (or poverty assessment tool) may allow indirect verifiability of the clothing and education expenditures and the subjective indicators through comparing these answers with the answers to the readily verifiable indicators.

- **Model 9** — Model 9 incorporates variables that are usually available in LSMS surveys. In addition to the seven control variables, it includes 124 regressors related to demographic, asset, expenditure, housing, and credit and financial asset information.

Annex B presents a description of the 257 regressors entered into the different models. For each model, the corresponding column (M*) indicates the specific regressors included in the model type. Figure 2.5.1 presents an overview of the eight regression models tested.

In conclusion, the models differ in the sets of poverty indicators being submitted to regression analysis. The result of the regression analysis, i.e. the identified set of best regressors (be it 5, 10, or 15) could potentially be used as a tool in nationally representative surveys in Uganda for assessing whether a household is below or above the poverty line. The models differ in the number and type of regressors that are considered, and models 1 to 7 represent increasingly simple tools that appear progressively less prone to the risk of strategic answers and verification problems.

Figure 2.5.1 — Schematic representation of the models' construction process.



3 RESULTS FROM SINGLE-STEP REGRESSION MODELS

In the following chapter, the results are summarized by listing

- Regressors that were among the BEST5, BEST10, and BEST15 models
- The adjusted R-square achieved (for example, an R-square of 0.6 indicates that 60 percent of the observed variance in the dependent variable is explained by the regressors)

It is important to note that the set of BEST regressors is statistically determined by the SAS MAXR technique, which searches for the best model fit. The term BEST should not be misunderstood as a value statement that implies as being best for the total accuracy of a regression model, or for any of the other measures of performance listed below. The set of BEST5, BEST10, or BEST15 regressors simply refers to the best model fit, given the constraints on the set of available regressors and on the maximum number of regressors for inclusion (for example five regressors in a BEST5 model).

In order to assess the prediction power of a regression model (or tool) for poverty assessment, we also present the following seven measures of accuracy performance for each model (IRIS, 2005):

- **Total Accuracy** — The percentage of the total sample of 788 households whose poverty status is correctly predicted by the regression model
- **Poverty Accuracy** — Accuracy among the very poor, which refers to the households correctly predicted as very poor, expressed as a percentage of the total very poor
- **Non-Poverty Accuracy** — The accuracy among the not-very poor, which refers to the households correctly predicted as not-very poor, expressed as percentage of the total number of not-very poor
- **Undercoverage** — The error of predicting very-poor households as being not-very poor, expressed as a percentage of the total number of very poor
- **Leakage** — The error of predicting not very-poor households as very poor, expressed as a percentage of the total number of very poor.
- **Poverty Incidence Error** (PIE) — The difference between the predicted and the actual (observed) poverty incidence, measured in percentage points
- **Balanced Poverty Accuracy Criterion** (BPAC) — Poverty Accuracy minus the absolute difference between Undercoverage and Leakage, each expressed as a percentage of the total number of very poor. When Undercoverage and Leakage are equal, the BPAC is equal to the Poverty Accuracy. BPAC is measured in percentage points. The application of the BPAC criteria is based on three assumptions about the valuation of errors concerning the predictions of the very poor and the not-very poor (IRIS, 2005).

Leakage and Undercoverage are often used in the literature to assess the poverty targeting performance of development and safety net policies, institutions, or projects. PIE indicates the precision of a model (or poverty assessment tool) in correctly predicting the poverty incidence. Ideally, the value of PIE should be zero, implying that the predicted poverty rate equals the observed poverty rate. Positive values of PIE indicate an underestimation of poverty, whereas negative values imply an overestimation of the poverty headcount index. PIE is useful if the objective of the poverty assessment is to measure the poverty outreach of an entire institution that provides microfinance or business development services. Hence, the evaluation question is: What percentage institution X's clients are very poor? Note that a good PIE value may be reached by a combination of a low Poverty Accuracy and a low Non-Poverty Accuracy. This is because errors in predicting the very poor may cancel out with errors made in predicting the not-very poor, and then still lead to a good PIE value. Thus, a good PIE value (close to zero) may be achieved by a model that has low Poverty Accuracy, low Non-Poverty Accuracy, and high Leakage and Undercoverage. Because of this, the selection of a model on the basis of PIE entails the risk of choosing a model with low Poverty Accuracy and high Undercoverage and Leakage. BPAC considers these three accuracy measures, and models with a higher positive value for BPAC indicate a higher Poverty Accuracy, adjusted by the absolute difference between Leakage and Undercoverage. There may exist trade-offs between PIE and BPAC in the selection of models. A perfect prediction model would have a value of zero for PIE and a value of 100 for BPAC. In such a perfect model, Leakage and Undercoverage would have a value of zero, and Total Accuracy, Poverty Accuracy, and Non-Poverty Accuracy a value of 100.

In section 3.1 through 3.9, we present results from models using the OLS regression technique applied to the different sets of regressors identified as Model 1 through 9. In section 3.10, we use three alternative regression techniques. For each of the regression models presented in this chapter, we provide the performance measures described above for predicting the very poor and not-very poor. These model performance measures are illustrated below, using the results from Model 1.

3.1 MODEL 1

Model 1 includes all 250 regressors available for the regression analysis, in addition to the seven control variables. Table 3.1.1 presents the number of households classified as very poor and not-very poor by the international poverty line, as well as the predicted poverty status of the households within both groups.

Table 3.1.1 — Observed vs. Predicted poverty status for the BEST15 regressors set

Poverty status (as measured by benchmark questionnaire in survey)	Predicted poverty status		
	Not-very poor	Very poor	Total
Not-very poor	461	72	533
Very poor	93	162	255
Total	**554**	**234**	**788**

Observed poverty status:

- Actual Poverty Incidence = (255 / 788) * 100 = 32.36%
- Actual Non-Poverty Incidence = (533 / 788) * 100 = 67.64%

Predicted poverty status:

- Predicted Poverty Incidence = (234 / 788) * 100 = 29.69%
- Predicted Non-Poverty Incidence = (554 / 788) * 100 = 70.31%

Model performance:

- Total Accuracy = (461 + 162) / 788) * 100 = 79.06%
- Poverty Accuracy = (162 / 255) * 100 = 63.53%
- Non-Poverty Accuracy = (461 / 533) * 100 = 86.49%
- Undercoverage = (93 / 255) * 100 = 36.47%
- Leakage = (72 / 255) * 100 = 28.23%
- Poverty incidence error (PIE) = -2.67 percentage points
- Balanced Poverty Accuracy Criterion (BPAC) = 63.53-abs(36.47-28.23)= 55.29 percentage points

From Table 3.1.2, it can be observed that the highest performance in terms of Total Accuracy of Model 1 is actually achieved in the BEST15 set (79.06%). Furthermore, the variables incorporated in each set present a balanced multidimensional nature on which community characteristics appear important already in the BEST5 set.

Model 1 achieves on average a Total Accuracy of 78 percent. Poverty Accuracy was on average 61.4 percent. With respect to the not-very poor, the accuracy in prediction is around 85 percent. Model 1 underestimates the Actual Poverty Incidence by 2.7 percentage points.

In comparison with the rest of the models presented in this chapter, Model 1 BEST15 achieved the highest levels of Total Accuracy and Poverty Accuracy, as well as the lowest Undercoverage figures. Nevertheless, this set achieved only the second best PIE and the second highest BPAC.

Model 1 has the highest R^2 of all models. However, as a precursor to the results of this chapter, all OLS models performed relatively similar in terms of the accuracy measures Undercoverage and Leakage. Considering the variables included in the BEST* sets, ten variables appeared at least in five out of the eight models, a situation which contributes to the similarity in model performance. Significant improvements in accuracy were achieved by using two-step models, as shown in Chapter 4.

In addition, the use of Model's 1 regressor set on a Linear Probability Model (LPM) presented higher accuracy results than the OLS alternative. The same pattern was observed for the rest of the models (see Annex E, Table E2).

In general, some of the selected indicators of Model 1 may not be optimal in terms of the difficulty of obtaining information or verifying the indicators. For example, the "annualized food expenditures," "annualized-non food expenditures" (education, health and other expenditures), and "number of days with any superior food eaten" indicators would require detailed questioning about the different expenditures or consumption incurred by the households, and they would not provide easily verifiable information.

Table 3.1.2 — Summary of accuracy results for Model 1

Variables	Model Performance (%, percentage pt)	
Best 5 indicators: R2 adjusted = 0.517		
Cooking fuel is charcoal or paraffin	Total Accuracy	76.39
Annualized non food expenditures 2 (education, health, and others)	Poverty Accuracy	58.82
Value of panga (bushknife/machete)	Non-Poverty Accuracy	84.80
Number of days in past 7 days any of 4 superior food eaten (max. 28)	Undercoverage	41.17
Do you have nursery/pre primary center in your community?	Leakage	31.76
	Pred. Pov. Incidence	29.31
	PIE	-3.05
	BPAC	49.40
Best 10 indicators: R2 adjusted = 0.561		
Next best five indicators:	Total Accuracy	78.04
Whether spouse owns any type of shoes	Poverty Accuracy	61.96
Are you or members of household denied service or do you/they have only limited opportunity to sanitation services?	Non-Poverty Accuracy	84.74
	Undercoverage	38.04
Annualized food expenditures, recall 1 week	Leakage	29.80
Maximum education level of any household member is post secondary/superior level (completed)	Pred. Pov. Incidence	29.69
	PIE	-2.67
Maximum education level of males is secondary/post primary/J1 level (completed)	BPAC	53.72
Best 15 indicators: R2 adjusted =0.587		
***Next best five indicators*:**	Total Accuracy	79.06
Roof with banana leaves, fiber, grass, bamboo or wood	Poverty Accuracy	63.52
Household never ate sweet potato because other food was scarce	Non-Poverty Accuracy	86.49
Total of household members in farmers group	Undercoverage	36.47
Do you have market in your community?	Leakage	28.23
Household head is a widow	Pred. Pov. Incidence	29.69
	PIE	-2.67
	BPAC	55.29

3.2 MODEL 2

This model excludes all expenditure or expenditure-derived variables (section C of the composite questionnaire), with the exception of “average daily per-capita clothing expenditures” and “expenditures on others” (social events, leisure) in the past 12 months.

The highest R2 was achieved by the BEST15 set, while the best performance on the accuracy measures, as well as the lowest Undercoverage and Leakage levels were registered in the BEST10 regressor set. This set yielded a Total Accuracy of 76.27 percent. Poverty Accuracy was 58.43 percent and Non-Poverty Accuracy was 84.8 percent. The Predicted Poverty Incidence was 29.19 percent.

Compared to Model 1, Model 2 shows generally lower accuracy figures in each of the three sets and a higher degree of Undercoverage and Leakage. The Predicted Poverty Incidence was slightly lower (further underestimating the poverty rate), yielding a lower PIE level. Poverty Accuracy decreased from 63.53 (Model 1, BEST15) to 58.43 percent.

In terms of poverty dimensions, this model also presents multidimensional indicators in each of the BEST* sets. Half of the indicators present in Model 2 were also present in Model 1 (BEST15), while the new indicators are from the same type as the indicators which they replaced.

Table 3.2.1 — Summary of accuracy results, Model 2

Variables	Model performance (%, percentage pt)	
BEST 5 indicators: R2 adjusted = 0.478		
Roof with banana leaves, fiber, grass, bamboo or wood	Total Accuracy	74.36
Amount that household needs per month to live	Poverty Accuracy	53.33
Do you have nursery/pre primary center in your community?	Non-Poverty Accuracy	84.42
Average daily per-capita clothing expenditures	Undercoverage	46.67
Expenditures on other expenditures(social events, leisure), last 12months	Leakage	32.54
	Pred. Pov. Incidence	27.79
	PIE	-4.57
	BPAC	39.22
Best 10 indicators: R2 adjusted = 0.540		
Next best five indicators:	Total Accuracy	76.26
Cooking fuel is charcoal or paraffin	Poverty Accuracy	58.43
House size: large	Non-Poverty Accuracy	84.80
Do you have immunization center in your community?	Undercoverage	41.57
Value of panga	Leakage	31.76
Maximum education level of any household member is post secondary/superior level (completed)	Pred. Pov. Incidence	29.19
	PIE	-3.17
Maximum education of males is secondary/post primary/J1 level (completed)	BPAC	48.62
Removed indicators:		
Roof with banana leaves, fiber, grass, bamboo or wood		
Best 15 indicators: R2 adjusted =0.569		
Next best five indicators:	Total Accuracy	75.89
Age of youngest household member	Poverty Accuracy	58.04
In the last 7 days, how many days did a main meal consist only of staple food, plant protein and any vegetables?	Non-Poverty Accuracy	84.43
	Undercoverage	41.96
Are you or members of household denied service or do you/they have only limited opportunity to sanitation services?	Leakage	32.55
	Pred. Pov. Incidence	29.31
Household head is widow	PIE	-3.05
Maximum education level of any household member is no schooling/not completed grade1	BPAC	48.63

3.3 MODEL 3

As mentioned earlier, this model was not created because the variable total value of household assets was not chosen by the MAXR procedure within the best 250 indicators for the regression analysis.

3.4 MODEL 4

This model is similar to Model 2, but excludes the variables “average daily per-capita clothing expenditures” and “other expenditures” (e.g. social events, leisure) in the past 12 months. Adjusted R2

levels are practically the same as in model 2 for each of the BEST* sets, being the maximum 0.564 at the BEST15 set.

In spite of these similarities, Model 4 has a better performance (around 3 percentage points higher) than Model 2 in all accuracy measures. Furthermore, from all models presented in this chapter Model 4 registers the highest Non-Poverty Accuracy (87.43%) and the lowest Leakage level (26.27%). The model achieves the best performance at the BEST15 set, with a Total Accuracy of 78.93% and a Poverty Accuracy of 61.18%. The Predicted Poverty Incidence was 28.3%, 4.06 percentage points lower (PIE) than the Actual Poverty Incidence.

With respect to the indicators selected for the model, the type and nature of the variables selected is similar to the previous models. The BEST15 set presents a multidimensional combination on which housing, community, education and food security variables continue to play a predominant role.

Table 3.4.1 — Summary of accuracy results, Model 4

Variables	Model performance (%, percentage pt)	
BEST 5 indicators: R2 adjusted = 0.477		
Cooking fuel is charcoal or paraffin	Total Accuracy	74.75
Value of panga	Poverty Accuracy	53.33
Do you have nursery/pre primary center?	Non-Poverty Accuracy	84.99
Maximum education level of any household member is post secondary/superior level (completed)	Undercoverage	46.67
	Leakage	31.37
Maximum education level of males is secondary/post primary/J1 level (completed)	Pred. Pov. Incidence	27.41
	PIE	-4.95
	BPAC	38.03
Best 10 indicators: R2 adjusted = 0.535		
Next best five indicators:	Total Accuracy	76.40
House size: large	Poverty Accuracy	57.65
Do you have immunization center in your community?	Non-Poverty Accuracy	85.37
Amount that household needs per month to live	Undercoverage	42.35
Number of days in past 7 days with any of 4 superior food eaten (max. 28)	Leakage	30.59
Household head is widow	Pred. Pov. Incidence	28.55
	PIE	-3.81
	BPAC	45.89
Best 15 indicators: R2 adjusted =0.564		
Next best five indicators:	Total Accuracy	78.93
Roof with banana leaves, fiber, grass, bamboo, or wood	Poverty Accuracy	61.18
Household never ate sweet potato because other food was scarce	Non-Poverty Accuracy	87.43
Total household members in farmers group	Undercoverage	38.82
Are you or members of your household denied service or do you/they have only limited opportunity to sanitation services?	Leakage	26.27
	Pred. Pov. Incidence	28.30
Do you have market in your community?	PIE	-4.06
	BPAC	48.63

3.5 MODEL 5

Model 5 is based on Model 4, but excludes all subjective variables. All variables related to subjective self-assessment of the adequacy of food consumption (i.e., Food Security Scale variables from Freedom from Hunger), vulnerability, the respondents' own poverty assessment, as well as the assessment of the dwelling's condition were dropped, leaving out some important dimensions. In spite of this, the adjusted R^2 values presented only a minimal decrease while the accuracy figures lowered in less than 1 percentage point.

The best performance was achieved in the BEST15 set with 78.17 percent Total Accuracy, 59.61 percent Poverty Accuracy and 87.05 percent Non-Poverty Accuracy. Undercoverage was 40.39 percent, while Leakage registered the second best level from all models with 27.05 percent. In addition, from all BEST15 models, this set presented the lowest Predicted Poverty Incidence (28.05%), yielding a PIE value of -4.31 percentage points.

The exclusion of subjective variables caused previously neglected housing and community related variables to enter into the best combinations, making this model more verifiable than the previous ones. Interesting to note is that the BEST5 set selected the same variables as the BEST5 from Model 4. Following the results observed up to now, the model incorporated a similar set of multidimensional indicators into BEST* sets.

Table 3.5.1 — Summary of accuracy results, Model 5

Variables	Model performance (%, percentage pt)	
Best 5 indicators: R2 adjusted = 0.477		
Cooking fuel is charcoal or paraffin	Total Accuracy	74.75
Value of panga	Poverty Accuracy	53.33
Do you have nursery/pre primary center in your community?	Non-Poverty Accuracy	84.99
Maximum education level of any household member is post secondary/superior level (completed)	Undercoverage	46.67
	Leakage	31.37
Maximum education level of males is secondary/post primary/J1 level (completed)	Pred. Pov. Incidence	27.41
	PIE	-4.95
	BPAC	38.03
Best 10 indicators: R2 adjusted = 0.527		
Next best five indicators:	Total Accuracy	76.65
House size: large	Poverty Accuracy	56.47
Do you have immunization center in your community?	Non-Poverty Accuracy	86.30
Do you have market in your community?	Undercoverage	43.53
Household head is widow	Leakage	28.63
Median education level of household members is primary level (completed)	Pred. Pov. Incidence	27.54
	PIE	-4.82
	BPAC	41.57
Best 15 indicators: R2 adjusted =0.553		
Next best five indicators:	Total Accuracy	78.17
Household cooks in one of the rooms in the house	Poverty Accuracy	59.61
Lighting source: Gas lamp or electricity (neighbor, public or own socket)	Non-Poverty Accuracy	87.05
Roof with banana leaves, fiber, grass, bamboo or wood	Undercoverage	40.39
Total household members in farmers group	Leakage	27.06
Costs of home improvements	Pred. Pov. Incidence	28.05
	PIE	-4.31
	BPAC	46.27

3.6 MODEL 6

This model excludes all monetary variables. The adjusted R^2 dropped only slightly, ranging from 0.471 to a maximum of 0.550 in the BEST15 set.

