

**An Improved Data set for Demographic Research:
The KwaZulu-Natal Income Dynamics Survey (KIDS) 3rd Wave**

Julian May
School of Development Studies, University of KwaZulu-Natal

**Paper prepared for the “Forum on African Development and Poverty Reduction:
The “Macro-Micro Linkage”, Cape Town, 13-15 October, 2004¹**

Final

5 October 2004

¹/ An earlier version of this paper was made available at the Joint Population Conference, Durban, 6-8 October, 2004.

INTRODUCTION

Qualitative research such as the South African Participatory Poverty Appraisal has shown the extent to which change matters for those who are poor (May and Norton, 1997). Concern about future vulnerability and shocks, expectations that some event might dramatically transform their lives such as births, deaths and entry into the labour market, and anticipation of obtaining entitlements such as government grants are frequently described as being either features of poverty or as strategies that might offer pathways out of poverty. Each of these events is integrally caught up with the demographic and socio-economic life-cycle that individuals and families undergo as time passes (Chayanov, 1966).

In South Africa, the analysis of such change has relied upon cross-sectional studies or upon census data. Although useful, these data are unable to address a variety of questions, particularly those concerning dynamic processes and causal linkages. To address this gap, the KwaZulu-Natal Income Dynamics Study (KIDS) was undertaken by a consortium of South African and international researchers in 1998 which re-interviewed 1100 households first surveyed in 1993 as a part of the national Project for Statistics on Living Standards and Development (PSLDS). KIDS has just been extended by a further 5 years with a resurvey conducted in 2004. The study has also expanded to include adult children of the original sample who have established their own households and children who are being cared for by others. The methodology has also broadened to encompass both quantitative and qualitative methodologies. The quantitative component now provides a three period panel study that should yield a unique insight through the collection of survey data that spans South Africa's transition over a decade, the introduction of many policies intended to reduce poverty, as well as the era in which the impact of rapid HIV/AIDS infection is beginning to be felt. The in-depth qualitative research associated with the current study further focuses on some of these important issues by investigating changing household structure, caring patterns in households in which there are long-term ill members and livelihood activities.

This paper begins with a brief review of the usefulness of panel data before detailing the methodology and progress of this study. Procedures followed for KIDS 2004 will be described in detail and a very preliminary analysis of the attrition rates achieved in

1998 and 2004 is presented. The paper concludes with an assessment of the research program including information about public access of the KIDS 2004 data.

THE CASE FOR PANEL SURVEYS

Although opinions differ on the extent of change, few analysts would deny that South Africa has undergone a dramatic economic, political and social transformation during the first 10 years of democratic government. Recent debates over how best to reflect on these changes show the concern with finding ways of measuring the impact of both a new political and economic regime, as well as changing macro-trends that include globalisation and the HIV/AIDS epidemic. This is something that the cross-sectional surveys usually undertaken by official statistics agencies cannot readily do. While cross-sectional surveys can tell us whether poverty rates are decreasing, increasing or holding level, they cannot tell us about the fate of individual households. As an example, should cross-sectional surveys at two points in time reveal that the poverty rate is the same in each period, this could be the result of the same households having been in poverty in both periods. Alternatively it may be that some households managed to exit poverty over the period while an equal number entered. Such distinctions are likely to be very important in determining an effective policy response that would differ for chronic (the first case) versus transitory (the latter) poverty (Chaudhuri and Ravallion, 1994).

A different methodology is required if we are to understand what is happening to individual households over time. Typically referred to as longitudinal or panel surveys, the same households interviewed in first period are re-interviewed in the subsequent survey. With this sort of information one can determine whether the same or different households are in poverty in the two periods while also analysing the causes of these transitions.

Several well-known panel data sets exist that have been used to analyse a range of issues related to change. Examples include the determinants of income mobility using the Cote d'Ivoire Living Standards Survey (Grootaert and Kanbur, 1995), access to rural assets using the International Crops Research Institute Semi-Arid Tropics Village Level Studies in India (Gaiha and Deolalikar, 1993) and the influence of family history on children's well-being using the Panel Study of Income Dynamics in the United States of America (Brooks-Gunn et al, 1993). Baulch and Hoddinott,