In general, Model 6 performed similarly to Model 5. The best performance was observed in the BEST15 set where Total Accuracy reached 78.81 percent. Poverty Accuracy increased to 62.74 percent. On the contrary, Non-Poverty Accuracy decreased to 86.49 percent. As a result, Undercoverage decreased to 37.25 percent and Leakage increased to 28.23 percent. The Predicted Poverty Incidence was 29.44 percent.

In terms of the variables incorporated into the BEST* sets, this model used the same type of variables as Model 5. The main difference between the two is the replacement of value-based asset variables by those referring to the incidence of ownership or number of units owned of such assets. At this point, housing, community, and education variables constitute the majority of the variables included in the BEST sets.

Table 3.6.1 — Summary of accuracy results, Model 6

Variables	Model performance (%, percentage point)	
Best 5 indicators: R2 adjusted = 0.471		
Cooking fuel is charcoal or paraffin	Total Accuracy	74.36
Do you have nursery/pre primary center in your community?	Poverty Accuracy	52.55
Panga ownership	Non-Poverty Accuracy	84.80
Maximum education level of any household member is post secondary/superior level (completed)	Undercoverage	47045
	Leakage	31.76
Maximum education level of males is secondary/post primary/J1 level (completed)	Pred. Pov. Incidence	27.28
	PIE	-5.08
	BPAC	36.86
Best 10 indicators: R2 adjusted = 0.523		
Next best five indicators:	Total Accuracy	76.40
House size: large	Poverty Accuracy	56.08
Do you have immunization center in your community?	Non-Poverty Accuracy	86.12
Do you have market in your community?	Undercoverage	43.92
Household head is widow	Leakage	29.02
Median education level of household members is primary level (completed)	Pred. Pov. Incidence	27.54
	PIE	-4.82
	BPAC	41.18
Best 15 indicators: R2 adjusted =0.550		
Next best five indicators:	Total Accuracy	78.81
Household cooks in one of the rooms in the house	Poverty Accuracy	62.74
Lighting source: Gas lamp or electricity (neighbor, public or own socket)	Non-Poverty Accuracy	86.49
Roof with banana leaves, fiber, grass, bamboo, or wood	Undercoverage	37.25
Total household members in farmers group	Leakage	28.23
Poultry ownership	Pred. Pov. Incidence	29.44
	PIE	-2.92
	BPAC	53.72

3.7 MODEL 7

Model 7 incorporates 165 poverty indicators that have been rated as easily verifiable (score 4 or 5) by NIDA, based on their experience in conducting field research and surveys in Uganda. (See Verifiability score of the variables, Annex D, Table 1, for NIDA's ratings for all 250 regressors.)

In terms of adjusted R^2, Model 7's performance is similar to Model 6, ranging from 0.476 to 0.556 in the BEST15 set.

Model 7 achieved the best performance on the BEST10 set, which presented slightly better results than the BEST15 set. Total Accuracy was 77.28 percent, while Poverty Accuracy was 59.21 percent. Non-Poverty Accuracy was 85.92 percent. The Predicted Poverty Incidence was 28.68 percent, deriving a PIE value of -3.68 percentage points.

Considering that the indicators were easily verifiable as deemed by NIDA, Model 7's BEST10 represents the most practical and operational approach for poverty assessment. Nevertheless, the loss of accuracy — especially among the very poor — is significant (63.53% in BEST15 Model 1 versus 58.82 % in BEST15 Model 7). This trade off between practicality and accuracy should be seriously considered by practitioners who want to target the very poor.

Table 3.7.1 — Summary of accuracy results, Model 7

Variables	Model performance (%, percentage point)	
Best 5 indicators: R2 adjusted = 0.476		
Cooking fuel is charcoal or paraffin	Total Accuracy	74.74
Value of panga	Poverty Accuracy	53.33
Do you have nursery/pre primary center in your community?	Non-Poverty Accuracy	84.99
Maximum education level of any household member is post secondary/superior level (completed)	Undercoverage	46.66
	Leakage	31.37
Maximum education level of males is secondary/post primary/J1 level (completed)	Pred. Pov. Incidence	27.41
	PIE	-4.95
	BPAC	38.04
Best 10 indicators: R2 adjusted = 0.531		
Next best five indicators:	Total Accuracy	77.28
House size: large	Poverty Accuracy	59.21
Do you have immunization center in your community?	Non-Poverty Accuracy	85.92
Amount household needs per month to live	Undercoverage	40.78
Do you have market in your community?	Leakage	29.41
Household head is widow	Pred. Pov. Incidence	28.68
	PIE	-3.68
	BPAC	47.84
Best 15 indicators: R2 adjusted =0.556		
Next best five indicators:	Total Accuracy	77.15
Age of youngest household member	Poverty Accuracy	58.82
Lighting source: Gas lamp or electricity (neighbor, public or own socket)	Non-Poverty Accuracy	85.92
Roof with banana leaves, fiber, grass, bamboo, or wood	Undercoverage	41.18
Whether spouse owns any type of shoes	Leakage	29.41
Are you or members of your household denied service or do you/they have only limited opportunity to sanitation services?	Pred. Pov. Incidence	28.55
	PIE	-3.81
	BPAC	47.06

The indicators incorporated in Model 7's best set have a more multidimensional nature than indicators selected indicators in Model 6. Housing characteristics, as well as community and education variables continue to play an important role, but their proportion within the best set is lower. This facilitates the incorporation of other variables such as participation on organizations, access to services, and monetary variables.

3.8 MODEL 8

Model 8 is similar to Model 7, but includes the best two expenditure variables plus five powerful subjective variables:

- Did you or any other adult in your household stop eating meals for an entire day because you didn't have enough money to buy food?
- Position on the ladder of a household with a monthly income of 70,000 Ug.Sh.
- Number of days in past seven days with any of four superior food eaten (maximum = 28)
- Amount household needs per month to live
- Change in number of steps (ladder) from seven years ago to today (if minus = worsened)

Including these variables resulted in a minor increase in adjusted R2 levels compared to Model 7. Compared to Model 7, Model 8's performance for the BEST15 set substantially improved.

The BEST15 set achieved a Total Accuracy of 77.91 percent, a Poverty Accuracy of 62.74 percent (second best from all models) and a Non-Poverty Accuracy of 85.17 percent. Furthermore, the Predicted Poverty Incidence of 30.33 percent was the highest among all models, being only 2.03 percentage points (PIE) lower than the Actual Poverty Incidence. The BEST15 of Model 8 yielded the best PIE as well as the highest BPAC among all models in this chapter.[21]

With regard to the variables incorporated into the BEST* sets, it can be observed that expenditure and food security related variables became more relevant and replaced some of the housing, education, and community variables which were used in Model 7. This issue reflects the importance that expenditure and subjective variables may have when assessing poverty in Uganda. Nevertheless, the inclusion of these variables caused a decrease on model performance on the BEST5 and BEST10 sets.

These results indicate that Model 7 is better than Model 8 when few indicators are preferred (BEST5 or 10). For a larger set of indicators, however, Model 8 achieves the highest performance.

Table 3.8.1 — Summary of accuracy results, Model 8

Variables	Model performance (%, percentage point)	
Best 5 indicators: R2 adjusted = 0.481		
Cooking fuel is charcoal or paraffin	Total Accuracy	73.6
Value of panga	Poverty Accuracy	51.76
Do you have nursery/pre primary center in your community?	Non-Poverty Accuracy	84.05
Maximum education level of males is secondary/post primary/J1 level (completed)	Undercoverage	48.23
	Leakage	33.33
Average daily per-capita clothing expenditures	Pred. Pov. Incidence	27.53
	PIE	-4.83

	BPAC	36.86
Best 10 indicators: R2 adjusted = 0.536		
Next best five indicators:	Total Accuracy	75.88
House size: large	Poverty Accuracy	56.86
Do you have immunization center in your community?	Non-Poverty Accuracy	84.99
Amount household needs per month to live	Undercoverage	43.13
Household head is widow	Leakage	31.37
Maximum education level of any household member is post secondary/superior level (completed)	Pred. Pov. Incidence	28.55
	PIE	-3.81
	BPAC	45.10
Best 15 indicators: R2 adjusted =0.567		
Next best five indicators:	Total Accuracy	77.91
Age of youngest household member	Poverty Accuracy	62.74
Roof with banana leaves, fiber, grass, bamboo, or wood	Non-Poverty Accuracy	85.17
Are you or members of your household denied service or do you/they have only limited opportunity to sanitation services?	Undercoverage	37.25
	Leakage	30.98
Do you have market in your community?	Pred. Pov. Incidence	30.33
Number of days in past 7 days any of 4 superior food eaten (max. 28)	PIE	-2.03
	BPAC	56.47

3.9 MODEL 9

Model 9 uses a set of 124 regressors which are usually found in World Bank LSMS surveys. [22] Compared to Model 8, Model 9 shows lower adjusted R^2 levels, ranging from 0.474 to 0.535 in the BEST15 set.

The highest Total Accuracy was achieved by the BEST10 set with 74.49 percent and 84.43 percent Non-Poverty Accuracy. However, the highest Poverty Accuracy (54.51%) and the lowest Undercoverage (45.49%) were achieved by the BEST15 set. The Leakage level was on its lowest level in the BEST10 set (32.55%). These results suggest that the BEST10 set can better predict not very-poor status, while the BEST15 better identifies the very poor.

The Predicted Poverty Incidence was higher for the BEST15 set (28.55%) than for the BEST10 set (27.92%). PIE was -3.81 and -4.44 percentage points, respectively.

In comparison with the previous models, Model 9 had the lowest accuracy performance. This result was expected as Model 9 has the most restrictive set of regressors. For the BEST15 set, the model achieved the lowest accuracy and BPAC figures, as well as the highest Undercoverage and Leakage levels. However, the BEST15 set did not achieve the worst PIE value, again indicating that relying on the PIE criterion alone may lead to the selection of models with low Poverty Accuracy and Non-Poverty Accuracy as Leakage and Undercoverage errors may cancel each other out.

In terms of the variables composing the best sets, housing, assets, and expenditure related variables constitute the majority of the variables included in the three sets, and are followed by education variables. Furthermore, the indicators selected in this model are similar in type and nature to the selected indicators in the previous models. These results are summarized in Annex F, Table 1.

Table 3.9.1 — Summary of accuracy results, Model 9

Variables	Model performance (%, percentage point)	
Best 5 indicators: R2 adjusted = 0.474		
Cooking fuel is charcoal or paraffin	Total Accuracy	73.60
Average daily per-capita clothing expenditures	Poverty Accuracy	51.76
Value of panga	Non-Poverty Accuracy	84.05
Household head completed only secondary/post primary education	Undercoverage	48.23
Remittances sent to relatives, last 12 months	Leakage	33.33
	Pred. Pov. Incidence	27.54
	PIE	-4.82
	BPAC	36.86
Best 10 indicators: R2 adjusted = 0.519		
Next best five indicators:	Total Accuracy	74.49
Cooking fuel is leaves, husk, or cow dung	Poverty Accuracy	53.72
Lighting source: Gas lamp or electricity (neighbor, public, or own socket)	Non-Poverty Accuracy	84.43
Roof with banana leaves, fiber, grass, bamboo, or wood	Undercoverage	46.27
Household head is widow	Leakage	32.55
Number of household members with completed superior education, excluding head	Pred. Pov. Incidence	27.92
	PIE	-4.44
	BPAC	40.00
Best 15 indicators: R2 adjusted =0.535		
Next best five indicators:	Total Accuracy	74.36
Toilet: shared or own ventilated, improved latrine or flush toilet	Poverty Accuracy	54.51
Value of agricultural land	Non-Poverty Accuracy	83.86
Value of CD players	Undercoverage	45.49
Number of poultry owned by the household	Leakage	33.72
Number of household members with no schooling or incomplete grade one, excluding head	Pred. Pov. Incidence	28.55
	PIE	-3.81
	BPAC	42.74

3.10 RESULTS FROM OTHER SINGLE-STEP REGRESSION TECHNIQUES: QUANTILE, PROBIT, AND LINEAR PROBABILITY MODEL

In this chapter, we presented nine models with single-step regressions. Each of these models were estimated in a single run with the Ordinary Least Squares (OLS) regression technique using a continuous dependent variable logarithm of daily per-capita expenditures. Annex E, Table 1 summarizes their results, whereas Annex F, Table 1, shows the BEST15 regressors for each of the nine models. (In the next chapter, we will present two-step models.)

Alternative single-step regression techniques include Probit, Quantile, and Linear Probability Models. The Linear Probability Model (LPM) and the Probit model have as dependent variable a dummy variable that is coded one if the household is very poor and zero otherwise. The LPM model is also estimated with OLS using the SAS package, and the selection of BEST15 regressors is done using the MAXR procedure.

In the Probit and the Quantile regressions, it is not feasible to use the MAXR procedure to select the set of BEST15 regressors. In order to test the accuracy performance of the Quantile regression model (which uses the log of daily per-capita expenditures as the dependent variable), the BEST15 regressors set (as it

was determined by the OLS-MAXR regression) is used. The Quantile regression models are estimated with the STATA package, whereas the Probit model is estimated with SAS. For the Probit model, where the dependent variable is a dummy variable similar to the LPM, we use the BEST15 regressors that were identified in the LPM model with the SAS MAXR procedure. The Probit model (like the LPM model) estimates the probability of a household being below the poverty line.

Annex E, Table 3, presents the accuracy performance for these alternative single-step regression techniques. We restricted the testing of these alternative regression techniques to four sets of regressors — models 1, 4, 7, 9 — and estimate the models only with a set of fifteen regressors.

Table 3.10.1 shows the complete results for Model 1's accuracy performance.

Table 3.10.1 — Summary of the accuracy results from the single-step regression techniques for Model 1

Model 1 Poverty rate: 32.36%	Adj. R2	% Total Accuracy	% Poverty Accuracy	% Under-coverage	% Leakage	PIE (% points)	BPAC (% points)
Single-step methods — MAXR variable selection							
OLS	0.587	79.06	63.53	36.47	28.24	-2.66	55.29
Quantile Regression (estimation point: 46)		79.44	68.24	31.76	31.76	0	68.24
Linear Probability	0.337	79.95	61.18	38.82	23.14	-5.08	45.50
Probit		80.71	63.92	36.08	23.53	-4.06	51.37

For the set of regressors used in Model 1, the best single-step regression technique in terms of maximizing BPAC is the Quantile technique. Through an iterative procedure involving a series of regressions with the given set of BEST15 regressors, alternative percentile points of estimation for the quantile model are tested in order to maximize BPAC. With an optimal point of estimation set at the 46^{th} percentile, the Quantile technique achieves a PIE of 0 percentage points. In other words, this model perfectly predicts the observed poverty rate. Moreover, the value for Poverty Accuracy, and hence also for BPAC, is 68.24 percentage points. In comparison with the single-step OLS, LPM, and Probit regression techniques, the gains in the Balanced Poverty Accuracy Criterion from using the Quantile technique are considerable.

Similar results concerning the differences in accuracy performance between the four alternative single-step regression techniques are shown next. The tables below present the results for the sets of regressors used in models 4, 7, and 9.

Table 3.10.2 — Summary of the accuracy results from the single-step regression techniques for Model 4

Model 4 Poverty rate: 32.36%	Adj. R2	% Total Accuracy	% Poverty Accuracy	% Under-coverage	% Leakage	PIE (% points)	BPAC (% points)
Single-step methods — MAXR variable selection							
OLS	0.564	78.93	61.18	38.82	26.28	-4.07	48.63
Quantile Regression (estimation point: 45)		79.82	69.41	30.59	31.76	0.38	68.24
Linear Probability	0.321	79.44	62.75	37.25	26.28	-3.55	51.77
Probit		79.57	63.92	36.08	27.06	-2.92	54.90

Model 4 represents the set of regressors that excludes the total value of household assets and all expenditure variables. It includes all subjective poverty indicators and most indicators from the

practitioner tools. Table 3.10.2 compares the accuracy performance of four single-step regression techniques for the set of regressors used in Model 4. The best regression technique is the single-step Quantile. This technique achieves a value for BPAC of 68.24 and a value of PIE of 0.38 percentage points. Compared to the other three regression techniques, this result is a significant improvement.

Model 7 represents the set of regressors that only includes poverty indicators that NIDA, the survey firm, rated as easily verifiable (score 4 or 5). Annex D, Table 1, provides the ratings for all 250 regressors. Table 3.10.3 compares the accuracy performance of four single-step regression techniques for the set of regressors used in Model 7. The best regression technique in terms of BPAC is again the Quantile technique, with a value for BPAC of 65.10 percentage points and a value of PIE of 0.63 percentage points. As with the previous models, when compared to the other techniques, this represents a considerable improvement with respect to BPAC.

Table 3.10.3 — Summary of the accuracy results from the single-step regression techniques for Model 7

Model 7 Poverty rate: 32.36%	Adj. R2	% Total Accuracy	% Poverty Accuracy	% Under-coverage	% Leakage	PIE (% points)	BPAC (% points)
Single-step methods — MAXR variable selection							
OLS	0.556	77.15	58.82	41.18	29.41	-3.81	47.06
Quantile Regression (estimation point: 44)		78.05	67.06	32.94	34.90	0.63	65.10
Linear Probability	0.309	79.69	61.18	38.82	23.92	-4.83	46.27
Probit		79.69	61.57	38.43	24.31	-4.57	47.45

Model 9 represents the set of regressors that is usually contained in World Bank LSMS surveys. Table 3.10.4 compares the accuracy performance of the four single-step regression techniques for the set of regressors used in Model 9. The best regression technique in terms of BPAC is, as in the previous models, the Quantile model. This technique achieves a value for BPAC of 61.57 percentage points and a value of PIE of -0.12 percentage points. Compared to the other regression techniques, this result of the best single-step regression technique again constitutes a considerable improvement with respect to BPAC.

Table 3.10.4 — Summary of the accuracy results from the single-step regression techniques for Model 9

Model 7 Poverty rate: 32.36%	Adj. R2	% Total Accuracy	% Poverty Accuracy	% Under-coverage	% Leakage	PIE (% points)	BPAC (% points)
Single-step methods — MAXR variable selection							
OLS	0.535	74.36	54.51	45.49	33.72	-3.81	42.74
Quantile Regression (estimation point: 44)		75.51	61.96	38.04	37.65	-0.12	61.57
Linear Probability	0.278	76.78	52.94	47.06	24.71	-7.23	30.59
Probit		77.66	56.86	43.14	25.88	-5.58	39.60

In conclusion, the use of the Quantile regression technique allows us to considerably improve the accuracy performance compared to single-step Ordinary Least Squares (OLS). This result holds true for all four sets of regressors which were tested in this section (i.e., Models 1, 4, 7, and 9). In the next chapter, the accuracy performance of the two-step methods is explored.

4 TWO-STEP MODELS

4.1 INTRODUCTION

The accuracy measures presented in Models 1 through 9 refer to the models' accuracy of predicting the average poverty status *for all expenditure percentiles* in the full sample. However, they do not take into account the differences in accuracy observed at different levels of expenditure (benchmark indicator "daily expenditures per capita"). These models exhibit a high Total Accuracy, but a lower Poverty Accuracy. They underestimated the Actual Poverty Incidence by a margin as large as 7.23 percentage points. The relatively low performance of the OLS models is partly driven by the level of the actual incidence of the very poor (at 32.3%). The more a country's poverty rate deviates from a level of 50 percent, the more the single-step OLS models tend to show a weaker performance.

In order to improve the estimation of poverty status, we employed a two-step approach (see Grootaert et al., 1998) that breaks down the differences in the accuracy measures by percentiles of the benchmark indicator. In step one, we estimate the original model for the full sample. Predicted daily expenditures per capita are compared against a wide range of benchmark cut-offs at different percentile levels. In step two, the model is estimated using a subsample that only includes those households whose predicted expenditures fall below the different cutoffs, in order to identify the best regressor set for that subsample. The estimation in step two is repeated with OLS, using SAS's MAXR technique. Finally, the combined accuracy level of the two models is calculated by considering the predicted status from step one for households with predicted expenditures above the different cutoffs and the predicted status from step two for the subsample of predicted expenditures below the cutoffs. The percentile cutoff that achieves the highest combined BPAC measure is considered the best two-step model.

In the remaining part of section 4.1, we present in detail first the results of the two-step OLS approach for the BEST15 regressor set of Model 1. Sections 4.2 to 4.4 present the results for similar two-step OLS approaches but using alternative sets of regressors, as defined by Models 4, 7 and 9. Overall, the two-step OLS models perform better than the single-step OLS models presented in Chapter 3. The models lead to an improvement in Poverty Accuracy and a reduction of the Poverty Incidence Error (PIE). Also, the Balanced Poverty Accuracy Criterion (BPAC) increases noticeably (see Annex E, Table 2). Section 4.5 combines the two-step approach with the three alternative regression techniques, i.e. Quantile, Probit, and the Linear Probability Model. With one set of regressors tested (Model 9), the two-step Quantile regression technique achieves the highest BPAC of over 67 percentage points and PIE values near 0 percentage points. For the regressors used in Models 1, 4, and 7, we observed the highest BPAC with the two-step LPM technique, with values above 70 percentage points. The results show that, in general, for the case of Uganda, the Quantile and LPM regression techniques perform well and can achieve a fairly good accuracy performance.

4.1.1 FIRST STEP: MODEL 1 – BEST15 SET ON FULL SAMPLE

We first evaluate the performance of Model 1 with the BEST15 regressors. Table 4.1.1 presents the results, which correspond to the results already shown in Table 3.1.2.

Table 4.1.1 — Accuracy level for the BEST 15 regressor set

Measure	Level
Total Accuracy	79.06 %
Poverty Accuracy	63.52 %
Non poverty Accuracy	86.49 %

Measure	Level
Undercoverage	36.47 %
Leakage	28.23 %
PIE	-2.67 percentage points
BPAC	55.29 percentage points

Table 4.1.2 presents a comparison of the predicted expenditures and the actual expenditures, both expressed in Ugandan shillings. The model tended to overestimate the level of expenditures, especially in the extremes of the distribution, while it underestimated the expenditures around the mean. Nevertheless, the predicted poverty incidence was 2.67 percentage points lower than the actual poverty incidence (PIE).

Table 4.1.2 — Comparison between predicted and actual expenditures.

Variable	Minimum	Maximum	Mean	Std. Deviation
Actual daily expenditures per capita, Ug.Sh. (benchmark)	42.65	11 545.66	1 293.77	1 167.71
Predicted daily expenditures per capita, Ug.Sh.	209.03	11 618.92	1 169.61	910.77

4.1.2 SECOND STEP AND COMBINED ACCURACY OF THE TWO-STEP MODEL

By testing the set of variables from Model 1[1] on the different subsamples (e.g. all expenditure percentiles above the headcount rate), the new BEST15 regressor sets were identified. Afterwards, the combined accuracy measures for all subsamples were determined and the optimal subsample was selected. For this, the main evaluation criterion was the maximization of BPAC.

Following this approach, the highest BPAC level was found when using the 51st percentile as the cutoff point for the subsample estimated in the second step. The combined accuracy measures from the two-step model are presented in Table 4.1.3.

Table 4.1.3 — Combined accuracy from two-step estimation, Model 1

Measure	51st Percentile
Number of observations in the subsample	444
Adjusted R2 for the subsample	0.333
Total Accuracy (%)	79.44
Poverty Accuracy (%)	67.84
Undercoverage (%)	32.16
Leakage (%)	31.37
PIE (% points)	-0.25
BPAC (% points)	67.06

[1] For the second step, only the BEST15 variables were used (selected by MAX R) for regression analysis on each of the subsamples.

It can be observed for the subsample that while the adjusted R2 value was lower than in the BEST15 set from the first step (0.587), the accuracy results improved noticeably. Total Accuracy slightly increased in the two-step approach, from 79.06 to 79.44 percent. Also, Poverty Accuracy increased 4.32 percentage points. While Undercoverage decreased, Leakage increased by 3.13 percentage points. The predicted poverty incidence increased, causing a reduction of the PIE level by 2.41 percentage points.

BPAC increased from 55.29 percentage points in the single-step model to 67.06 percentage points, implying an increase by 11.77 percentage points.

Table 4.1.3 presents the BEST15 regressor set obtained for the subsample at the 51st percentile. The BEST15 set from the first step (corresponding to BEST15 in Table 3.1.2) is presented for comparison. Six variables (shaded in gray) appear to be important in both steps of the model. These were:

- Roof with banana leaves, fiber, grass, bamboo, or wood
- Annualized food expenditures, recall 1 week
- Value of panga
- Do you have market/bazaar in your community?
- Do you have nursery/pre-primary center in your community?
- Household head is widow

With the exception of the annualized food expenditures, these variables were considered to be somewhat easy or easily verifiable by NIDA.

Table 4.1.4 — BEST15 regressor sets derived from the second step

Variables	Model 1	51st Percentile
Average age of household members, except head	X	
Cooking fuel is charcoal or paraffin	X	
Roof with banana leaves, fiber, grass, bamboo, or wood	X	X
Household never ate sweet potato because other food was scarce	X	
Total household members in farmers group	X	
Are you or members of your household denied service or do you/they have only limited opportunity to sanitation services?	X	
Annualized food expenditures, recall 1 week	X	X
Annualized non food expenditures 2 (education, health, and others)	X	
Value of panga	X	X
Number of days in past 7 days any of 4 superior food eaten (max. 28)	X	
Do you have market/bazaar in your community?	X	X
Do you have nursery/pre primary center in your community?	X	X
Household head is widow	X	X
Maximum education of any household member is post secondary/superior level (completed)	X	
Maximum education of males is secondary/post primary/J1 level (completed)	X	
Household sometimes ate beer type bananas because other food was scarce		X

Variables	Model 1	51st Percentile
Whether spouse owns any type of shoes		X
Maximum education of any household member is no schooling/not completed grade one		X
Toilet: shared or own ventilated, improved latrine or flush toilet		X
Poultry ownership		X
Amount that household needs per month to live		X
Distance to the bodaboda (motorcycle) station (km)?		X
Distance to the private health centre with midwife/nurse (km)?		X
Distance to access a network phone (km)?		X
Number of new regressors	15	9

The last row in Table 4.1.4 shows the number of new regressors that substitute for some of the original regressors used in step 1. The BEST15 set for the subsample rely mostly on community, housing, and asset variables.

A practitioner tool based on a two-step model would have to include questions which obtain information on the 15 regressors selected by the BEST15 model of the first step. In addition, the practitioner tool would need to obtain information about the new additional poverty indicators that have been identified among the BEST15 regressors (51st percentile) of the second step (see Table 4.1.4).

In practice, all questions related to the first and second step (15 plus 9 indicators) can be integrated into a single interview with each household. The interviewer could begin with the BEST15 indicators in the first step and then compute an estimated per capita daily expenditure. If the estimated expenditure falls above the cutoff value for the 51st percentile the household would be rated as not-very poor and the interview could be terminated. If, however, the predicted per capita expenditure value falls below this cutoff, the interview would need to be continued by asking questions related to the nine additional regressors of the second step. Based on the values obtained for the nine regressors (plus the original regressors from the first step), a second value for predicted per capita daily expenditures would be computed. If this second value is below the applicable poverty line, the household would be rated as very poor. In practice, however, it is recommended not to interrupt the interview for the calculation based on the first fifteen indicators, but rather to continue with the questions for the remaining nine poverty indicators. In this case, the calculations of one (or two) expenditure values could be done after the interview.

4.2 TWO-STEP MODEL 4

As mentioned in section 2.5, Model 4 excluded the variable “total value of household assets” as well as all expenditure variables. With this, it was possible to create a set of regressors containing all subjective poverty indicators, along with most of the indicators from the practitioners’ tools. Table 4.2.1 presents the performance of the two-step approach for this set of regressors.

Among the subsamples, the highest combined BPAC was found at the 55th percentile. Compared to the single-step model, Total Accuracy and Poverty Accuracy increased by 1.15 and 6.27 percentage points, respectively. Also, Leakage increased by 2.74 percentage points.

The two-step approach predicted a higher incidence of poverty than the single-step model. PIE changed from -4.07 to -1.14 percentage points, reducing the difference between the predicted and observed poverty headcount. The gain in BPAC is 15.29 percentage points.

Table 4.2.1 — Accuracy results for Model 4

Measure	Model 4	55th Percentile
Number of observations	788	485
Adjusted R2 for the sample/subsample	0.564	0.302
Total Accuracy (%)	78.93	80.08
Poverty Accuracy (%)	61.18	67.45
Undercoverage (%)	38.82	32.55
Leakage (%)	26.28	29.02
PIE (% points)	-4.07	-1.14
BPAC (% points)	48.63	63.92

4.3 TWO-STEP MODEL 7

As explained in Chapter 2, Model 7 was constructed using the variables that NIDA rated as easily verifiable. Table 4.3.1 presents the performance of the two-step approach for this set of regressors.

Table 4.3.1 — Accuracy results for Model 7

Measure	Model 7	57th Percentile
Number of observations	788	498
Adjusted R2 for the sample/subsample	0.556	0.319
Total Accuracy (%)	77.15	79.70
Poverty Accuracy (%)	58.82	68.24
Undercoverage (%)	41.18	31.76
Leakage (%)	29.41	30.98
PIE (% points)	-3.81	-0.25
BPAC (% points)	47.06	67.45

The highest combined BPAC was found at the 57th percentile. Total Accuracy increased 2.55 percentage points. Also, Poverty Accuracy registered a considerable increase of 9.42 percentage points. While Undercoverage decreased by around 10 percentage points, the Leakage level increased by around 1 percentage point.

The two-step approach predicted a higher incidence of poverty than that predicted by the single step model. PIE decreased from -3.81 to -0.25 percentage points, reducing the difference between predicted and observed poverty headcount. The gain in BPAC is 20.39 percentage points.

4.4 TWO-STEP MODEL 9

As presented in the previous chapter, Model 9 incorporated 124 variables which are commonly found in LSMS datasets.

The highest combined BPAC was found in the 54th percentile. For the second step, the subsample consisted of 464 households and the model yielded an adjusted R2 of 0.210, considerably lower than for the single-step (full sample) model.

Table 4.4.1 — Accuracy results for Model 9

Measure	Model 9	54th Percentile
Number of observations	788	464
Adjusted R2 for the sample/subsample	0.535	0.210
Total Accuracy (%)	74.36	77.79
Poverty Accuracy (%)	54.51	63.92
Undercoverage (%)	45.49	36.08
Leakage (%)	33.72	32.55
PIE (% points)	-3.81	-1.14
BPAC (% points)	42.74	60.39

With respect to the accuracy measures, a general improvement can be observed. Total Accuracy increased 3.49 percentage points. Poverty Accuracy increased from 54.51 to 63.92 percent, yielding a change of 9.41 percentage points, which was similar to the change registered in Model 7.

In contrast to Models 1 and 7, both Undercoverage and Leakage decreased. In addition, this model showed the lowest improvement in the predicted headcount, with a PIE level of minus 1.14 percentage points. The gain in BPAC was 17.65 percentage points.

4.5 RESULTS FROM OTHER TWO-STEP REGRESSION TECHNIQUES: QUANTILE, PROBIT, AND LINEAR PROBABILITY MODEL

The previous three sections have presented models that were estimated with the Ordinary Least Squares (OLS) regression technique using the continuous dependent variable logarithm of daily per-capita expenditures. Annex E, Table 2, summarizes their results, whereas Annex F, Table 2, shows the BEST15 regressors for each of the three regressions using different sets of regressors (i.e., Models 1, 4, 7, and 9).

Similar to single-step regression techniques, alternative formulations of the two-step approach again consist of using the Probit, Quantile, and Linear Probability Models as alternative regression techniques. For example, in a two-step modeling framework, a two-step Probit model consists of running two Probit regressions. Similar to the above OLS models, the first run includes the full sample, whereas the second includes a subsample of poorer households.

As already mentioned in section 3.10, the Linear Probability Model (LPM) and the Probit model have as dependent variable a dummy variable that is coded one if the household is very poor and zero otherwise. Similar to the OLS regression technique presented in sections 4.1 to 4.4, the Quantile regression model uses the log of daily per-capita expenditures as the dependent variable. Similar to the single-step models, the regressors used in the two-step Quantile regressions are the same as those identified with the SAS MAXR technique for the two-step OLS regressions. In addition, the percentile cutoff point for the second-step subsample in Quantile regressions is the same as the one determined in the two-step OLS model. Moreover, the point of estimation for the first step Quantile regression is set similar to the one found optimal for the single-step Quantile model presented in Chapter 3. To identify the optimal second point of estimation for the second-step Quantile regression, we again employ an iterative procedure that runs a series of regressions with the given set of BEST15 regressors (as determined by the second-step OLS regression). Also similar to the single-step models, the regressors used in the two-step Probit regressions

are the same as those identified by the SAS MAXR technique for the two-step LPM regressions, and the cutoff point for the subsample in the two-step Probit is the same as in the LPM model.

We restrict the testing of the three alternative two-step regression techniques to four sets of regressors, namely Models 1, 4, 7 and 9. Again, the models are estimated with a set of BEST15 regressors.

For Model 1, the results concerning the accuracy performance of the four two-step regression techniques are shown in Table 4.5.1.

Table 4.5.1 — Summary of the accuracy results of two-step regression techniques for Model 1

Model 1 Poverty rate: 32.36%	Adj. R2	% Total Accuracy	% Poverty Accuracy	% Under-coverage	% Leakage	PIE (% points)	BPAC (% points)
Two-step methods — MAXR variable selection							
OLS 51st Percentile	0.333 subsample	79.44	67.84	32.16	31.37	-0.25	67.06
Quantile Regression (estimation points 46, 24) - 51% cutoff		79.19	68.24	31.76	32.54	0.25	67.45
Linear Probability 54th Percentile	0.257 subsample	83.12	75.29	24.70	27.45	0.88	72.54
Probit 54th Percentile		81.59	69.80	30.20	26.67	-1.14	66.27

Table 4.5.1 shows the accuracy performance of the three alternative two-step regression techniques. The OLS model is similar to the one presented in section 4.1.

For the set of regressors as identified by Model 1, the above table shows that the best two-step regression technique in terms of maximizing BPAC is the Linear Probability Model. Using the poorest 54 percent, the LPM achieves a PIE of 0.88 percentage points which is noticeably behind the PIE values of 0.25 and –0.25 from the two-step OLS and Quantile models. The value for Poverty Accuracy is 75.29 %, and for BPAC it is 72.54 percentage points. In comparison with the two-step OLS regression technique, gains in Balanced Poverty Accuracy Criterion from using the LPM technique are achieved. Annex E.3 shows results from all four single-step and two-step models in one table.

For Model 4, the results concerning the accuracy performance of the four two-step regression techniques are shown in Table 4.5.2.

Table 4.5.2 — Summary of the accuracy results of two-step regression techniques for Model 4

Model 1 Poverty rate: 32.36%	Adj. R2	% Total Accuracy	% Poverty Accuracy	% Under-coverage	% Leakage	PIE (% points)	BPAC (% points)
Two-step methods — MAXR variable selection							
OLS 55th Percentile	0.302 subsample	80.08	67.45	32.55	29.02	-1.14	63.92
Quantile Regression (estimation points 45, 26) - 55% cutoff		80.46	70.59	29.41	30.98	0.51	69.02
Linear Probability 55th Percentile	0.214 subsample	81.73	71.76	28.24	28.24	0	71.76
Probit 55th Percentile		81.35	70.20	29.80	27.84	-0.63	68.24

For the set of regressors used in Model 4, Table 4.5.2 shows that the best two-step regression technique in terms of maximizing BPAC is the Linear Probability Model. Using the poorest 55 percent, the LPM achieves a PIE of 0 percentage points, meaning that the model perfectly predicts the observed poverty rate. The value for Poverty Accuracy and hence also for BPAC is 71.76 percentage points. In comparison with the other two-step regression techniques, the gains in Balanced Poverty Accuracy Criterion (BPAC) from using LPM are noticeable.