(1999) provide a useful review of further examples while May and Roberts (2001) discuss the case for panel data in South Africa and the poverty transitions that could be identified in the first two waves of KIDS between 1993 and 1998. A particular advantage of such longitudinal surveys is that they allow us to analyse the dynamic behaviour of individual households, something not possible with standard cross-sectional surveys. A second advantage is that in many econometric analyses, longitudinal data enable us to control for unobserved, time-invariant characteristics of households that may bias efforts to estimate causal relationships using cross-sectional data. To illustrate this, rarely do we observe or measure in a survey a family's preferences and priorities for educating their children. It is quite likely that families that put a high priority on education will work extra hard to obtain income needed to pay school fees. However, if cross-sectional data is used in an effort to discover the impact that family income has on education, it is likely that biased results will be obtained because the families that are observed with the highest income may also be those who prioritise education the most. Estimates derived from such data will thus tend to overstate the impact that income transfers would have on educational decisions of families that give only an average priority to education. In contrast, with longitudinal data, panel data methods can be used to control for these non-time varying preferences and other family characteristics and thereby obtain unbiased estimates of the impact of income on education.²

Despite these advantages, it must be stressed that unlike cross-sectional surveys that are always representative samples for their particular point in time, longitudinal surveys, including the one described below, cease to be representative of the overall population after their first survey round. The representativeness of subsequent rounds is diminished further in rapidly changing countries and due to the sample attrition that inevitably occurs over time. Of equal concern are the problems associated with distinguishing real trends in the data from the noise produced by respondent, fieldworker and other sources of error. Care is necessary in both fieldwork and the subsequent analysis if these problems are not to result in misleading conclusions. At issue then is whether the 10 year KIDS project meets these criteria.

²/ This advantage must be balanced with an associated disadvantage that many of estimators that take advantage of this capacity of panel data to control for unobserved characteristics are quite sensitive to measurement error.

PROJECT FOR STATISTICS ON LIVING STANDARDS AND DEVELOPMENT

The PSLDS was the first nationally representative household survey in South Africa to investigate poverty, inequality and socio-economic dynamics. The study was undertaken in the last half of 1993 by a consortium of South African survey groups and universities under the leadership of the South African Labour and Development Research Unit (SALDRU) at the University of Cape Town (PSLSD 1994).³ Similar to the Living Standards Measurement Surveys (LSMS) that have been undertaken by the World Bank in more than 100 developing countries the main instrument was a comprehensive household survey collecting an array of information on the socio-economic circumstances of households. In addition to the household questionnaire, a community questionnaire was administered in each cluster of the sample to collect information common to households in an area such as school availability, health care facilities and prices for a selected group of commodities.

The 1993 sample was selected using a two-stage self-weighting proportional-to-population design using a sample frame based on the 1985 Census (PSLSD, 1994). An important component of the design was the definition of a household. A two-tiered definition for household members, resident or non-resident, was formulated based on time spent in residence. According to PSLSD (1994), resident household members, were defined as (i) those who had lived "...under this roof for more than 15 days of the last 30 days *and* (ii) when they are together they share food from a common source (i.e., they cook and eat together); *and* (iii) contribute to or share in, a common resource pool (i.e., they contribute to the household through wages and salaries or other cash and in-kind income or they may be benefiting from this income but not contributing to it, e.g., children, and other non-economically active people in the household)." The household was also defined to include non-resident members who were those that satisfied conditions (ii) and (iii) but who needed only to have lived "...under this 'roof' or within the same compound/homestead/stand at least 15 days out of the past year." Whilst a large set of information collected in the household roster was for resident members (age, gender, relationship to head, education and so forth), limited information was collected from non-resident household members.

³/ The PSLSD has also been referred to as the South African Integrated Household Survey (SAIHS), the South African Living Standards Measurement Survey (SA-LSMS), and the PSLSD/World Bank survey.

The PSLD has since been widely used for both policy analysis and the design of official statistics. Numerous research papers have been prepared using these data, while they have also formed the basis for international comparison (World Bank, 1998; UNDP, 1999). Although by now dated, it remains a benchmark in the history of South African poverty statistics.

KIDS 1998

With the aim of addressing research questions such as those concerning the dynamics of poverty already mentioned, households surveyed by the PSLSD in KwaZulu-Natal province were re-surveyed from March to June, 1998 by the KwaZulu-Natal Income Dynamics Survey (KIDS). The re-survey was directed by a research consortium including the University of Natal, the University of Wisconsin, and the International Food Policy Research Institute. Although the methodology has been detailed elsewhere (May *et al*, 2000), this section provides a summary of the major features of the study as these are pertinent to the most recent wave of data collection.⁴

The choice of KwaZulu-Natal for the resurvey was partly due to practical considerations including a confluence of research interests, resources, and the feasibility of relocating the households interviewed in 1993. Moreover, as KwaZulu-Natal and the Free State are the only regions in South Africa in which the provincial boundaries had not changed, the underlying sample frame of the study could be described as being representative of a meaningful policy planning structure.