For Model 7, the results concerning the accuracy performance of the four two-step regression techniques are shown in Table 4.5.3.

Table 4.5.3 — Summary of the accuracy results of two-step regression techniques for Model 7

Model 7 Poverty rate: 32.36%	Adj. R2	% Total Accuracy	% Poverty Accuracy	% Under-coverage	% Leakage	PIE (% points)	BPAC (% points)
Two-step methods — MAXR variable selection							
OLS 57th Percentile	0.319 subsample	79.70	68.24	31.76	30.98	-0.25	67.45
Quantile Regression (estimation points 46, 27) - 57% cutoff		78.93	67.84	32.16	32.94	0.25	67.06
Linear Probability 42nd Percentile	0.152 subsample	81.97	73.33	26.66	29.01	0.76	70.98
Probit 42nd Percentile		80.45	70.58	29.41	30.98	0.51	69.01

For the set of regressors used in Model 7, Table 4.5.3 shows that the best two-step regression technique in terms of maximizing BPAC is again the Linear Probability Model. Using the poorest 42 percent, the LPM achieves a PIE of 0.76 percentage points. In other words, this model almost perfectly predicts the observed poverty rate. Moreover, the value for Poverty Accuracy is 73.33 percent and the value for BPAC is 70.98 percentage points. In comparison with the other two-step regression techniques, gains in Poverty Accuracy and Balanced Poverty Accuracy Criterion (BPAC) from using the LPM are achieved.

For Model 9, the results concerning the accuracy performance of the four two-step regression techniques are shown in Table 4.5.4.

Table 4.5.4 — Summary of the accuracy results of two-step regression techniques for Model 9

Model 9 Poverty rate: 32.36%	Adj. R2	% Total Accuracy	% Poverty Accuracy	% Under-coverage	% Leakage	PIE (% points)	BPAC (% points)
Two-step methods — MAXR variable selection							
OLS 54th Percentile	0.210 subsample	77.79	63.92	36.08	32.55	-1.14	60.39
Quantile Regression (estimation points 44,24) - 54% cutoff		79.06	67.84	32.16	32.55	0.13	67.45
Linear Probability 49th Percentile	0.159 subsample	79.44	64.71	35.29	28.23	-2.28	57.64
Probit 49th Percentile		79.44	64.31	35.69	27.84	-2.53	56.47

For the set of regressors used in Model 9 (i.e., the regressors usually contained in LSMS data sets), Table 4.5.4 shows that the best two-step regression technique in terms of maximizing BPAC is Quantile regression. With points of estimation set at the 44th percentile for the first step and at the 24th percentile for the second step and using the poorest 54 percent (similar to the two-step OLS) as the subsample for the second step, the Quantile regression achieves a PIE of 0.13 percentage points. In other words, this model almost perfectly predicts the observed poverty rate. Moreover, the value for Poverty Accuracy is 67.84 percent, and for BPAC it is 67.45 percentage points. In comparison with the other two-step regression techniques, the gains in Poverty Accuracy and Balanced Poverty Accuracy Criterion (BPAC) from using the two-step quantile regression are considerable.

Annex E, Table 3, presents the accuracy performance for these alternative two-step regression techniques and shows the results for the four single-step regression techniques that were already presented in section 3.10. The table shows that the two-step Linear Probability technique achieves the highest BPAC for the first three sets of regressors (i.e. Models 1, 4, and 7), whereas in Model 9 the highest BPAC is achieved using the two-step Quantile regression technique. Annex F, Table 3, shows the set of regressors that were used by these two-step techniques with the best BPAC results.

5 SUMMARY

This report first presented nine single-step regression models, each with a set of best 5, 10, and 15 regressors. These models were estimated with the Ordinary Least Squares (OLS) regression technique. The single-step OLS models did not efficiently predict the percentage of very-poor households, in spite of a Total Accuracy exceeding 70 percent. Whereas tests of the single-step OLS regression technique were performed for different sets of regressors — Models 1 through 9 — the tests of alternate single-step and two-step regression techniques in this report were limited to four sets of regressors — Models 1, 4, 7, 9. As in the other three test countries (Bangladesh, Peru, and Kazakhstan), Model 1 includes all poverty indicators. In comparison, Model 4 excludes all expenditure categories and the total value of household assets but still contains most poverty indicators from practitioners' tools as well as subjective poverty indicators. Model 7 includes only indicators that are deemed by experienced survey firm staff as being highly verifiable and easy to ask; it is thus the most practical model. Model 9 uses indicators similar to those found in the World Bank's LSMS data sets.

This report also contains tests of three alternative single-step regression techniques, namely the Probit, Quantile, and Linear Probability Model (LPM) techniques. Among these, the single-step Quantile regression technique yielded the highest value for BPAC for all four sets of regressors. In addition, we tested two-step models using the four different regression techniques.

The single-step OLS models show satisfactory levels of Total Accuracy, that is, they accurately predict a large percentage of households which actually fall into a given category. However, all nine models show lower Poverty Accuracy levels, and all consistently underestimate the poverty headcount, yielding negative values of PIE. The values for BPAC range between 42.74 and 56.47 percentage points. Annex E, Table 1, provides a summary of accuracy results for all single-step OLS models. Annex F, Table 1, summarizes the variables used as the 15 regressors in the different models.

We made considerable improvements by using Quantile regression technique in a single-step framework. For Models 1, 4, 7, and 9, we obtained positive values for BPAC of over 60 percentage points, while the values for PIE were close to zero, indicating an almost perfect prediction of the observed poverty rate. These results represent vast improvements compared to the results obtained with the single-step OLS, Probit, and LPM techniques.

From Annex E, Table 1, we observe that all models estimated with a single-step using Ordinary Least Squares (OLS) were less accurate for the very poor than for the not-very poor. This implies that the inaccuracies in prediction are not equally distributed over all expenditure percentiles but are systematically higher for the very poor. This problem of unbalanced accuracies can be potentially reduced by the use of two-step models, following a method pioneered by Grootaert et al. (1998). The computational costs of these models, however, are higher than for single-step models.

The results of the two-step OLS models presented in Chapter 4 compare favorably with the single-step OLS models presented in Chapter 3. While the Total Accuracy of the two-step models is only marginally higher than for the single-step OLS models (and in certain cases it is even slightly lower), the two-step models have a clear advantage in estimating the proportion of the population that is very poor, thereby improving estimates of headcount indices.

We explored three alternative regression techniques in a two-step regression framework for the four different sets of regressors. Again, these are: Model 1 (full set of regressors), Model 4 (all regressors except total value of assets and expenditure categories), Model 7 (the model thought to be most practical), and Model 9 (with a set of regressors usually contained in LSMS data sets). The alternative regression techniques are Quantile, Probit, and the Linear Probability Model (LPM). When considering all eight different regression techniques tested (i.e. the four single-step and the four two-step techniques), the two-step Linear Probability regression model achieved the highest BPAC for three sets of regressors (Models 1, 4, and 7; with 72.54, 71.76, and 70.98 percentage points respectively), while the two-step Quantile regression achieved the highest BPAC in Model 9 with 67.45 percentage points. The BPAC value of Model 7 is only 2 percentage points lower than that for Model 1. Thus, the use of indicators that have received high verifiability scores (instead of complex monetary indicators or subjective indicators) implies relatively minor losses in accuracy performance.

Annex F, Table 3, lists the set of best regressors that were used in the model achieving the highest value of BPAC. The regression models identified a set of multi-dimensional regressors. The most important dimensions of poverty were demography, education; specific consumer, agricultural, or financial assets; housing characteristics; clothing expenditures; food consumption and insecurity; social capital; and the infrastructure and institutional characteristics of the community. With regard to specific regressors, it is first of note that the value of total assets turned out to be insignificant (which is not what we found in Bangladesh and Kazakhstan). In rural Uganda, many assets are home-produced and seldom bought or sold. According to NIDA, the respondents had difficulty in estimating the value for many types of assets, indicating the possibility of large measurement errors and therefore explaining the insignificance of this poverty indicator in the preceding analysis. Therefore, the total value of assets is not considered an adequate indicator for Uganda. Second, like in Bangladesh, Kazakhstan and Peru, the per-capita expenditure on clothing is the best or second-best single expenditure category out of the total of 13 expenditure categories. Hence, clothing expenditures has been consistently identified as the most robust indicator among the possible expenditure categories in all of the four IRIS test countries. The good news is that with sufficient training by interviewers, estimates on clothing expenditures might be fairly easily obtained from respondents as suggested by the verifiability score of 4 given by NIDA for the clothing expenditure variables (see Annex D).

In conjunction with tests in Bangladesh, Kazakhstan, and Peru, the accuracy tests in Uganda show that the choice of a suitable regression technique is an empirical issue. The choice is influenced by the level of the poverty rate. In countries with a low poverty rate, two-step techniques appear to yield better results in terms of the Balanced Poverty Accuracy Criterion (BPAC). Overall, the results for Uganda highlight that the challenge of finding tools capable of accurately identifying very-poor households can only be satisfyingly addressed with the use of more complex regression techniques.

6 BIBLIOGRAPHY

Ahmad, Sultan. 2003. Purchasing power parity for international comparisons of poverty: Sources and methods. Mimeograph, Development Data Group, The World Bank, www.worldbank.org/data/ICP.

Appleton, S., and S. Ssewanjana. 2003. Poverty analysis in Uganda, 2002/03. Mimeo, Economic Policy Research Centre, Makerere University, Kampala, Uganda (cited from Ssewanjana et al., 2004).

Deininger, Klaus, and A. J. Okidi. 2003. Growth and poverty reduction in Uganda, 1992-2000. Panel data evidence. Development Policy Review, Vol. 21, pp. 481-509.

Heston, Alan, Robert Summers and Bettina Aten. 2002. Penn World Table Version 6.1, Center for International Comparisons at the University of Pennsylvania (CICUP), October 2002.

IRIS. 2005. Note on Assessment and Improvement of Tool Accuracy. Mimeograph, Revised version from June 2, 2005. IRIS center, University of Maryland.

Kappel, Robert, Jann Lay, and Susan Steiner. 2005. Uganda: No more pro-poor growth? Development Policy Review, Vol. 23 (forthcoming).

Karshenas, Massoud. 2004. Global poverty estimates and the millennium development goals: Towards a unified framework. International Labor Organization (ILO), Employment Strategy Department, Geneva.

Chen, Shaohua, and Martin Ravallion. 2004. How have the worlds' poorest fared since the early 1980s? World Bank Policy Research Paper No. 3341, June 2004, Washington, D.C.

Reddy, Sanjay G., and Thomas W. Pogge. 2003. How not to count the poor. Unpublished mimeograph, Columbia University, see http://www.socialanalysis.org.

Ssewanjana, N.S., Okidi, A.J., Angemi, D., and V. Barungi. 2004. Understanding the determinants of income inequality in Uganda. CSAE /WPS 2004/ Number 29. Economic Policy Research Centre, Makerere University, Kampala, Uganda.

World Bank. 2005. World development indicators 2004. Section 2.5 on Poverty. The World Bank, Washington, D.C.

Grootaert, Christiaan, and Jeanine Braithwaite. 1998. "Poverty correlates and indicator-based targeting in Eastern Europe and the Former Soviet Union." Poverty Reduction and Economic Management Network, Environmentally and Socially Sustainable Development Network. Washington D.C.: The World Bank.

NIDA. 2004. Poverty assessment survey in Uganda. Field technical report submitted by Nkoola Institutional Development Associates (NIDA) to the IRIS Center, University of Maryland.

Uganda Bureau of Statistics (UBOS). 2002. Uganda population and housing census. Provisional results. Kampala, Uganda, November 2002.

WHO. 1985. Energy and protein requirements. WHO Technical Report Series No. 724. Geneva, Switzerland.

Zeller, Manfred. 2004. Review of Poverty Assessment Tools. Report submitted to IRIS and USAID as part of the Developing Poverty Assessment Tools Project. January 2004. (Available online at http://www.povertytools.org).

Zeller, Manfred, Gabriela V. Alcaraz, and Julia Johannsen. 2005. Developing and testing poverty assessment tools: Results from accuracy tests in Bangladesh. IRIS Center, University of Maryland, College Park. First report in November 2004. Final report in August 2005 (Available online at http://www.povertytools.org).

Zeller, Manfred, Julia Johannsen, and Gabriela V. Alcaraz. 2005. Developing and testing poverty assessment tools: Results from accuracy tests in Peru. IRIS Center, University of Maryland, College Park. First report in February 2005. Final report in August 2005 (Available online at http://www.povertytools.org).

Zeller, Manfred, and Gabriela V. Alcaraz. 2005. Developing and testing poverty assessment tools: Results from accuracy tests in Kazakhstan. IRIS Center, University of Maryland, College Park. First report in June 2005, and final report in August 2005. (Available online at http://www.povertytools.org).

ANNEXES

ANNEX A

Annex A, Table 1. Distribution and size of the sample

Region	District	County	Sub-county	Code of enumeration area (Source: UBOS)	Local Council (LC1)	Urban (1=Yes, 2 = No)	Sample siz (Number of households
Northern	Arua	Maracha	Oluffe	EA05	AYIVU A	2	32
Northern	Nebbi	Jonam	Panyango	EA09	ANYAMA	2	32
Northern	Apac	Maruzi	Apac	EA02	OMELLE	2	32
Western	Bushenyi	Buhweju	Burere	EA02	MABANGA	2	32
Western	Hoima	Bugahya	Kitoba	EA03	KILRANGOBE	2	32
Western	Kabarole	Burahya	Karambi	EA09	FUTI BUTANGWA LCI A	2	32
Western	Kibaale	Buyaga	Kyanaisoke	EA03	KISENGYA	2	32
Western	Masindi	Kibanda	Mutunda	EA06	FUNGUA MACO	2	32
Western	Mbarara	Municipality	Kamukuzi	EA10	KASHANYARAZI	1	32
Western	Rukungiri	Rujumbura	Kagunga	EA02	KAMURI	2	32
Western	Kyenjojo	Mwenge	Kyarusozi	EA05	MAKARRA	2	32
Central	Kampala	City Council	Nakawa Division	EA26	KISENYI II ZONE 'B'	1	32
Central	Luwero	Katikamu	Butuntumula	EA01	NGOGOLO	2	32
Central	Masaka	Kalungu	Bukulula	EA04	BUTONGO	2	32
Central	Mubende	Buwekula	Kiyuni	EA07	KIJUMBA EAST	2	32
Central	Mukono	Mukono	Goma	EA03	GONGOBE	1	32
Central	Rakai	Kyotera	Kalisizo	EA08	LUWAWULO	2	32
Central	Wakiso	Busiiro	Nsangi	EA08	KIVU	2	32
Eastern	Bugiri	Bukooli	Buwunga	EA01	BUSOWA CENTRAL 'A'	2	32
Eastern	Iganga	Kigulu	Nakalama	EA06	BUKOBOOLI 'A'	2	32
Eastern	Kamuli	Budiope	Kidera	EA08	NAKAWA B 'B'	2	32
Eastern	Kumi	Bukedea	Malera	EA01	KOTIOKOT	2	32
Eastern	Mbale	Manjiya	Bukigai	EA04	BUNAMOSO	2	32
Eastern	Soroti	Soroti	Asuret	EA01	ORAJAI	2	32
Eastern	Mayuge	Bunya	Malongo	EA10	NAMADHI C	2	32
Total sample							800

Note: Ninety-six (i.e. 12 percent) of the sample households reside in urban areas, the remainder in rural areas. This is consistent with the urban share in Uganda's population.

Annex A, Table 2. Equivalence scales and daily calorific requirements

Age	Male (kcal)	Female (kcal)	Equivalence scale
1	820	820	0.273
1-2	1150	1150	0.383
2-3	1350	1350	0.450
3-5	1550	1550	0.517
5-7	1850	1750	0.617
7-10	2100	1800	0.700
10-12	2200	1950	0.733
12-14	2400	2100	0.800
14-16	2650	2150	0.883
16-18	2850	2150	0.950
18-30	3000	2100	1
30-60	2900	2150	0.977
> 60	2450	1950	0.845

Source: WHO 1985, cited from Appleton et al. (1999)

Note: Age ranges are of the form x (inclusive) to y (exclusive) years e.g. 1 refers to below one year of age. Appleton et al. (1999) use WHO's requirements for "moderate" work as shown in the above table. It can be seen that the WHO recommends 3000 calories a day for men aged 18-30 engaged in moderate work. WHO's energy requirements differ by age and sex. Appleton et al (1999) allow for varying needs by age by using calorific equivalence scales as listed above. The equivalence scale for a person of a given age and sex category is set equal to the ratio of the recommended intake for a male of the relevant age divided by 3000, the requirements for the reference category of men aged 18-30. Appleton et al (1999) do not allow for sex differences in calculating the equivalence scales but use the WHO calorie requirements for males to derive equivalence scales which are then applied to both males and females. As this method for computing adult equivalence scales adopted by Appleton et al. (1999) and Appleton and Ssewanjana (2003) has been used in estimating the official poverty rates published by the Ugandan Government, we as well apply the same male-based equivalent scales for women and men alike (see third column) for the calculation of adult-equivalents in the IRIS data set.

Annex A, Table 3.1. Monthly consumption expenditures per adult equivalent (in 1997 prices), by survey year and decile

Decile	1992/93	1997 (MS-4)	1999/00	2002/03	September 2004 (IRIS sample)
1	8745	11119	12415	11938	8920
2	11376	14036	16246	15363	14510
3	13917	16875	19599	18413	18566
4	16410	19603	22751	21675	22348
5	19197	22778	26498	25125	27483
6	22484	26041	30662	29359	33345
7	26695	31547	36036	35005	39008
8	32289	37928	44369	44911	49272
9	43015	49829	62725	65665	64680

Source: For year 2002/03, data provided by Economic Policy Research Centre, Makerere University, Kampala. For prior years, see Appleton et al. (1999).

Note: For the above table, the expenditures from the IRIS data were deflated based on changes in the consumer price index between June 1997 and September 2004.

Annex A, Table 3.2. National poverty line (in 1997 prices), expressed as monthly expenditure per adult equivalent

Region	Poverty line Ug.Sh. (in 1997 prices)
CENTRAL RURAL	21322
CENTRAL URBAN	23150
EASTERN RURAL	20652
EASTERN URBAN	22125
WESTERN RURAL	20872
WESTERN URBAN	21800
NORTHERN RURAL	20308
NORTHERN URBAN	21626

Source: Data provided by Economic Policy Research Centre, Makerere University, Kampala.

Note: For the analysis of the IRIS data, the above poverty lines which are expressed in 1997 prices were inflated using the inflation rate between June 1997 and September 2004 so as to calculate the poverty lines valid for September 2004.

ANNEX B. DESCRIPTORS OF ALL REGRESSORS

Annex B, Table 1. Descriptive statistics of all regressors (n= 257), by type of model

Variable Label	Min.	Max.	Mean	St. Dev.	M1	M2	M3	M4	M5	M6	M7	M8	M9
Household size	1	18	5.83	2.97	X	X		X	X	X	X	X	X
Household size squared	1	324	42.81	42.67	X	X		X	X	X	X	X	X
Age of household head	18	102	43.09	15.43	X	X		X	X	X	X	X	X
Western Region	0	1	0.32	0.47	X	X		X	X	X	X	X	X
Northern Region	0	1	0.12	0.33	X	X		X	X	X	X	X	X
Eastern Region	0	1	0.27	0.45	X	X		X	X	X	X	X	X
Urban location	0	1	0.12	0.33	X	X		X	X	X	X	X	X
Head of household has any account	0	1	0.11	0.31	X	X		X	X	X	X	X	X
Head and spouse have a joint account	0	1	0.02	0.12	X	X		X	X	X	X	X	X
Total agricultural assets owned (out of 6)	0	6	2.28	1.77	X	X		X	X	X	X	X	X
Average age of adult household members	19	87	37.39	12.60	X	X		X	X	X	X	X	X
Squared age of household head	324	10404	2094.17	1523.38	X	X		X	X	X	X	X	X
Age of oldest household member	18	102	45.01	16.67	X	X		X	X	X	X	X	
Age of youngest household member	1	87	8.15	13.38	X	X		X	X	X	X	X	
Average of household members, except head	0.5	87	13.96	11.09	X	X		X	X	X	X	X	
Bicycle ownership	0	1	0.40	0.49	X	X		X	X	X			X
Blanket ownership	0	1	0.83	0.38	X	X		X	X	X			X
Household had access to formal loans in the past	0	1	0.05	0.22	X	X		X	X	X	X	X	X
The percentage of amount you would spend on food if you were given an additional 100 Ug.Sh. tomorrow?	0.08	500	82.03	29.73	X	X		X					
Cars ownership	0	1	0.01	0.10	X	X		X	X	X	X	X	X
CD player ownership	0	1	0.01	0.08	X	X		X	X	X	X	X	X
Contribution other groups	0	1	0.01	0.09	X	X		X	X	X	X	X	
Contribution traders association	0	1	0.00	0.04	X	X		X	X	X	X	X	
What is the size of these rooms in square feet?	13	5600	329.31	365.45	X	X		X	X	X	X	X	X

Variable Label	Min.	Max.	Mean	St. Dev.	M1	M2	M3	M4	M5	M6	M7	M8	M9
Have you made a recent home improvement in the last three years?	0	1	0.36	0.48	X	X		X	X	X			
Do you have Mobile (cell phone) in the house?	0	1	0.09	0.28	X	X		X	X	X			X
Head has a daily agricultural labor	0	1	0.05	0.22	X	X		X	X	X	X	X	X
Household sometimes ate beer type bananas because other food was scarce	0	1	0.02	0.15	X	X		X					
Head of household sleeps on floor or thin sleeping mat	0	1	0.18	0.38	X	X		X	X	X			
Head of household sleeps thick mattress	0	1	0.12	0.33	X	X		X	X	X			
Household sometimes ate bukupa because other food was scarce	0	1	0.05	0.22	X	X		X			X	X	
Household mostly ate bukupa because other food was scarce	0	1	0.05	0.22	X	X		X			X	X	
Household rarely ate cassava because other food was scarce	0	1	0.13	0.33	X	X		X					
Cooking fuel is leaves, husk, or cow dung	0	1	0.02	0.14	X	X		X	X	X	X	X	X
Cooking fuel is charcoal or paraffin	0	1	0.13	0.34	X	X		X	X	X	X	X	X
Household often didn't have enough food	0	1	0.17	0.37	X	X		X					
Percentage of dependents younger than 14 and older than 60 years (in relation to household size)	0	100	48.66	23.45	X	X		X	X	X	X	X	
Percentage of dependents younger than 15 and older than 64 years (in relation to household size)	0	100	50.45	23.49	X	X		X	X	X	X	X	
Percentage of dependents younger than 18 and older than 60 years (in relation to household size)	0	100	56.87	22.83	X	X		X	X	X	X	X	
Flooring material is cow dung	0	1	0.28	0.45	X	X		X	X	X	X	X	X
Floor is brick/stone, cement or cement with additional covering	0	1	0.26	0.44	X	X		X	X	X	X	X	X
Do you agree that people are always interested only in their own welfare?	0	1	0.66	0.47	X	X		X					
Do you agree that if you have a problem, there is always someone to help you?	0	1	0.81	0.40	X	X		X					
Do you agree that you feel accepted as a member of this village/neighborhood?	0	1	0.94	0.25	X	X		X					
Household feels that food expenses are above need	0	1	0.27	0.44	X	X		X			X	X	

Variable Label	Min.	Max.	Mean	St. Dev.	M1	M2	M3	M4	M5	M6	M7	M8	M9
Household feels that clothing expenses are average	0	1	0.18	0.39	X	X		X			X	X	
Household feels that health care expenses are average	0	1	0.17	0.37	X	X		X					
Household feels that child educational expenses are average	0	1	0.17	0.37	X	X		X			X	X	
Household feels that housing expenses are above need	0	1	0.13	0.34	X	X		X			X	X	
Household feels living standard improved significantly compared to 1996	0	1	0.05	0.21	X	X		X			X	X	
Household cooks in one of the rooms in the house	0	1	0.07	0.25	X	X		X	X	X	X	X	
Lighting source: Gas lamp or electricity (neighbor, public or own socket)	0	1	0.09	0.28	X	X		X	X	X	X	X	X
No lock in main entrance door or wood or metal bar to close from inside	0	1	0.42	0.49	X	X		X	X	X	X	X	
Household never ate maize meals because other food was scarce	0	1	0.46	0.50	X	X		X			X	X	
Household rarely ate maize meals because other food was scarce	0	1	0.16	0.36	X	X		X					
Household sometimes ate maize meals because other food was scarce	0	1	0.13	0.34	X	X		X					
Household often ate maize meals because other food was scarce	0	1	0.08	0.27	X	X		X					
Household rarely ate mukene because other food was scarce	0	1	0.26	0.44	X	X		X					
Household sometimes ate mukene because other food was scarce	0	1	0.31	0.46	X	X		X					
Household ate mostly mukene because other food was scarce	0	1	0.00	0.04	X	X		X					
Household borrows from neighbors/relatives rarely or sometimes	0	1	0.01	0.08	X	X		X					
Household head is domestic worker	0	1	0.08	0.27	X	X		X	X	X	X	X	X
Donkey ownership	0	1	0.31	0.46	X	X		X	X	X	X	X	X
Other natural disaster had not much or some but manageable damage	0	1	0.35	0.48	X	X		X					
Quality of walls: good	0	1	0.22	0.41	X	X		X	X	X	X	X	

Variable Label	Min.	Max.	Mean	St. Dev.	M1	M2	M3	M4	M5	M6	M7	M8	M9
Roof with banana leaves, fiber, grass, bamboo or wood	0	1	0.21	0.40	X	X		X	X	X	X	X	X
House size: large	0	1	0.06	0.24	X	X		X	X	X	X	X	
Household had to skip meals less than 30 days during the past 12 months	0	1	0.13	0.34	X	X		X					
Household mostly ate sorghum because other food was scarce	0	1	0.21	0.41	X	X		X					
Household usually purchases staple food fortnightly or monthly	0	1	0.13	0.33	X	X		X					
Household usually purchases staple food twice a week	0	1	0.23	0.42	X	X		X					
Household usually purchases staple food once a week	0	1	0.16	0.37	X	X		X					
Household usually purchases staple food less frequently than monthly	0	1	0.08	0.28	X	X		X					
Household never ate sweet potato because other food was scarce	0	1	0.28	0.45	X	X		X					
Toilet: shared or own ventilated, improved latrine or flush toilet	0	1	0.10	0.30	X	X		X	X	X	X	X	X
Water source: Open spring/well	0	10	4.27	1.49	X	X		X	X	X	X	X	X
Water source: Private borehole or piped water	0	7	0.64	1.07	X	X		X	X	X	X	X	X
How many meals were served to the household members during the last 2 days?	0	7	4.10	2.36	X	X		X			X	X	
Number of days in the last 7 days where a main meal consisted on large fish	0	30	6.37	9.20	X	X		X					
Number of days in the last 7 days where a main meal consisted on staple food, plant protein and any vegetables	0	162	11.36	18.91	X	X		X					
In the last 30 days, for how many days did your household not have enough to eat everyday?	0	1	0.28	0.45	X	X		X					
For how many weeks will your stock of staple food last?	0	1	0.35	0.48	X	X		X					
In last 12 months did you or any other adult in your household stop eating meals for an entire day because you did not have enough money to buy food?	0	1	0.26	0.44	X	X		X				X	
Did you or any other adult household member lose weight in last 12 months because you did not have enough money to buy food?	0	1	0.31	0.46	X	X		X					

Variable Label	Min.	Max.	Mean	St. Dev.	M1	M2	M3	M4	M5	M6	M7	M8	M9
Do you have access to electricity in this community?	0	1	0.32	0.47	X	X		X	X	X	X	X	
How far away the access to electricity (km)?	0	64	12.19	17.60	X	X		X	X	X	X	X	
Whether household head owns any type of shoes	0	1	0.87	0.34	X	X		X	X	X	X	X	X
Whether spouse owns any type of shoes	0	1	0.66	0.48	X	X		X	X	X	X	X	X
Head of household is farmer	0	1	0.57	0.49	X	X		X	X	X	X	X	X
Number of female adult household members	0	5	1.19	0.69	X	X		X	X	X	X	X	
Total females in household	0	11	3.00	1.80	X	X		X	X	X	X	X	X
Number of members, relatives working in Kampala and sending money	0	2	0.10	0.34	X	X		X	X	X			
Food expenditure as share from total household expenditures	0	1.88	1.22	0.39	X								
Number of members, relatives working abroad (foreign country) and sending money	0	2	0.02	0.13	X	X		X	X	X			
In last 3 years, how many marriages of a child belonging to the household occurred?	0	3	0.17	0.50	X	X		X	X	X	X	X	
In last 3 years how many adoptions of a child occurred?	0	14	0.43	1.18	X	X		X	X	X	X	X	
Serious (but not chronic) illness of a second working adult member in last 3 years	0	60	0.71	3.58	X	X		X			X	X	
Occurrence of a serious chronic illness or major disability of any household member in last 3 years	0	1	0.07	0.26	X	X		X			X	X	
Death of a dependent member (Household member) in last 3 years	0	6	0.33	0.72	X	X		X	X	X	X	X	
Relocation of residence/house because of violence in last 3 years	0	1	0.04	0.19	X	X		X	X	X	X	X	
Relocation of residence/household because of other reasons such as natural disasters in last 3 years	0	1	0.04	0.20	X	X		X	X	X	X	X	
Did your household have a very serious problem or failure in your own crop production in last 3 years?	0	1	0.58	0.49	X	X		X			X	X	
Did your household have a very serious problem or failure in your own animal production in the last 3 years?	0	1	0.34	0.47	X	X		X			X	X	
Did your household become a victim of severe robbery in the last 3 years?	0	1	0.16	0.37	X	X		X	X	X	X	X	
Total household members in civic group	0	1	0.01	0.08	X	X		X	X	X	X	X	

Variable Label	Min.	Max.	Mean	St. Dev.	M1	M2	M3	M4	M5	M6	M7	M8	M9
Household belongs to cultural group	0	1	0.00	0.05	X	X		X	X	X	X	X	
Total household members in farmers group	0	2	0.04	0.22	X	X		X	X	X	X	X	
Total household members in neighbors association	0	4	0.08	0.37	X	X		X	X	X	X	X	
Household belongs to religious group	0	1	0.01	0.10	X	X		X	X	X	X	X	
Household belongs to women's group	0	1	0.02	0.13	X	X		X	X	X	X	X	
Are you or members of household denied service or do you/they have only limited opportunity to sanitation services?	0	1	0.09	0.29	X	X		X			X	X	
Are you or members of household denied service or do you/they have only limited opportunity to agricultural extension?	0	1	0.13	0.33	X	X		X					
Are you or members of household denied service or do you/they have only limited opportunity to security/police services?	0	1	0.10	0.30	X	X		X					
Did households have access to GOAPS?	0	1	0.16	0.37	X	X		X	X	X			
Goat ownership	0	1	0.36	0.48	X	X		X	X	X	X	X	X
Where on the ladder would you locate a household, who has an income equal to Ug.Sh. 70,000 per month?	1	10	4.77	2.47	X	X		X				X	
Do you have immunization center in your community?	0	1	0.64	0.48	X	X		X	X	X	X	X	
Do you or your partner/spouse or anyone else in your household have a withdrawable savings account?	0	1	0.13	0.34	X	X		X	X	X	X	X	X
Change in number of steps from 7 years ago to today, if minus worsened	-7	7	-0.15	2.13	X	X		X				X	
Average daily per-capita clothing expenditures	-4.05	6.17	2.41	2.37	X	X						X	X
Leather shoes ownership	0	1	0.65	0.48	X	X		X	X	X			X
Leather shoes (spouse) ownership	0	1	0.50	0.50	X	X		X	X	X			X
Value of homestead	5.003	18.826	12.70	2.06	X	X		X	X		X	X	X
Value of agricultural land	6.098	19.114	12.60	2.93	X	X		X	X		X	X	X
Annualized food expenditures, recall 1 week	7.243	16.724	13.86	0.82	X								
Number of literate female adults in household	0	5	0.71	0.80	X	X		X	X	X			X
Number of literate male adults in household	0	5	0.87	0.84	X	X		X	X	X			

Variable Label	Min.	Max.	Mean	St. Dev.	M1	M2	M3	M4	M5	M6	M7	M8	M9
Homestead and agricultural land price	6.74	19.67	13.91	1.91	X	X		X	X		X	X	X
Remittances received	-2.46	26.71	0.02	6.19	X	X		X	X				X
Weekly expenditures for buying food	1.73	11.29	8.74	1.53	X								
Monthly expenditures on utilities	-4.16	12.21	4.09	4.42	X								
Monthly expenditures on transport	1.26	13.02	6.80	3.58	X	X		X	X	X			
Expenditures on school/education, last 12 months	3.19	15.61	9.70	3.08	X							X	
Expenditures on health, last 12 months	3.66	15.23	10.62	1.70	X								
Expenditures on home, last 12 months	0.38	15.89	5.93	3.78	X								
Remittances sent to relatives, last 12 months	1.10	14.73	6.76	3.79	X								X
Expenditures on other expenditures (social events, leisure), last 12 months	0.69	15.20	8.00	3.22	X	X						X	
Sum of household clothing expenditures in past 12 months	4.49	13.68	9.92	2.44	X								X
Costs of home improvements	5.34	15.89	7.25	2.75	X	X		X	X				
Do you have access to mainline phone in your community?	0	1	0.20	0.40	X	X		X	X	X	X	X	
Value of inherited major funds or assets, past 3 years	0	17.91	4.93	1.41	X	X		X	X				
Amount household needs per month to live	6.91	14.91	11.75	1.01	X	X		X			X	X	
Value of jewelry	0.48	14.22	6.64	3.75	X	X		X	X				
Debt owed to other households by your household at present	-1.27	13.46	5.14	3.74	X	X		X	X				
Annualized non food expenditures 2 (education, health, and others)	6.12	15.92	12.19	1.39	X								
Value of total savings first respondent	6.75	13.66	7.29	1.60	X	X		X	X				X
Value of total savings respondent and spouse	6.95	15.18	7.04	0.74	X	X		X	X				X
Value of household members total savings	7.69	15.18	8.24	1.47	X	X		X	X				X
Total value of all animals	5.64	16.52	9.68	3.06	X	X		X	X		X	X	X
Value of bicycles	3.08	12.47	6.18	3.79	X	X		X	X		X	X	X
Value of blanket	3.15	12.61	8.56	2.64	X	X		X	X		X	X	X
Value of boats	-2.30	11.16	-2.27	0.63	X	X		X	X		X	X	X
Value of local cattle	5.21	16.52	6.89	3.24	X	X		X	X		X	X	X