In 1993, the KwaZulu-Natal portion of the PSLSD sample was representative at the province level, conditional on the accuracy of the 1985 census and other information used as the sampling frame. The data referred to 1558 households of all races located in 73 sampling points or clusters, 23 in the former 'white' province of Natal and 50 in the former Zulu homeland of KwaZulu. For KIDS, the 160 white and coloured households were excluded from the sample frame of KIDS due to the sampling biases that seemed likely given the small sample size and the distribution of the clusters that were sampled. While there were minor advantages to retaining these groups, such as

⁴/ Financial support for the 1998 wave of KIDS was provided by the United States Agency for International Development (Office of Women in Development - Grant Number FAO-0100-G-00-5050-00 Strengthening Development Policy through Gender Analysis; the BASIS/CRSP project at the University of Wisconsin-Madison, and a University Partnership Grant); the Ford Foundation; a Centre for Science Development population studies grant to the University of Natal; and the Development Bank of Southern Africa..

the maintenance of overall sample size and the value of sampling all ethnic groups in the province, in fact the sample size of these two sets of households was too small to permit comparative ethnic analyses. Moreover the households in these groups are entirely located in a small number of clusters (due to the general lack of spatial integration of the population) which appear to be non-representative at the ethnic group level.

During fieldwork, a further 216 households could not be located by KIDS, of which 164 were African. For more than one third of these households that were not re-interviewed, information collected verified the household had moved but was not detailed enough to allow tracking to a new residence. For the remaining households, however, there was simply no trace, i.e., no one approached in the community recognised the name of any household members when presented with the 1993 household roster. Maluccio (2000) shows that it is important to distinguish between these two groups, those who are known to have moved and those who seemingly left no trace. While the loss of the former group may be regarded as attrition, the prospect that the latter group may represent bogus interviews has to be acknowledged.

Furthermore, return visits to two clusters revealed that 36 households had been fabricated in both 1993 and 1998. An investigation was launched in which members of the 1998 field team were interviewed, and eventually 25 clusters were identified in which fabrication might have occurred. All 25 clusters were visited and the sampled households were located where possible. As a result of this scan, 138 mostly African households in five clusters were removed from the data set as these were suspected as having been fabricated either in 1993 or 1998. Careful checking of the data and of the field records suggests that this event was unique. Finally, 36 households had fragmented, splitting into two, and in 3 cases, 3 households. This meant that an additional 39 households were surveyed that contained core household members. Eventually, the matched 1993 and 1998 waves of KIDS contained data from 1031 households with an additional 39 split-off households.

To ensure comparability, the 1998 household questionnaire largely followed the 1993 version, although there were some important changes. One of these was a greater focus on individual (as opposed to household) ownership of assets and control over their use so that gender-differentiated analysis would be possible. A second change was the expanded emphasis on the set of individuals not living in the household but

economically linked to it. Four new sections were added which gathered data on economic shocks (both positive and negative), social capital (including group membership, kin networks, civic engagement, and trust), assets brought to marriage, and household decision making. The household questionnaire was necessarily complex and to ensure data collection accuracy, survey enumerators were trained for two weeks including practice interviewing on non-sample households. In the field, the questionnaire took close to three hours on average to complete, often requiring repeat visits in order to overcome respondent fatigue. To the extent possible, the new sections on economic shocks and social capital were replicated in 69 community level surveys.

Given the various purposes of the study, the identification of "main" decision-makers within households was very important to enable the collection of longitudinal data on them. In 1993, PSLSD recorded a head for each household. The head, a resident or a non-resident member of the household, was that person designated by the survey respondent to be the household head. The 1993 enumerator training manual offers no additional guidance or criteria for this designation, however. While in many instances it might be correct to assume that this "self-declared household head" corresponds to a main decision-maker, given the cultural diversity and complexity of households in South Africa, this may not always be accurate. For example, in three generation households, the survey respondent might declare the eldest the head but an employed middle aged child is really the primary decision-maker. To capture some of these complexities, the research team felt an expansion on the self-declared concept was necessary. This was done in an ex-ante fashion, through analysis of the 1993 data, and ex-post during the 1998 survey.

These individuals who were likely to be key decision-makers were termed 'Core' persons, an important feature of the 1998 wave of KIDS, and of course for all future waves since this concept determines who are what is to be followed. As May *et al.* (2000) document, a household member was designated ex-ante as a Core person if he/she satisfied any of the following criteria:

- 1) A self-declared head of household (from 1993)
- 2) Spouse/partner of self-declared head of household (from 1993)
- 3) Lived in a three generation household and all of the following were true:
 - Child, child-in-law, or niece/nephew of self-declared head
 - At least 30 years old
 - Have at least one child living in household
- 4) Spouse/partner of person satisfying criterion.