Variable Label	Min.	Max.	Mean	St. Dev.	M1	M2	M3	M4	M5	M6	M7	M8	M9
Value of CD players	0.49	12.90	0.56	0.95	X	X		X	X		X	X	X
Value of color TVs	2.10	13.46	5.80	1.15	X	X		X	X		X	X	X
Value of donkeys	-2.76	9.90	-2.69	0.88	X	X		X	X		X	X	X
Value of food processing assets	-3.28	10.31	-3.26	0.48	X	X		X	X		X	X	X
Value of goat	3.39	13.59	6.12	3.66	X	X		X	X		X	X	X
Value of gomesi	-7.20	14.00	5.47	8.04	X	X		X	X		X	X	X
Value of leather shoes	2.64	12.61	7.06	3.37	X	X		X	X		X	X	X
Value of leather shoes (spouse)	0	13.017	5.69	3.58	X	X		X	X		X	X	X
Value of metal pots	2.77	14.00	8.41	2.31	X	X		X	X		X	X	X
Value of mosquito nets	1.05	11.92	3.09	3.49	X	X		X	X		X	X	X
Value of motorcycles	3.29	14.51	3.57	1.64	X	X		X	X		X	X	X
Value of other vehicles	-2.23	11.00	-2.18	0.75	X	X		X	X		X	X	X
Value of ox-implements	1.06	13.59	1.35	1.66	X	X		X	X		X	X	X
Value of panga	0.46	9.95	5.04	3.47	X	X		X	X		X	X	X
Value of rubber shoes	0.30	10.71	1.75	3.19	X	X		X	X		X	X	X
Value of tire shoes	-2.08	10.13	-1.89	1.38	X	X		X	X		X	X	X
Total amount borrowed from formal institutions	2.96	14.51	3.38	1.98	X	X		X	X				X
Number of days in past 7 days any of 4 superior food eaten (max. 28)	0	18	4.97	3.42	X	X		X				X	
Value of transport assets	4.95	17.11	7.54	3.10	X	X		X	X		X	X	X
Number of male adults in household	0	5	1.15	0.82	X	X		X	X	X	X	X	X
Ratio of male to females	0	8	1.17	0.99	X	X		X	X	X	X	X	X
Total males in household	0	11	2.83	1.85	X	X		X	X	X	X	X	X
Do you have market in your community?	0	1	0.16	0.37	X	X		X	X	X	X	X	
How far away the bodaboda (motorcycle) station (km)?	0	24	4.45	5.79	X	X		X	X	X	X	X	
Metal cooking pots ownership	0	1	0.88	0.32	X	X		X	X	X	X	X	X
How far away the private health center with midwife/nurse (km)?	0	24	4.86	5.79	X	X		X	X	X	X	X	
How far away the access to network phone (km)?	0	80	9.09	17.25	X	X		X	X	X	X	X	

Variable Label	Min.	Max.	Mean	St. Dev.	M1	M2	M3	M4	M5	M6	M7	M8	M9
Number of bicycles owned by the household	0	10	0.45	0.67	X	X		X	X	X	X	X	X
Number of black/white TVs owned by the household	0	1	0.03	0.18	X	X		X	X	X	X	X	X
Number of donkeys owned by the household	0	7	0.03	0.36	X	X		X	X	X	X	X	X
Household has any source of electricity in house	0	1	0.26	0.44	X	X		X	X	X	X	X	X
Number of goats owned by the household	0	32	1.32	2.83	X	X		X	X	X	X	X	X
Number of gomesi owned by the household	0	21	2.17	2.33	X	X		X	X	X			X
Number of hand hoes owned by the household	0	13	2.41	1.99	X	X		X	X	X	X	X	X
Number of leather shoes owned by the household	0	10	1.10	1.23	X	X		X	X	X	X	X	X
Number of local cattle owned by the household	0	50	0.79	2.85	X	X		X	X	X	X	X	X
Number of motorcycles owned by the household	0	2	0.03	0.18	X	X		X	X	X	X	X	X
Household declares to not have a savings habit	0	1	0.11	0.32	X	X		X			X	X	
Number of ox-implements owned by the household	0	5	0.05	0.34	X	X		X	X	X	X	X	X
Number of poultry owned by the household	0	230	4.73	16.22	X	X		X	X	X	X	X	X
Number of radios owned by the household	0	4	0.69	0.60	X	X		X	X	X	X	X	X
Number of ram presses owned by the household	0	2	0.00	0.08	X	X		X	X	X	X	X	X
Number of rubber shoes (spouse) owned	0	2	0.05	0.23	X	X		X	X	X	X	X	X
Number of sheep owned by the household	0	10	0.16	0.86	X	X		X	X	X	X	X	X
Number of sleepers owned by the household	0	169	1.16	8.03	X	X		X	X	X			X
Number of tire shoes (spouse) owned	0	2	0.03	0.19	X	X		X	X	X	X	X	X
Do you have nursery/pre primary center in your community?	0	1	0.36	0.48	X	X		X	X	X	X	X	
During the past 5 years, did you face other disaster?	0	1	0.24	0.42	X	X		X					
Other vehicle ownership	0	1	0.00	0.06	X	X		X	X	X	X	X	X
Degree of participation in health group	0	1	0.00	0.05	X	X		X			X	X	
Degree of participation in neighbors group	0	1	0.06	0.23	X	X		X			X	X	
Panga ownership	0	1	0.64	0.48	X	X		X	X	X	X	X	X
Degree of participation in NGO for entrepreneurial services	0	1	0.00	0.04	X	X		X			X	X	
Degree of participation in other NGOs	0	1	0.01	0.07	X	X		X					

Variable Label	Min.	Max.	Mean	St. Dev.	M1	M2	M3	M4	M5	M6	M7	M8	M9
Degree of participation in parents group	0	1	0.01	0.09	X	X		X			X	X	
Degree of participation in religious group	0	1	0.01	0.09	X	X		X			X	X	
Sum of degree of participation out of 21 institutions	0	2	0.19	0.41	X	X		X			X	X	
Household is active in at least one institution	0	1	0.18	0.38	X	X		X			X	X	
Degree of participation in women group	0	1	0.02	0.12	X	X		X			X	X	
Degree of participation in youth group	0	1	0.00	0.06	X	X		X			X	X	
Poultry ownership	0	1	0.52	0.50	X	X		X	X	X	X	X	X
How far away the private health center with a regular doctor (km)?	0	24	5.51	6.00	X	X		X	X	X	X	X	
Do you have primary school in your community?	0	1	0.40	0.49	X	X		X	X	X	X	X	
Ram press ownership	0	1	0.00	0.05	X	X		X	X	X	X	X	X
Dependency ratio younger than 15 or older than 64 years	0	8	1.33	1.05	X	X		X	X	X	X	X	
Dependency ratio younger than 18 or older than 60 years	0	8	1.73	1.31	X	X		X	X	X	X	X	
Percentage of adult household members who read only (in relation to household size)	0	100	2.82	9.66	X	X		X	X	X	X	X	X
Household head can read only	0	1	0.06	0.24	X	X		X	X	X			X
Number of household members who can read only	0	8	0.31	0.73	X	X		X	X	X	X	X	X
Spouse can read only	0	1	0.05	0.22	X	X		X	X	X			X
Remittances received/ total household expenditures	-0.32	1.32	-0.12	0.37	X								X
Remittances sent/ total household expenditures	0.08	0.96	0.47	0.26	X	X		X	X	X			X
Ratio of male adults to female adults	0	5	1.00	0.73	X	X		X	X	X	X	X	
Rooms per person	0.07	6	0.69	0.69	X	X		X	X	X	X	X	X
Rubber/plastic shoes ownership	0	1	0.17	0.38	X	X		X	X	X			X
Total number of days sick by females	0	1131	65.82	116.28	X	X		X	X	X			X
Average number of days sick by females	0	365	22.07	42.26	X	X		X	X	X			X
Household head is single	0	1	0.07	0.25	X	X		X	X	X	X	X	X
Head is student	0	1	0.00	0.04	X	X		X	X	X	X	X	X
Total land area= agricultural land + homestead	0	50001	72.73	1784.18	X	X		X	X	X	X	X	X
Household head is self-employed in trade	0	1	0.07	0.25	X	X		X	X	X	X	X	X

Variable Label	Min.	Max.	Mean	St. Dev.	M1	M2	M3	M4	M5	M6	M7	M8	M9
Household head is unemployed	0	1	0.01	0.10	X	X		X	X	X	X	X	X
Household head is widow	0	1	0.14	0.34	X	X		X	X	X	X	X	X
Wood shoes ownership	0	1	0.02	0.14	X	X		X	X	X	X	X	X
Household head has no schooling or did not complete grade1	0	1	0.21	0.41	X	X		X	X	X	X	X	
Household head: incomplete secondary education	0	1	0.20	0.40	X	X		X	X	X	X	X	X
Household head completed superior education	0	1	0.03	0.16	X	X		X	X	X	X	X	X
Number of household members with incomplete secondary education, excluding head	0	9	0.50	0.98	X	X		X	X	X	X	X	X
Number household members who completed secondary/post primary education only, excluding head	0	2	0.06	0.27	X	X		X	X	X	X	X	X
Median education of household members is primary level (complete)	0	1	0.05	0.21	X	X		X	X	X	X	X	
Median education of household members is secondary/post primary/J1 level (complete)	0	1	0.02	0.14	X	X		X	X	X	X	X	
Median education of household members is post secondary/superior level (complete)	0	1	0.01	0.07	X	X		X	X	X	X	X	
Maximum education of any household member is no schooling/not complete grade1	0	1	0.05	0.21	X	X		X	X	X	X	X	
Maximum education of any household member is post secondary/superior level (complete)	0	1	0.05	0.21	X	X		X	X	X	X	X	
Median education of adult household members is secondary/post primary/J1 level (complete)	0	1	0.06	0.23	X	X		X	X	X	X	X	
Maximum education of any adult is not schooling/not complete grade one	0	1	0.12	0.33	X	X		X	X	X	X	X	
Education level of spouse is not schooling/not complete grade one	0	1	0.20	0.40	X	X		X	X	X	X	X	
Education level of spouse is post secondary/superior level (complete)	0	1	0.01	0.08	X	X		X	X	X	X	X	
Median education of household females is secondary/post primary/J1 level (complete)	0	1	0.05	0.23	X	X		X	X	X	X	X	
Median education of household males is primary level (complete)	0	1	0.11	0.31	X	X		X	X	X	X	X	

Variable Label	Min.	Max.	Mean	St. Dev.	M1	M2	M3	M4	M5	M6	M7	M8	M9
Median education of household males is post secondary/superior level (complete)	0	1	0.01	0.09	X	X		X	X	X	X	X	
Maximum education of males is not schooling/not complete grade one	0	1	0.13	0.34	X	X		X	X	X	X	X	
Maximum education of males is secondary/post primary/J1 level (complete)	0	1	0.06	0.24	X	X		X	X	X	X	X	
Maximum education of males is post secondary/superior level (complete)	0	1	0.03	0.18	X	X		X	X	X	X	X	

ANNEX C. GENDER-SPECIFIC VARIABLES

Annex C, Table 1. Gender-specific variables used in regression analyses

Variable Label	Min.	Max.	Mean	St. Dev.
Whether spouse owns any type of shoes	0	1	0.66	0.48
Number of female adult household members	0	5	1.19	0.69
Total number of females in household	0	11	3.00	1.80
Household belongs to women's group	0	1	0.02	0.13
Leather shoes (spouse) ownership	0	1	0.50	0.50
Number of literate female adults in household	0	5	0.71	0.80
Number of literate male adults in household	0	5	0.87	0.84
Value of leather shoes (spouse)	0	13.02	5.69	3.58
Number of male adults in household	0	5	1.15	0.82
Ratio of male to females	0	8	1.17	0.99
Total number of males in household	0	11	2.83	1.85
Degree of participation in women group	0	1	0.02	0.12
Spouse can read only	0	1	0.05	0.22
Ratio of male adults to female adults	0	5	1.00	0.73
Total number of days sick by females	0	1131	65.82	116.28
Average number of days sick by females	0	365	22.07	42.26
Education level of spouse is not schooling/not completed grade one	0	1	0.20	0.40
Education level of spouse is post secondary/superior level (completed)	0	1	0.01	0.08
Median education of household females is secondary/post primary/J1 level (completed)	0	1	0.05	0.23

ANNEX D. VERIFIABILITY SCORES PROVIDED BY NIDA

Variable assessment scale: 1=very hard, 5=easily verifiable

Note: The indicators with verifiability scores of 4 or 5 have been included in Model 7 and 8

Annex D, Table 1. Verifiability score of the variables

Variable	Measurement	Verifiability	Difficulty to ask
Head of household has any account	Yes/No	4	5
Head and spouse have a joint account	Yes/No	4	5
Total agricultural assets owned (out of 6)	Number	4	5
Average age of adult household members	Years	4	5
Squared age of household head	Years	4	5
Age of oldest household member	Years	4	5
Age of youngest household member	Years	4	5
Average of household members, except head	Years	4	5
Bicycle ownership	Yes/No	3	5

Variable	Measurement	Verifiability	Difficulty to ask
Blanket ownership	Yes/No	3	5
Household had access to formal loans in the past	Yes/No	5	5
The percentage of amount you would spend on food if you were given an additional 100 Ug.Sh. tomorrow?	Percentage	1	3
Cars ownership	Yes/No	4	5
CD player ownership	Yes/No	4	5
Contribution other groups	Cash or kind/ None	4	5
Contribution traders association	Cash or kind/ None	4	5
What is the size of these rooms in square feet?	Sq. ft.	4	3
Have you made a recent home improvement in the last three years?	Yes/No	3	5
Do you have Mobile (cell phone) in the house?	Yes/No	3	5
Head has a daily agricultural labor	Yes/No	5	5
Household sometimes ate beer type bananas because other food was scarce	Yes/No	3	4
Head of household sleeps on floor or thin sleeping mat	Yes/No	2	1
Head of household sleeps thick mattress	Yes/No	2	4
Household sometimes ate bukupa because other food was scarce	Yes/No	4	3
Household mostly ate bukupa because other food was scarce	Yes/No	4	3
Household rarely ate cassava because other food was scarce	Yes/No	2	5
Cooking fuel is leaves, husk, or cow dung	Yes/No	4	4
Cooking fuel is charcoal or paraffin	Yes/No	4	5
Household often didn't have enough food	Yes/No	3	4
Percentage of dependents younger than 14 and older than 60 years (in relation to household size)	Percentage	4	5
Percentage of dependents younger than 15 and older than 64 years (in relation to household size)	Percentage	4	5
Percentage of dependents younger than 18 and older than 60 years (in relation to household size)	Percentage	4	5
Flooring material is cow dung	Yes/No	5	5
Floor is brick/stone, cement or cement with additional covering	Yes/No	5	5
Do you agree that people are always interested only in their own welfare?	Yes/No	3	5
Do you agree that if you have a problem, there is always someone to help you?	Yes/No	2	5
Do you agree that you feel accepted as a member of this village/neighborhood?	Yes/No	1	5
Household feels that food expenses are above need	Yes/No	4	5
Household feels that clothing expenses are average	Yes/No	4	5
Household feels that health care expenses are average	Yes/No	3	5
Household feels that child educational expenses are average	Yes/No	4	5
Household feels that housing expenses are above need	Yes/No	4	5

Variable	Measurement	Verifiability	Difficulty to ask
Household feels living standard improved significantly compared to 1996	Yes/No	4	5
Household cooks in one of the rooms in the house	Yes/No	5	5
Lighting source: Gas lamp or electricity (neighbor, public or own socket)	Yes/No	5	2
No lock in main entrance door or wood or metal bar to close from inside	Yes/No	5	2
Household never ate maize meals because other food was scarce	Yes/No	4	5
Household rarely ate maize meals because other food was scarce	Yes/No	3	5
Household sometimes ate maize meals because other food was scarce	Yes/No	3	5
Household often ate maize meals because other food was scarce	Yes/No	3	5
Household rarely ate mukene because other food was scarce	Yes/No	3	5
Household sometimes ate mukene because other food was scarce	Yes/No	2	4
Household ate mostly mukene because other food was scarce	Yes/No	2	4
Household borrows from neighbors/relatives rarely or sometimes	Yes/No	3	3
Household head is domestic worker	Yes/No	4	2
Donkey ownership	Yes/No	4	5
Other natural disaster had not much or some but manageable damage	Yes/No	3	3
Quality of walls: good	Yes/No	4	2
Roof with banana leaves, fiber, grass, bamboo or wood	Yes/No	5	2
House size: large	Yes/No	5	5
Household had to skip meals less than 30 days during the past 12 months	Yes/No	2	3
Household mostly ate sorghum because other food was scarce	Yes/No	2	5
Household usually purchases staple food fortnightly or monthly	Yes/No	2	5
Household usually purchases staple food twice a week	Yes/No	2	5
Household usually purchases staple food once a week	Yes/No	2	5
Household usually purchases staple food less frequently than monthly	Yes/No	2	4
Household never ate sweet potato because other food was scarce	Yes/No	2	5
Toilet: shared or own ventilated, improved latrine or flush toilet	Yes/No	4	4
Water source: Open spring/well	Yes/No	5	5
Water source: Private borehole or piped water	Yes/No	4	5
How many meals were served to the household members during the last 2 days?	Number	4	5
Number of days in the last 7 days where a main meal consisted of large fish	Number of days	3	4

Variable	Measurement	Verifiability	Difficulty to ask
Number of days in the last 7 days where a main meal consisted on staple food, plant protein and any vegetables	Number of days	3	5
In the last 30 days, for how many days did your household not have enough to eat everyday?	Number of days	3	4
For how many weeks will your stock of staple food last?	Number of weeks	3	5
In last 12 months did you or any other adult in your household stop eating meals for an entire day because you did not have enough money to buy food?	Yes/No	2	4
Did you or any other adult household member lose weight in last 12 months because you did not have enough money to buy food?	Yes/No	3	5
Do you have access to electricity in this community?	Yes/No	5	5
How far away the access to electricity (km)?	Km.	5	5
Whether household head owns any type of shoes	Yes/No	4	4
Whether spouse owns any type of shoes	Yes/No	4	4
Head of household is farmer	Yes/No	5	5
Number of female adult household members	Number	5	5
Total females in household	Number	5	5
Number of members, relatives working in Kampala and sending money	Number	3	5
Food expenditure as share from total household expenditures	Number	3	5
Number of members, relatives working abroad (foreign country) and sending money	Number	2	5
In last 3 years, how many marriages of a child belonging to the household occurred?	Number	4	5
In last 3 years how many adoptions of a child occurred?	Number	5	5
Serious (but not chronic) illness of a second working adult member in last 3 years	Yes/No	5	5
Occurrence of a serious chronic illness or major disability of any household member in last 3 years	Yes/No	5	5
Death of a dependent member (Household member) in last 3 years	Number	5	4
Relocation of residence/house because of violence in last 3 years	Yes/No	5	5
Relocation of residence/household because of other reasons such as natural disasters in last 3 years	Yes/No	5	5
Did your household have a very serious problem or failure in your own crop production in last 3 years?	Yes/No	5	5
Did your household have a very serious problem or failure in your own animal production in the last 3 years?	Yes/No	5	5
Did your household become a victim of severe robbery in the last 3 years?	Yes/No	5	5
Total household members in civic group	Number	4	5
Household belongs to cultural group	Yes/No	4	5
Total household members in farmers group	Number	4	5

Variable	Measurement	Verifiability	Difficulty to ask
Total household members in neighbors association	Number	4	5
Household belongs to religious group	Yes/No	4	5
Household belongs to women's group	Yes/No	5	5
Are you or members of household denied service or do you/they have only limited opportunity to sanitation services?	Yes/No	5	5
Are you or members of household denied service or do you/they have only limited opportunity to agricultural extension?	Yes/No	3	5
Are you or members of household denied service or do you/they have only limited opportunity to security/police services?	Yes/No	3	5
Did households have access to GOAPS?	Yes/No	3	5
Goat ownership	Yes/No	4	5
Where on the ladder would you locate a household, who has an income equal to Ug.Sh. 70,000 per month?	Step in ladder	3	3
Do you have immunization center in your community?	Yes/No	5	5
Do you or your partner/spouse or anyone else in your household have a withdrawable savings account?	Yes/No	5	5
Change in number of steps from 7 years ago to today, if minus worsened	Step in ladder	3	3
Average daily per-capita clothing expenditures	Ug.Sh.	2	4
Leather shoes ownership	Yes/No	3	4
Leather shoes (spouse) ownership	Yes/No	3	4
Value of homestead	Ug.Sh.	4	2
Value of agricultural land	Ug.Sh.	4	3
Annualized food expenditures, recall 1 week	Ug.Sh.	3	5
Number of literate female adults in household	Number	3	4
Number of literate male adults in household	Number	3	4
Homestead and agricultural land price	Ug.Sh.	4	5
Remittances received	Ug.Sh.	2	5
Weekly expenditures for buying food	Ug.Sh.	3	5
Monthly expenditures on utilities	Ug.Sh.	3	4
Monthly expenditures on transport	Ug.Sh.	3	4
Expenditures in school/education, last 12 months	Ug.Sh.	3	4
Expenditures on health, last 12 months	Ug.Sh.	3	4
Expenditures on home, last 12 months	Ug.Sh.	2	4
Remittances sent to relatives, last 12 months	Ug.Sh.	3	5
Expenditures on other expenditures (social events, leisure), last 12months	Ug.Sh.	2	3
Sum of household clothing expenditures in past 12 months	Ug.Sh.	2	4
Costs of home improvements	Ug.Sh.	3	4
Do you have access to mainline phone in your community?	Yes/No	5	5
Value of inherited major funds or assets, past 3 years	Ug.Sh.	2	5
Amount household needs per month to live	Ug.Sh.	4	5

Variable	Measurement	Verifiability	Difficulty to ask
Value of jewelry	Ug.Sh.	2	5
Debt owed to other households by your household at present	Ug.Sh.	2	5
Annualized non food expenditures 2 (education, health, and others)	Ug.Sh.	3	4
Value of total savings first respondent	Ug.Sh.	1	3
Value of total savings respondent and spouse	Ug.Sh.	2	3
Value of household members total savings	Ug.Sh.	3	4
Total value of all animals	Ug.Sh.	4	4
Value of bicycles	Ug.Sh.	4	4
Value of blanket	Ug.Sh.	4	5
Value of boats	Ug.Sh.	4	5
Value of local cattle	Ug.Sh.	4	5
Value of CD players	Ug.Sh.	5	4
Value of color TVs	Ug.Sh.	5	5
Value of donkeys	Ug.Sh.	4	5
Value of food processing assets	Ug.Sh.	4	5
Value of goat	Ug.Sh.	4	5
Value of gomesi	Ug.Sh.	4	5
Value of leather shoes	Ug.Sh.	4	5
Value of leather shoes (spouse)	Ug.Sh.	4	5
Value of metal pots	Ug.Sh.	4	5
Value of mosquito nets	Ug.Sh.	4	5
Value of motorcycles	Ug.Sh.	4	5
Value of other vehicles	Ug.Sh.	4	5
Value of ox-implements	Ug.Sh.	5	5
Value of panga	Ug.Sh.	4	5
Value of rubber shoes	Ug.Sh.	4	5
Value of tire shoes	Ug.Sh.	4	5
Total amount borrowed from formal institutions	Ug.Sh.	2	4
Number of days in past 7 days any of 4 superior food eaten (max. 28)	Number of days	3	5
Value of transport assets	Ug.Sh.	4	5
Number of male adults in household	Number	5	5
Ratio of male to females	Number	5	5
Total males in household	Number	5	5
Do you have market in your community?	Yes/No	5	5
How far away the bodaboda (motorcycle) station (km)?	Km.	5	5
Metal cooking pots ownership	Yes/No	4	3
How far away the private health center with midwife/nurse (km)?	Km.	5	5
How far away the access to network phone (km)?	Km.	5	5
Number of bicycles owned by the household	Number	4	5

Variable	Measurement	Verifiability	Difficulty to ask
Number of black/white TVs owned by the household	Number	4	5
Number of donkeys owned by the household	Number	4	5
Household has any source of electricity in house	Yes/No	5	5
Number of goats owned by the household	Number	4	5
Number of gomesi owned by the household	Number	3	5
Number of hand hoes owned by the household	Number	4	5
Number of leather shoes owned by the household	Number	4	5
Number of local cattle owned by the household	Number	4	5
Number of motorcycles owned by the household	Number	4	5
Household declares to not have a savings habit	Yes/No	4	5
Number of ox-implements owned by the household	Number	5	5
Number of poultry owned by the household	Number	4	5
Number of radios owned by the household	Number	4	5
Number of ram presses owned by the household	Number	4	5
Number of rubber shoes (spouse) owned	Number	4	5
Number of sheep owned by the household	Number	4	5
Number of sleepers owned by the household	Number	3	5
Number of tire shoes (spouse) owned	Number	4	4
Do you have nursery/pre primary center in your community?	Yes/No	5	5
During the past 5 years, did you face other disaster?	Yes/No	3	3
Other vehicle ownership	Yes/No	4	5
Degree of participation in health group	Active/No active or non participant	4	5
Degree of participation in neighbors group	Active/No active or non participant	4	5
Panga ownership	Yes/No	4	5
Degree of participation in NGO for entrepreneurial services	Active/No active or non participant	4	5
Degree of participation in other NGOs	Active/No active or non participant	3	5
Degree of participation in parents group	Active/No active or non participant	4	5
Degree of participation in religious group	Active/No active or non participant	4	5
Sum of degree of participation out of 21 institutions	Number	4	5
Household is active in at least one institution	Yes/No	4	5
Degree of participation in women group	Active/No active or non participant	4	5
Degree of participation in youth group	Active/No active or non participant	4	5
Poultry ownership	Yes/No	5	5
How far away the private health center with a regular doctor (km)?	Km.	5	5
Do you have primary school in your community?	Yes/No	5	5
Ram press ownership	Yes/No	5	4

Variable	Measurement	Verifiability	Difficulty to ask
Dependency ratio younger than 15 or older than 64 years	Ratio	4	5
Dependency ratio younger than 18 or older than 60 years	Ratio	4	5
Percentage of adult household members who read only (in relation to household size)	Percentage	5	5
Household head can read only	Yes/No	3	4
Number of household members who can read only	Number	4	5
Spouse can read only	Yes/No	3	5
Remittances received/ total household expenditures	Ratio	3	5
Remittances sent/ total household expenditures	Ratio	3	5
Ratio of male adults to female adults	Ratio	5	5
Rooms per person	Number	4	3
Rubber/plastic shoes ownership	Yes/No	3	4
Total number of days sick by females	Number of days	3	3
Average number of days sick by females	Number of days	3	4
Household head is single	Yes/No	5	5
Head is student	Yes/No	5	5
Total land area= agricultural land + homestead	Acres	4	5
Household head is self-employed in trade	Yes/No	4	5
Household head is unemployed	Yes/No	4	5
Household head is widow	Yes/No	5	5
Wood shoes ownership	Yes/No	4	4
Household head has no schooling or did not complete grade1	Yes/No	4	4
Household head: incomplete secondary education	Yes/No	4	4
Household head completed superior education	Yes/No	4	4
Number of household members with incomplete secondary education, excluding head	Number	4	4
Number household members who completed secondary/post primary education only, excluding head	Number	4	4
Median education of household members is primary level (complete)	Yes/No	4	4
Median education of household members is secondary/post primary/J1 level (complete)	Yes/No	4	4
Median education of household members is post secondary/superior level (complete)	Yes/No	4	4
Maximum education of any household member is no schooling/not complete grade1	Yes/No	4	4
Maximum education of any household member is post secondary/superior level (complete)	Yes/No	4	4
Median education of adult household members is secondary/post primary/J1 level (complete)	Yes/No	4	4
Maximum education of any adult is not schooling/not complete grade one	Yes/No	4	4
Education level of spouse is not schooling/not complete grade one	Yes/No	4	4

Variable	Measurement	Verifiability	Difficulty to ask
Education level of spouse is post secondary/superior level (complete)	Yes/No	4	4
Median education of household females is secondary/post primary/J1 level (complete)	Yes/No	4	4
Median education of household males is primary level (complete)	Yes/No	4	4
Median education of household males is post secondary/superior level (complete)	Yes/No	4	4
Maximum education of males is not schooling/not complete grade one	Yes/No	4	4
Maximum education of males is secondary/post primary/J1 level (complete)	Yes/No	4	4
Maximum education of males is post secondary/superior level (complete)	Yes/No	4	4

Source: Personal communication with NIDA (2005). The ratings were derived by the survey staff of NIDA on a scale from 1 to 5 (after the survey had been completed). Higher values indicate improved verifiability and ease in asking information on a poverty indicator. The above scores represent consensus estimates by the NIDA survey team.

ANNEX E. SUMMARY OF RESULTS

Annex E, Table 1. Single-step OLS models with per-capita daily expenditures as continuous dependent variable

Model	Description	Type	Adj. R2	% Total Accuracy	% Poverty Accuracy	% Under-coverage	% Leakage	PIE (% points)	BPAC (% points)
1	All 257 regressors (Ref. Table 3.1.2)	B-5	0.517	76.39	58.82	41.18	31.76	-3.05	49.40
		B-10	0.561	78.05	61.96	38.04	29.8	-2.67	53.72
		B-15	0.587	79.06	63.53	36.47	28.23	-2.67	55.29
2	Exclusion of expenditure variables except average daily per-capita clothing expenditures and expenditures on others (social events, leisure) in the past 12 months (Ref. Table 3.2.1)	B-5	0.478	74.36	53.33	46.66	32.55	-4.57	39.22
		B-10	0.54	76.27	58.43	41.57	31.76	-3.17	48.62
		B-15	0.569	75.89	58.04	41.96	32.55	-3.05	48.63
3	Exclusion of total value of household assets	B-5							
		B-10							
		B-15							
4	Exclusion of average daily per-capita clothing expenditures and expenditures on others (social events, leisure) in the past 12 months. (Ref. Table 3.4.1)	B-5	0.477	74.75	53.33	46.67	31.37	-4.95	38.03
		B-10	0.535	76.4	57.65	42.35	30.59	-3.81	45.89
		B-15	0.564	78.93	61.18	38.82	26.28	-4.07	48.63
5	Exclusion subjective variables (Ref. Table 3.5.1)	B-5	0.477	74.75	53.33	46.67	31.37	-4.95	38.03
		B-10	0.527	76.65	56.47	43.53	28.63	-4.82	41.57
		B-15	0.553	78.17	59.61	40.39	27.05	-4.31	46.27
6	Exclusion monetary variables (Ref. Table 3.6.1)	B-5	0.471	74.36	52.55	47.45	31.76	-5.08	36.86
		B-10	0.523	76.4	56.08	43.92	29.02	-4.82	41.18
		B-15	0.55	78.81	62.74	37.25	28.23	-2.92	53.72
7	Easily verifiable variables (NIDA) (Ref. Table 3.7.1)	B-5	0.476	74.74	53.33	46.66	31.37	-4.95	38.04
		B-10	0.531	77.28	59.21	40.78	29.41	-3.68	47.84
		B-15	0.556	77.15	58.82	41.17	29.41	-3.81	47.06

8	Model 7 plus strong subjective and expenditure regressors (Ref. Table 3.8.1)	B-5	0.481	73.6	51.76	48.23	33.33	-4.83	36.86
		B-10	0.536	75.88	56.86	43.13	31.37	-3.81	45.10
		B-15	0.567	77.91	62.74	37.25	30.98	-2.03	56.47
9	LSMS-type regressors (Ref. Table 3.9.1)	B-5	0.474	73.6	51.76	48.23	33.33	-4.82	36.86
		B-10	0.519	74.49	53.72	46.27	32.55	-4.44	40.00
		B-15	0.535	74.36	54.51	45.49	33.72	-3.81	42.74

Note: The results that are marked in bold font indicate the best estimation in terms of PIE and BPAC within a model type, e.g. the best single-step OLS regression for the set of regressors identified as model type 9 is achieved by 15 regressors. One can see that the version with fifteen regressors had always the best PIE and BPAC values, except for Model 7. Moreover, only one of the nine models, namely Model 4 shows a trade-off between PIE and BPAC criterion such that increasing BPAC (when moving from ten to fifteen regressors) leads to a higher (absolute) value of PIE at the same time.

Annex E, Table 2. Two-step models with a continuous dependent variable (OLS estimation) or Models 1, 4, 7, and 9

OLS 2-Step Poverty rate: 32.36%	Adj. R2	% Total Accuracy	% Poverty Accuracy	% Under-coverage	% Leakage	PIE (% points)	BPAC (% points)
Model 1 — 51st Percentile	0.333 subsample	79.44	67.84	32.16	31.37	-0.25	67.06
Model 4 — 55th Percentile	0.302 subsample	80.08	67.45	32.55	29.02	-1.14	63.92
Model 7 — 57th Percentile	0.319 subsample	79.70	68.24	31.76	30.98	-0.25	67.45
Model 9 — 54th Percentile	0.210 subsample	77.79	63.92	36.08	32.55	-1.14	60.39

Annex E, Table 3. Summary results for all single and two-step regressions for Models 1, 4, 7, and 9

	Adj. R2	% Total Accuracy	% Poverty Accuracy	% Under-coverage	% Leakage	PIE (% points)	BPAC (% points)
Model 1 Poverty rate = 32.36%							
Single-step methods -MAXR variable selection							
OLS	0.587	79.06	63.53	36.47	28.24	-2.66	55.29
Quantile Regression (estimation point: 46)		79.44	68.24	31.76	31.76	0	68.24
Linear Probability	0.337	79.95	61.18	38.82	23.14	-5.08	45.50
Probit		80.71	63.92	36.08	23.53	-4.06	51.37
Two-step methods — MAXR variable selection							
OLS 51st Percentile	0.333 subsample	79.44	67.84	32.16	31.37	-0.25	67.06
Quantile Regression (estimation points 46, 24) - 51% cutoff		79.19	68.24	31.76	32.54	0.25	67.45
Linear Probability 54th Percentile	0.257 subsample	83.12	75.29	24.70	27.45	0.88	72.54
Probit 54th Percentile		81.59	69.80	30.20	26.67	-1.14	66.27
Model 4 Poverty rate = 32.36%							
Single-step methods -MAXR variable selection							
OLS	0.564	78.93	61.18	38.82	26.28	-4.07	48.63
Quantile Regression (estimation point: 45)		79.82	69.41	30.59	31.76	0.38	68.24
Linear Probability	0.321	79.44	62.75	37.25	26.28	-3.55	51.77
Probit		79.57	63.92	36.08	27.06	-2.92	54.90
Two-step methods — MAXR variable selection							

	Adj. R2	% Total Accuracy	% Poverty Accuracy	% Under-coverage	% Leakage	PIE (% points)	BPAC (% points)
OLS 55th Percentile	0.302 subsample	80.08	67.45	32.55	29.02	-1.14	63.92
Quantile Regression (estimation points 45, 26) - 55% cutoff		80.46	70.59	29.41	30.98	0.51	69.02
Linear Probability 55th Percentile	0.214 subsample	81.73	71.76	28.24	28.24	0	71.76
Probit 55th Percentile		81.35	70.20	29.80	27.84	-0.63	68.24
Model 7 Poverty rate: 32.36%							
Single-step methods -MAXR variable selection							
OLS	0.556	77.15	58.82	41.18	29.41	-3.81	47.06
Quantile Regression (estimation point: 44)		78.05	67.06	32.94	34.90	0.63	65.10
Linear Probability	0.309	79.69	61.18	38.82	23.92	-4.83	46.27
Probit		79.69	61.57	38.43	24.31	-4.57	47.45
Two-step methods — MAXR variable selection							
OLS 57th Percentile	0.319 subsample	79.70	68.24	31.76	30.98	-0.25	67.45
Quantile Regression (estimation points 46, 27) - 57% cutoff		78.93	67.84	32.16	32.94	0.25	67.06
Linear Probability 42nd Percentile	0.152 subsample	81.97	73.33	26.66	29.01	0.76	70.98
Probit 42nd Percentile		80.45	70.58	29.41	30.98	0.51	69.01
Model 9 Poverty rate: 32.36%							
Single-step methods -MAXR variable selection							
OLS	0.535	74.36	54.51	45.49	33.72	-3.81	42.74

	Adj. R2	% Total Accuracy	% Poverty Accuracy	% Under-coverage	% Leakage	PIE (% points)	BPAC (% points)
Quantile Regression (estimation point: 44)		75.51	61.96	38.04	37.65	-0.12	61.57
Linear Probability	0.278	76.78	52.94	47.06	24.71	-7.23	30.59
Probit		77.66	56.86	43.14	25.88	-5.58	39.60
Two-step methods — MAXR variable selection							
OLS 54th Percentile	0.210 subsample	77.79	63.92	36.08	32.55	-1.14	60.39
Quantile Regression (estimation points 44, 24) - 54% cutoff		79.06	67.84	32.16	32.55	0.13	67.45
Linear Probability 49th Percentile	0.159 subsample	79.44	64.71	35.29	28.23	-2.28	57.64
Probit 49th Percentile		79.44	64.31	35.69	27.84	-2.53	56.47

ANNEX F. VARIABLES INCLUDED IN THE BEST 15 MODELS

Annex F, Table 1. Variables included in the single-step OLS models (BEST 15 sets), estimated with SAS MAXR procedure