Thus, all heads of households and spouses of heads are automatically included and in some three-generation households, adult children are included. The rationale behind this approach aims to go beyond what might more conventionally called heads of household when deciding who or what should be followed. While the 1993 survey identified such an individual for each dwelling, analysis of this headship data revealed that the head was almost inevitably the oldest resident of the dwelling. While the decision-making power and social status of these individuals is doubtlessly real, the concern was that by focussing subsequent waves of data collection solely on this group would overlook other relevant household decision-makers. This concern is supported by number of studies of the problems associated with the "head of household" concept in South Africa (Ardington and Lund, 1994; Budlender, 1999) although Posel (2001) presents a convincing argument in favour of using self-declared headship.

Prior to beginning the 1998 fieldwork, using the above criteria a list of the Core persons in the household were identified to be targeted for additional information and tracking purposes. This designation was pre-printed on the household roster. It was hoped that this methodology would mean that fewer key decision makers in the household are missed than if only the self-declared heads were focused upon. In addition, guidelines were provided for designating new core persons to avoid missing other key decision makers in the household. The combination of Core persons, of which there were often more than one in an original 1993 household, and tracking movers meant that it was possible for original households to split and for the split-offs to remain in the sample. Furthermore, this methodology meant that when possible core household members who had moved were tracked, followed, and re-interviewed.

While the tracking procedures were somewhat more involved (see KIDS 1998 field worker training manual for details), the main elements were that Core persons were to be followed if they had moved and were no longer household members.

Of course, as a longitudinal survey, it is not proper to treat the 1998 observations in KIDS as a representative sample of households in 1998. In particular, KIDS is likely to necessarily under-represent “younger” households formed since the first survey. While the data may be weighted according to the 1991 census,⁵ this would not be meaningful for the 1998 data. Incomes and expenditures may be expressed in current prices, or using a community specific price index, constant prices can be calculated using fixed weight Lespeyeres type indices where (normalized) expenditure shares have been calculated using the 1993 data.⁶

In 2001, KIDS was supplemented by an in-depth study of a sub-sample comprising 50 households. These households purposively selected from 7 of the KIDS sampling points according to their poverty status⁷. Respondent households were identified which fell into four different dynamic poverty groups identified by Carter and May (2001): those that were poor in both periods (chronic poor); poor in 1993 but not in 1998 (got ahead); not poor in 1993 but poor in 1999 (fell behind) and not poor in both years (never poor). Taking advantage of the methodology, between four to six hours was spent with each household; usually over several days. The interview was organized around the development of a Household Events Map with household members through which key events in the household could be identified over the period 1990 to 2001. This was supported by a semi-structured interview. These data are still being analysed although papers have already been completed looking at the nature of work (Adato *et al.*, 2003) and social capital and poverty (Adato *et al.*, 2004).

KIDS 2004

KIDS 2004 is the recently completed third wave of this panel study. Once again the University of KwaZulu-Natal, the University of Wisconsin-Madison and the International Food Policy Research Institute (IFPRI) are the collaborating

⁵/ The “rcweight” variable provided in the original PSLSD data release is to be used to this purpose.

⁶/ The price index is based on a set of 12 basic commodities. For each commodity, prices were measured at the community level at both informal and formal shopping places (to the extent that both were available). A simple average of the formal and informal prices was used to create the index.

⁷/ See Adato, Lund and Mhlongo (forthcoming) for a more detailed discussion of the methodology of the qualitative study.

institutions.⁸ However, in order to accommodate new areas of interest, the participating institutions have been broadened to include the London School of Hygiene and Tropical Medicine (LSHTM), the Norwegian Institute of Urban and Regional Studies (NIBR). The study also has a far stronger policy focus and has been formally supported by the South African Department of Social Development (DSD) and input has also been received from the South African National Treasury. The study is being funded by the UK Department for International Development (DfID) through DSD, the National Research Foundation (NRF), the Norwegian Research Council (NRC), USAID and the Mellon Foundation.

As before, the 2004 study is based upon the original 1993 household socio-economic questionnaire and includes the collection of anthropometric data from children 6 years and below. New modules include the administration of a literacy test to children 7-9 years, a health module that includes information on caring. Several existing modules have been expanded or amended, including the collection of information on deaths in the household and the information collected on children. The qualitative component has been strengthened to include key informant interviews, focus group discussions and community and household case studies that combine ethnographic and structured interviewing techniques. Households that have split or fractured during this period have been followed up wherever feasible. An important change to the tracking process has been the inclusion of children of core household members who have established their own households, and of dependent children of the cores who are being cared for by other households. This was done retrospectively to 1993 both as a way of refreshing the KIDS panel as well as to increase the number of children for whom information is collected. These data are being supplemented by secondary data showing population characteristics, access to services and geo-referenced facilities and boundaries.

The household and community questionnaires from the 1998 survey were redesigned in a series of workshops held in Washington, London and Durban, and a pilot study was eventually conducted in the Umzimkulu district of Eastern Cape and the Cato Manor settlement in Durban during November, 2003. After a second round of pilot studies in January and early February, the final question design workshop in which all

⁸/ In January, 2004 the former Universities of Natal and Durban-Westville were merged to form the University of KwaZulu-Natal.

stakeholder organizations, including DSD were represented was run at UKZN over three days in mid February, 2004. The questionnaire was then translated into Zulu, and was back-translated into English to ensure consistency of interpretation. The training manual was updated to include new questions and procedures.

An improved community level questionnaire was developed which collected information from key respondents using a focus group methodology. Unlike the previous waves, the collection of these data was administered as an independent contract. To complement this information, a secondary data base has been compiled using the Geographic Information System (GIS) developed by the HSRC. Geo-reference information for every household has been collected using Global Position System (GPS) devices and these data will now be mapped into the secondary data base. The original maps and survey information from 1993 and 1998 have been restored from the SALDRA and UKZN archives and new maps have been prepared where necessary.

After a tender process overseen by DSD and the UKZN contracts office, a survey company was appointed in December, 2003. A scan of eligible households was undertaken prior to the fieldwork in February, 2004 which visited all households to be surveyed to gather information about household membership and migration. This scan found that 916 core persons could be located, 469 children of core persons had established their own households, and 242 foster children were living with other households. This gave an estimated sample of 1713 households to be interviewed, 113 more than originally estimated in the proposal.

Two weeks of intensive training was undertaken in March 2004 and fieldwork commenced immediately thereafter. The first week focused on training the field supervisors and quality control (QC) team and all training was undertaken by the University of KwaZulu-Natal (UKZN) researchers. The second week included all field-workers and training was undertaken jointly by the UKZN researchers and Development Research Africa (DRA) staff. Attention was paid to the anthropometric data collection, specialist trainers were used, and practical demonstrations were carried out. All fieldworkers and supervisors completed a competency exam on the entire questionnaire, and had to achieve a 70 percent pass rate before being accepted. Training on the use of anthropometric equipment was undertaken by the UKZN

Medical School and UKZN is leading the analysis of the CSG impact data working with the teams from the different universities.

Each wave of the study has required ethics clearance, and the design of Wave 3 has been subjected to especially rigorous evaluation by ethics committees at the three universities involved, addressing issues such as confidentiality, anonymity, the right of refusal and signed informed consent. Members of the research team provided ethics and sensitivity training to the enumerators and all members of the UKZN research team were required to complete the on-line course on Human Participants Protections Education for Research Teams course offered by the National Institutes of Health. Respondents were asked during Wave 2 whether they were willing to be revisited. Only those that agreed were approached for Wave 3. Participation in the third wave was voluntary and respondents were given the opportunity to withdraw from the study at any point, or refuse to answer specific questions. An incentive was given to the respondent in each household irrespective of their participation. This included household cleaning products and food.

The questionnaire and an informed consent form were translated into the language of the respondent, either English or Zulu. Approval was obtained from legal guardian(s) for the participation of children under 14 years of age before any data collection can take place. Oral consent was sought from children that are old enough to understand the request that is being made. Information packs were provided to the administrative authorities (municipal offices and/or traditional leaders) in areas in which the study was administrated. These packs contained summaries of the results from previous rounds of research using KIDS and also included development related material sourced from government and NGO's. A user-friendly, local language leaflet outlining the study and previous results was provided to all respondents. Enumerators carried identification, introduced themselves to respondents and provided letters of support from the University research team and the Department of Social Development. Permission to work in the survey areas was obtained in advance from the relevant local and provincial authorities.

Finally, records of the previous KIDS study have been placed in a controlled access environment and may not be removed from this room. All addresses and names have been removed from any public domain data and documentation, and a newly

introduced algorithm will prevent inadvertent matching of the data to administrative records that might result in the disclosure of identifying information.