Variables	M1	M2	M3	M4	M5	M6	M7	M8	M9
Cooking fuel is charcoal or paraffin	X	X		X	X	X	X	X	X
Annualized non food expenditures 2 (education, health, and others)	X								
Value of panga	X	X		X	X		X	X	X
Number of days in past 7 days any of 4 superior food eaten (max. 28)	X			X				X	
Do you have nursery/pre primary center in your community?	X	X		X	X	X	X	X	
Whether spouse owns any type of shoes	X						X		
Are you or members of your household denied service or do you/they have only limited opportunity to sanitation services?	X	X		X			X	X	
Annualized food expenditures, recall 1 week	X								
Maximum education of any household member is post secondary/superior level (completed)	X	X		X	X	X	X	X	
Maximum education of males is secondary/post primary/J1 level (completed)	X	X		X	X	X	X	X	
Roof with banana leaves, fiber, grass, bamboo or wood	X			X	X	X	X	X	X
Household never ate sweet potato because other food was scarce	X			X					
Total household members in farmers group	X			X	X	X			
Do you have market in your community?	X			X	X	X	X	X	
Household head is widow	X	X		X	X	X	X	X	X
Amount household needs per month to live		X		X			X	X	
Average daily per-capita clothing expenditures		X						X	X
Expenditures on other expenditures (social events, leisure) last 12mo		X							
House size: large		X		X	X	X	X	X	
Do you have immunization center in your community?		X		X	X	X	X	X	
Age of youngest household member		X					X	X	
In the last 7 days, how many days did a main meal consist of staple food, plant protein and any vegetables only?		X							
Maximum education of any household member is no schooling/not completed grade one		X							
Median education of household members is primary level (completed)					X	X			
Household cooks in one of the rooms in the house					X	X			
Lighting source: Gas lamp or electricity (neighbor, public or own socket)					X	X	X		X

Variables	M1	M2	M3	M4	M5	M6	M7	M8	M9
Panga ownership						X			
Poultry ownership						X			
Cooking fuel is leaves, husk, or cow dung									X
Toilet: shared or own ventilated, improved latrine or flush toilet									X
Value of agricultural land									X
Value of CD players									X
Number of poultry owned by the household									X
Household head completed only secondary/post primary education									X
Number of household members with completed superior education, excluding head									X
Number of household members with no schooling or incomplete grade one, excluding head									X
Remittances sent to relatives, last 12 months									X

Annex F, Table 2. Variables included in the two-step OLS regressions (Models 1, 4, 7, and 9)

Variables	Model 1		Model 4		Model 7		Model 9	
Two-step OLS	1st step	2nd step	1st step	2nd step	1st step	2nd step	1st step	2nd step
Average age of household members, except head	X							
Cooking fuel is charcoal or paraffin	X		X		X		X	
Roof with banana leaves, fiber, grass, bamboo or wood	X	X	X		X	X	X	X
Household ate never sweet potato because other food was scarce	X		X					
Total household members in farmers group	X		X					
Have you or members of household are denied service or only limited opportunity to sanitation services	X		X	X	X	X		
Annualized food expenditures, recall 1 week	X	X						
Annualized non food expenditures for education, health, other	X							
Value of panga	X	X	X	X	X	X	X	X
Number of days in past 7 days any of 4 superior food eaten (max. 28)	X		X					
Do you have market/ bazaar?	X	X	X	X	X	X		
Do you have nursery/pre primary center?	X	X	X	X	X	X		
Household head is widow	X	X	X	X	X	X	X	X
Maximum education of household member is post secondary/superior level (comp)	X		X		X	X		
Maximum education of females is secondary/post primary/j1 level (comp)	X		X		X			
Household ate sometimes beer type bananas because		X						

Variables	Model 1		Model 4		Model 7		Model 9	
Two-step OLS	1st step	2nd step	1st step	2nd step	1st step	2nd step	1st step	2nd step
other food was scarce								
Toilet: shared or own ventilated, improved latrine or flush toilet		X					X	X
Whether Spouse owns any type of shoes		X		X	X	X		X
Amount household needs to live per months		X	X	X	X	X		
How far away the bodaboda (motorcycle) station (km)?		X						
How far away the private health center with midwife/nurse (km)?		X						
How far away the access to network phone (km)?		X						
Poultry ownership		X				X		
Maximum education of any household member is no schooling/not complete grade1		X				X		
House size: large			X		X	X		
Do you have immunization center?			X	X	X	X		
Do you have mobile (cell phone) in the house?				X				
Household ate mukene because other food was scarce				X				
In last 12 months did you or any other adult in your household stop eating meals for an entire day because you did not have enough money to buy?				X				
Poultry number				X			X	X
Rooms per person				X				
Median education of household members is primary level (comp)				X		X		
Age of youngest household member					X			
Lighting source: Gas lamp or electricity (neighbor, public or own socket)					X		X	
Household head: incomplete secondary education						X		
Cooking fuel is leaves, husk, or cow dung							X	X
Average daily per-capita clothing expenditures							X	
Value of agricultural land							X	X
Value of CD players							X	
Household head completed only secondary/post primary education							X	
Number of household members with completed superior education, except head							X	
Number of household members with no schooling or incomplete grade1,except head							X	
Remittances sent to relatives, last 12 months							X	X
Head has a daily agricultural labor								X
Water source: Open spring/well								X
Value of rubber shoes								X
Black/white TVs number								X
Number of household members who completed								X

Variables	Model 1		Model 4		Model 7		Model 9	
Two-step OLS	**1st step**	**2nd step**	**1st step**	**2nd step**	**1st step**	**2nd step**	**1st step**	**2nd step**
secondary/post primary education only, except head								
Number of household members with incomplete secondary education, except head								X

Annex F, Table 3. Poverty indicators used in the best model (in terms of maximization of BPAC)

Variable	Model 1	Model 4	Model 7	Model 9
Type of model	2-step LPM	2-step LPM	2-step LPM	2-step Quan
Household expenditures				
Annualized food expenditures, recall 1 week	X			
Annualized non food expenditures on education, health and others	X			
Cost of home improvements		X		
Clothing expenditures:				
Per-capita daily average				X
Education				
Household Head				
Household head completed only secondary/post primary education				X
Household Members				
Number of household members with incomplete secondary education, excluding household head	X	X	X	X
Maximum education level of any household member is post secondary/superior level (completed)	X		X	
Median education level of household members is completed primary level	X	X	X	
Maximum education of any household member is no schooling/not completed grade one			X	
Number of household members who completed secondary/post primary education only, excluding household head				X
Number of household members with no schooling or incomplete grade one, excluding household head				X
Number of household members with completed superior education (post secondary, specialized training, diploma or degree and above), excluding household head				X
Housing Characteristics				
Cooking fuel is leaves, husk or cow dung	X	X	X	X
Roof material is banana leaves, fiber, grass, bamboo or wood	X	X	X	X
Toilet: shared or own ventilated, improved latrine or flush toilet			X	X

Variable	Model 1	Model 4	Model 7	Model 9
Lighting source: Gas lamp or electricity (neighbor, public or own socket)			X	X
Water source is private borehole or piped water			X	
Cooking fuel is charcoal or paraffin				X
Water source: Open spring/well				X
Assets *				
Consumer durables				
Black and white TV ownership	X			
Whether spouse owns any type of shoes	X		X	X
Value of homestead	X	X	X	
Value of inherited funds	X			
Number of tire shoes owned by the spouse	X			
Value of leather shoes (spouse)		X		
Number of radios owned by the household			X	
Value of CD players				X
Value of rubber shoes				X
Number of Black/white TV's owned by the household				X
Agriculture				
Value of panga	X		X	X
Number of poultry owned by the household				X
Household head has agricultural daily labor				X
Value of not irrigated agricultural land				X
Financial				
Amount borrowed from formal institutions	X			
Household head has any type of account (savings, checking, fixed term deposit)		X		
Remittances sent to relatives in last 12 months				X
Other				
Household head is widow	X	X	X	X
Total number of days sick by females	X	X		
Number of marriages from household members in the past 3 years	X	X	X	
Total number of household members in farmers group	X	X	X	
Household head is unemployed			X	
Contribution to other groups			X	
Subjective Variables				

Variable	Model 1	Model 4	Model 7	Model 9
Have you or members of the household been denied service or only limited opportunity to security/police services?	X			
Amount that household needs per month to live	X	X	X	
Household mostly ate sorghum because other food was scarce	X	X		
In the last 7 days, how many days was large fish served to the household members?	X	X		
In last 12 months did you or any other adult in your household stop eating meals for an entire day because you did not have enough money to buy food?	X	X		
Have you or members of the household been denied service or only limited opportunity to sanitation services?			X	
Community				
Do you have immunization center in your community?	X	X	X	
Do you have market/ bazaar in your community?	X	X	X	
Do you have nursery/pre primary center in your community?	X	X	X	
Do you have access to electricity in your community?		X		

* For analysis, all monetary variables were used in logarithmic terms.

ENDNOTES

[1] This report consists of original work and data analysis. Citations of entire paragraphs or tables in published material by other authors is only permitted after prior consent with the authors and the IRIS Center. The cleaning and processing of data, as well as the entire analysis presented in this report, was carried out at the Institute of Rural Development, Georg-August-University of Göttingen, Germany. We gratefully acknowledge the valuable comments and support given by the IRIS project members Thierry van Bastelaer, Tresja Denysenko, Kate Druschel, and Anthony Leegwater; by Advisory Panel members Lauren Hendricks (CARE), Jonathan Murdoch (Princeton University), and Laura Foose (SEEP, PAWG); and by Stefan Schwarze, Isabelle Jaisli, Marinella Fader and Norbert Binternagel of the Institute of Rural Development at the University of Göttingen. The input by the SEEP Network and its Poverty Assessment Working Group (PAWG), the Advisory Panel for the Developing Poverty Assessment Tools project, and USAID is gratefully acknowledged. In particular, Christian Grootaert provided valuable comments and advice during all phases of the field research and data analysis, especially with regard to the choice of regression technique and the definition of accuracy measures. We gratefully acknowledge the excellent cooperation with Charity Irungu, the IRIS consultant for the field survey in Uganda, and the staff of the survey firm NIDA in Kampala, Uganda. All remaining errors are ours.

[2] Contact Information: Nkoola Institutional Development Associates (NIDA), 209 Upper Mawanda Rd., Old Mulango Hill, P. O. Box 22130, Kampala, Uganda. Tel./ Fax: 256 (0)41-530696, Email: consultus@nida.or.ug .

[3] Charity Irungu, Ph.D., is a lecturer at the Center for Disaster Management and Humanitarian Assistance (CDMHA) of Western University College of Science and Technology, P.O. Box 190, 50100, Kakamega, Kenya, http://www.wucst.org/.

[4] The Uganda Bureau of Statistics carried out the nationally representative household survey of 2002 in all districts in the Northern area. However, in some areas this was only feasible because of the protection of the survey teams by government security forces.

[5] Benchmark surveys were undertaken during the period August 4, 2004 thru October 25, 2004. About one third of survey households were interviewed in each of the months. The median survey month is September.

[6] Purchasing power parity exchange rates between US-Dollar and other currencies are available from the World Bank at www.worldbank.org/povmonitor/ppp1993.htm. According to data of the International Monetary Fund (Source: International Financial Statistics), the Ugandan consumer price index (CPI) is 106.042 for January 2003, and 115.440 for September 2004. Hence, inflation during the period of January 2003 to September 2004 was 8.87 percent (115.440 divided by 106.042). By adjusting the January 2003 purchasing power parity rate by this inflation rate, we calculate a purchasing power parity rate per 1 US-Dollar of 446.49

Uganda Shillings (1.0887 times 410.15 Ug.Sh.) valid for September 2004. This yields a value of 482.22 Ug.Sh. per 1.08 US-dollar for September 2004.

[7] Source of data is: Alan Heston, Robert Summers, and Bettina Aten, Penn World Table Version 5.6 and 6.1, Center for International Comparisons at the University of Pennsylvania (CICUP), October 2002.

[8] According to the International Financial Statistics data of the International Monetary Fund, the Ugandan consumer price index for June 1992 is 65.95, and for September 2004 it is 115.44. Hence, inflation during the period of June 1992 to September 2004 was 75.04 percent (115.44 divided by 65.952). By adjusting the 1992 purchasing power parity rate of 351.76 Ug.Sh. by this inflation rate, we calculate a purchasing power parity rate per 1 US-Dollar of 615.72 Uganda Shillings (1.7504 times 351.76 Ug.Sh.) valid for September 2004. This yields a value of 664.98 Ug.Sh. per 1.08 US-dollar for September 2004.

[9] The more recent Version 6.1 of the Penn World Data (Hester et al., 2002) yields a lower purchasing power parity rate for September 2004. The 1993 PPP rate of Version 6.1. is 326.06 Ug.Sh. per 1 US-Dollar, yielding a PPP rate of 573.94 Ug.Sh. per 1 US-Dollar for September 2004 after adjusting for inflation using the CPI data published by the IMF. The CPI index is 65.582 for June 1993 and 115.44 for September 2004, yielding an inflation of 76.02 percent (i.e., 326.06 Ug.Sh. multiplied by 1.7602). The international poverty line of 1.08 US-Dollar would then be 619.85 Ug.Sh. instead of the 664.97 Ug.Sh. calculated from the PWT, Version 5.6. Hence, the newer PWT Version 6.1 gives an estimate of the purchasing power parity rate that is about seven percent lower than the one derived from the older Version 5.6. Giving the benefit of the doubt to those hovering between 619 and 664 Ug.Sh. per day per capita, and following also Chen and Ravallion who used the PWT, version 5.6 for their international comparisons, the higher purchasing power parity rate and thereby higher international poverty line is chosen for this analysis.

[10] Based on the IRIS data, the poverty rate according to the national poverty line is almost the same with the poverty rate according to the one-dollar international poverty line. This result is consistent with the Appleton et al. (1999, page 14) who found that the national poverty line of Uganda when expressed in purchasing power parity is close to the international one-dollar-a-day poverty line.

[11] Annex A.3 compares the expenditure distributions of the IRIS survey with those obtained in the 2002/03 UNHS survey as well as prior surveys. Except for the first decile, which is noticeably poorer in the IRIS survey, the mean expenditures per adult equivalent in the remaining nine deciles differ by less than 11 percent between the two data sets. We thank Dr. Sarah Ssewanjana from the Economic Policy Research Centre of Makerere University for the provision of data on the expenditure distribution of households from the 2002/03 UNHS survey.

[12] The sampling frame excludes six poorer-than-average districts in the Northern region where 12 percent of the national population live. According to the most recent UNHS survey from 2002/03, the poverty rate in the Northern region was 65%. Using this poverty rate for the 12% of national population living in these areas excluded from the sampling frame, and using the

estimated poverty rate from the IRIS data set (i.e. for 88% of the national population) leads to an estimated poverty rate of 40.8% for the nation as a whole for September 2004. This suggests that the recent trend of rising poverty rates from 33.8% for 1999/00 to 37.7% for 2002/03 has continued during the past two years until September 2004.

[13] In each of the four IRIS test countries, we test nine different sets of regressors which are termed Model 1 thru 9. Figure 2.5.1. (see Section 2.5) describes the differences between the sets of regressors for these nine models. In the case of Uganda, we end up having only eight models. Model 3 was excluded because the variable total value of household assets was not selected by MAXR as one of the BEST15 variables.

[14] The best sets of poverty indicators identified for each of the nine models refer to the combination of BEST5, 10 or 15 indicators selected by the SAS-MAXR procedure.

[15] The terms regressor and poverty indicator are used interchangeably in this document. They refer to a certain type of variable used in the regression. The regressors can be derived from one or many questions from the composite questionnaire. For example, some regressors or poverty indicators are directly computed from the variable obtained in the survey, such as the age of the household head. Other regressors require computation (using information from one or several questions), as they are not asked directly but are derived from the responses to the questions asked. An example is the size of the household (which is calculated from the information given in section B of the questionnaire).

[16] For the case of zeros as original monetary values, these were replaced by the value of one pro mille of the mean in order to be able to compute the natural logarithm.

[17] Using the MAXR procedure of SAS, we selected in a prior model the two best regressors among 13 expenditure categories (referring to questions C1 to C12 as well as clothing expenditures of section B of the composite questionnaire). The inclusion of only the best two of the expenditure categories was done so as to avoid dominance of expenditure variables in subsequent models.

[18] It is therefore important to consider the framework of incentives for when, where, and by whom a poverty assessment is carried out (incentives for the respondent as well as the interviewer). The following quote taken from an email by Jan Maes (Trickle Up Program) highlights some of the issues involved here: “One way of preventing clients from exaggerating their poverty or otherwise responding in a way they think ’would help their case,’ is to conduct the poverty assessment survey after loan approval rather than to use it as part of the approval process. In other words, this implies that the USAID certified tools will be ex post poverty assessment tools rather than ex ante poverty targeting tools”… “If you use the assessment as part of the loan application or selection process, you will have to interview all potential clients, including of course those who ’fail the poverty test’. On the downside, since you only get your poverty results after clients have already entered the program, you might learn when it is already too late that you are not reaching the poorest.”

[19] The project director, supervisors and interviewers of NIDA were asked to rate the verifiability of each of the indicators on a scale from 1 to 5, where 1 is very difficult or impossible to

verify, and 5 stands for easy verifiability. In Annex D, we list the consensus rating given by the survey firm NIDA. In addition, NIDA rated the corresponding questions contained in the questionnaire according to their difficulty to ask. In model 7, we include only the regressors that have been rated as easily verifiable (i.e. a score of 4 or 5).

[20] These variables were identified by the SAS-MAXR procedure as the best five variables among all subjective variables. The set of regressors containing all subjective variables were excluded in Model 5.

[21] The PIE values from all models presented in this chapter are negative, implying an underestimation of the Actual Poverty Incidence. An overestimation of the Actual Poverty Incidence is expressed by positive PIE values . Values for PIE (regardless of a positive or negative sign) that are closer to zero indicate a higher accuracy in predicting the Actual Poverty Incidence.

[22] In order to replicate the LSMS regressor set, 7 education related variables which were not selected by the initial MAX R within the base regressor set of 250 indicators were incorporated in this model.