Quality control has also been emphasised. During the fieldwork stage, check backs to households were undertaken by both the contracted survey company and independently by the research team. A ratio of 1 supervisor to 4 fieldworkers, a requirement of one interview being completed each day, the collection and capturing of detailed administrative information concerning the interview and quality control processes, and a comprehensive in-office quality control procedure were introduced. Completed questionnaires were checked in the field by the supervisors and in the office by both the survey company and the research team. Checks were also run on the first 400 interviews and again after 600 interviews had been completed. In particular, the ages of those interviewed in 1998 were not disclosed to the survey company, and as a result, this could be used to verify that each interview had been undertaken with the same household previously surveyed and that no fabrication had occurred.

Field work was completed in early August, 2004, and the cleaned data were submitted to the research team on 22 September, 2004. Preliminary inspection of the data suggests that in all sections, outliers appear to be minimal, and the data appear to be internally consistent. Virtually all ages have been collected for household members. More explicitly so than the previous waves, the goal of the 3rd wave of KIDS is to improve evidence based policy making for pro-poor policy through the analysis of these KIDS data and by placing these data into the public domain for analysis by the wider research community in South Africa and elsewhere. More specifically, the project will use these data for an analysis of long term poverty trends in South Africa, an impact assessment of social grants, particularly the Child Support Grant (CSG) that was introduced in 1998. The previous waves of the KIDS project also contain data on the health status of children, educational histories, birth records, household income and expenditure patterns, employment histories and finally community level data. As such, the study can be used to analyse the individual, household and community effects of the new CSG, as well as the effects of other interventions or trends that might have occurred. An analysis of the impact of multiple deaths preceded by a long illness is also planned. Although testing for HIV status has not been undertaken, in the younger adult age groups, it is anticipated these deaths will be the result of the

AIDS pandemic and simple cause of death data have been gathered to identify deaths from violence or accidents. During the first phase of analysis, five research questions will therefore be addressed by the research team:

- What are the pathways into and out of poverty including accumulation effects, life-cycle changes and social capital?
- What are the micro-economic barriers to mobility out of poverty in South Africa and how do these operate to produce poverty traps?
- What have been the long-term effects and impacts of government policies on poverty reduction, especially policies relating to social security grants, access to services and employment?
- What are the social and economic costs of the HIV/AIDS epidemic in terms of the impact on household livelihoods, the burden of care and coping strategies?
- What are the appropriate methodologies for the analysis of persistent poverty and what is their application to policy analysis and to the monitoring and evaluation of government policy?

Finally, KIDS 2004 has strong capacity building objectives. All members of the University of Natal research team have completed an ethics training course, two members have completed training in the facilitation of research and two will complete training in statistical analysis in January, 2004. Five South African PhD students are involved in the study. One student has successfully defended her proposal and completed all of the formal requirements for registration. A second student attended a 3 month 'pre-doctoral' training program at the University of Oslo and NIBR and from October, 2004 the third will be completing a 9 month course at the London School of Economics. The fourth student has already completed one year of her PhD at the University of Pennsylvania and has been awarded a US grant that covers her remaining tuition costs. Finally, a fifth student has started preparing her proposal for a thesis Masters in the School of Public Health at UKZN. Two NRF bursaries have been given to UKZN Masters students for 2004 and two of the Research Assistants have been offered financial support to complete a Masters degree in 2005. In addition, research assistants for the qualitative component of the study completed a 9 week in-depth training program in ethnographic methodology administered by UKZN's Anthropology department. UKZN counselling services have been used to assist team members affected in dealing with the trauma associated with prolonged contact with terminally ill respondents.

ATTRITION RATES IN THE KIDS STUDY⁹

An important question for any analysis using longitudinal data is the extent and nature of sample attrition. In theory, three factors underlie the level of attrition in a survey: the mobility of the target population, the success with which those who move are followed and interviewed, and the number of refusals. In practice, there is also the possibility of problems or errors in the fieldwork (both in earlier rounds and in the current one). A number of protocols were put in place to minimise attrition in the 1998 and 2004 re-surveys.

The 1993 (and thus 1998 target) sample included 1389 households¹⁰, (215 Indian and 1174 African). Interview teams first went to the original location of a 1993 household. If it was learned that the household had moved, the team was instructed to get new location information. They sought an address from other family, neighbours, schools, employers, etc. If a new address was found, and was sufficiently detailed, a team (later) followed the household. Sixty-three households were followed successfully using this protocol.¹¹ Of the target sample, 1132 households (81 percent) were successfully re-interviewed. For comparison, the LSMS Cote d'Ivoire panel survey in the late 1980s had re-survey rates under 90 percent after only one year (Grootaert and Kanbur, 1995) while the Peruvian (Lima) LSMS retained less than 60 percent of the original sample after five years (Glewwe and Hall, 1998). The second wave of the Indonesian Family Life Survey was more successful, re-interviewing over 93 percent of the sample after four years (Thomas, Frankenberg, and Smith 1999). Given the span of time and the mobility of the South African population, a re-survey rate of 81 percent seems acceptable.

⁹/ The data from KIDS 2004 are still preliminary and are currently being evaluated. All data in this and the subsequent sections are thus still tentative and should not be used for citation. A revised and final version of this paper will be released in April, 2005 once the evaluation period has been completed.

¹⁰/ This discussion excludes four households all of whose members died prior to the 1998 re-survey.

¹¹/ A small number of households moved out of KwaZulu-Natal; a few of these were followed but only one successfully interviewed.

Table 1: Attrition Rates in the 1993-1998 Waves (% of column)

Status	African (Non-Urban)	African (Urban)	Indian (All)	Total
Located and interviewed	703 79.5%	261 88.8%	168 78.1%	1132 81.3%
Moved, and could not be located	54 6.1%	13 4.4%	14 6.5%	81 5.8%
Not known in the area, refusal, fabricated or death	127 14.4%	20 6.8%	33 15.4%	180 12.9%
Total	884	294	215	1393

In most surveys of this type in developing countries, refusal rates are low. This is true in the KIDS survey: only 11 re-contacted households refused an interview. Many surveys in developing countries do not attempt to track movers. Had that strategy been followed, only 80 percent of the target households would have re-interviewed in 1998. Put another way, the tracking procedures yielded a 25 percent reduction in the level of attrition between the surveys.

Re-interview rates were higher in urban areas, where 90 percent of the target households were re-contacted. In large metropolitan areas, a sub-set of the urban sample which are characterised by more permanent housing structures and street addresses, re-interview rates were highest (not shown). There was less success when re-interviewing Indian households over 20 percent of which had moved between the survey rounds.

For more than one third of the households that were not re-interviewed, information collected verified the household had moved but was not detailed enough to allow tracking to a new residence. For the remaining households, however, there was simply no trace, i.e., no one approached in the community recognised the name of any household members when presented with the 1993 household roster. Maluccio *et al* (2000) show that it is important to distinguish between these two groups, those who are known to have moved and those who seemingly left no trace. While the loss of the former group may be regarded as attrition, the prospect that the latter group may represent bogus interviews has to be acknowledged. Furthermore, return visits to two clusters revealed that 36 households had been fabricated in both 1993 and 1998. Careful checking of the data and of the field records suggests that this event was unique. Overall, those not re-interviewed were significantly more likely to be Indian rather than African, have higher per capita income, higher per capita expenditure,

more educated household heads, and more durable assets. Obviously these variables tend to be correlated, and the comparisons suggest that the attrition present in KIDS between 1993 and 1998 was non-random (Maluccio *et al.*, 2000).

Although it is too soon to assess attrition in 2004, preliminary inspection of the data suggest that attrition rates are satisfactory with 87% of the households from the second wave successfully re-interviewed, equal to 64% of the original sample from 1993. Attrition rates for adult children of original participants who had established their own households are also satisfactory, with 71% of those correctly identified by the scan being interviewed. A number of areas were identified in which response rates could be improved with more effort and the field team has returned to these areas. Already more than 50 additional interviews have been completed. It is also worth noting that in 2004, the questionnaire was completed over two visits for more than three quarters of the households, with a mean 1.9 visits. The average time for the first visit was 2 hours. The average for the second visit was 1.2 hours, with 6 interviews being longer than 5 hours, and the average for the third visit was 48 minutes with 2 interviews longer than 5 hours. Average total contact time was therefore 2.8 hours, and 138 household had to endure more than 5 hours of interviewing and 5 households managed more than 10 hours of interviewing time.

CONCLUSION

The analysis of demographic, social and economic change imposes special data requirements. While useful, cross-sectional surveys of the type undertaken by most national statistical agencies do not provide the type of information that is required. In South Africa, the research based KwaZulu-Natal Income Dynamics Study (KIDS) does offer this opportunity. Although limited to one of South Africa's nine provinces, the study has gathered information from region characterised by relatively high levels of poverty, from households located in both urban and rural areas, as well as in former Homeland and white controlled districts. Between 1993 and 2004, attrition rates appear to be within acceptable limits, although small, and perhaps poorer households are more likely to have been missed in the second wave of the survey. Although never intended to become a panel survey, the effort put into the original PSLDS has enabled an important and unique data resource to be developed for socio-economic and demographic analysis. The revised three wave KIDS data will be placed into the public domain in the first half of 2005 once all cleaning processes have been

completed and a user-friendly meta-data document has been prepared. It is anticipated that the planning put into the most recent wave of data collection will make at least one further wave of KIDS an option for 2008.

REFERENCES

- Adato, M., F. Lund and P. Mhlongo. 2004. "Methodological Innovations in Research on the Dynamics of Poverty: A Longitudinal Study in KwaZulu-Natal, South Africa," paper presented at Q-Squared in Practice: A Conference on Experiences of Combining Qualitative and Quantitative Methods in Poverty Appraisal, University of Toronto, May 15-16.
- Adato, M., F. Lund and P. Mhlongo. Forthcoming. "Capturing 'Work' In South Africa: Evidence From A Study Of Poverty And Well-Being In KwaZulu-Natal," in Chen, M., Jhabvala, R. and Standing, G. eds. **Rethinking Work and Informality**, Geneva: ILO Publications.
- Ardington, E. and Lund, F. (1995), Pensions and Development: Social Security as Complementary to Programmes of Reconstruction and Development. **Development Southern Africa**, 12(4).
- Baulch, R. & Hoddinott, J., (2000). Economic Mobility and Poverty Dynamics in Developing Countries. **Journal of Development Studies**, forthcoming.
- Brooks-Gunn, J., Duncan, G.J., Klebanov, P.K. & Sealander, N., (1993). Do Neighborhoods Influence Child and Adolescent Development? **American Journal of Sociology**, 99(2): 353-95.
- Budlender, D., (1999b), The debate over household headship, Occasional Paper, Statistics South Africa, (<http://statssa.gov.za/debating/debating.htm>).
- Carter, M.R. and May, J., (2001), One Kind of Freedom: Poverty Dynamics in Post-Apartheid South Africa. **World Development**, 29(12): 198-2006.
- Chaudhuri, S. & Ravallion, M., (1994), How Well Do Static Indicators Identify the Chronically Poor? **Journal of Public Economics**, 53(3): 327-65.
- Chayanov, A.V., (1966), **On the Theory of Peasant Economy**, Richard D. Irwin for the American Economic Association, Homewood.
- Gaiha, R. & Deolalikar, A., (1993), Persistent, expected and innate poverty: estimates for semi-arid rural South India, 1975-1984. **Cambridge Journal of Economics**, 17(4): 409-21.
- Glewwe, P. and G. Hall (1998). Are some groups more vulnerable to macroeconomic shocks than others? Hypothesis tests based on panel data from Peru, **Journal of Development Economics**, 56(1):181-206.
- Grootaert, C. & Kanbur, R., (1995), The lucky few amidst economic decline: Distributional change in Cote d'Ivoire as seen through Panel Data Sets, 1985-88. **Journal of Development Studies**, 31(4): 603-19
- Maluccio, J.A., (2000), Attrition in the KwaZulu-Natal Income Dynamics Study 1993-1998. Food Consumption and Nutrition Division Discussion Paper. Washington D.C.: International Food Policy Research Institute.
- May, J. and A. Norton (1997), "A Difficult Life": The Perceptions and Experiences of Poverty in South Africa, **Social Indicators Research**, Vol. 41, Nos 1-3. pp.95-118 (23p).
- May, J., and B. Roberts (2001): Panel Data and Policy Analysis in South Africa: Taking a long view, **Social Dynamics**, Vol. 27. No. 1, 96-119.

- May, J., Carter, M.R., Haddad, L. and J. Maluccio, (2000), KwaZulu-Natal Income Dynamics Study (KIDS) 1993-1998: A longitudinal household data set for South African policy analysis, **Development Southern Africa**, Vol. 17(4), pp. 567-581.
- PSLSD. 1994. Project for Statistics on Living Standards and Development: South Africans Rich and Poor: Baseline Household Statistics. South African Labour and Development Research Unit, University of Cape Town, South Africa.
- Thomas, D., Frankenberg, E. & Smith, J., (1999), Lost But Not Forgotten: Attrition in the Indonesian Family Life Survey. RAND Labor and Population Program Working Paper No. 99-01. Santa Monica, California. RAND.
- UNDP, (1999): **Human Development Report for South Africa**, United Nations Development Programme, Pretoria.
- World Bank, (1998), **World Development Report 1998**. Oxford University Press, New York.