THE KWAZULU-NATAL INCOME DYNAMICS STUDY (KIDS) 3RD WAVE: METHODS, FIRST FINDINGS AND AN AGENDA FOR FUTURE RESEARCH

Julian D. May, Jorge Agüero, Michael R. Carter & Ian M. Timæus[1]

Submitted: 07 February, 2006
Revised: 02 April, 2007

Correspondence Address: Professor Julian May, School of Development Studies, Howard College Campus, University of KwaZulu-Natal, Durban, 4001. Tel: + 27 31 2602841; Fax: +27 31 2602359; Email: mayj@ukzn.ac.za

[1]/ Associate Professor, School of Development Studies, University of KwaZulu-Natal; Assistant Professor, Department of Economics, University of California, Riverside; Professor, Department of Agricultural and Applied Economics, University of Wisconsin-Madison; and Professor, Centre for Population Studies, London School of Hygiene & Tropical Medicine. In 2004, the KIDS team included Michelle Adato, Tania Boler, Einar Braathen, Thabani Buthelezi, Suriya Dawaad, Lawrence Haddad, Nina Hunter, Masinghita Khandlela, Mimi Ndokweni, Catherine van Ruit, Myra Taylor, Myriam Velia, Ingrid Woolard and Futhoshi Yamauchi. Particular acknowledgement is due to Thildé Stevens and Anna van den Berg of the Department of Social Development, Bridget Dillon of the Department for International Development and Neil Cohen of USAID for their support of this study. The efforts of the team at Development Research Africa who undertook the field work must also be commended. Earlier versions of this paper were presented at the Joint Population Conference, Durban, 6-8 October, 2004 and the Forum on African Development and Poverty Reduction: The "Macro-Micro Linkage", Cape Town, 13-15 October, 2004.

THE KWAZULU-NATAL INCOME DYNAMICS STUDY (KIDS) 3RD WAVE: METHODS, FIRST FINDINGS AND AN AGENDA FOR FUTURE RESEARCH

ABSTRACT

The panel study known as the KwaZulu-Natal Income Dynamics Study (KIDS) has been extended by a new wave of data collection conducted in 2004. The third wave of the study interviewed 865 households containing core adult members from 760 of the households contacted in 1993. It also conducted interviews in next-generation households that have split off from the parental households and in the current households of children who have been fostered out. The study finds that the proportion of people aged 20-44 dying between the second and third waves was nearly three times the proportion dying between the first two waves. The pattern of income distribution is one of increasing poverty and inequality since 1993, although the partial reversal of these trends in the post-1998 period is hopeful, as are signs of relative prosperity among those who established independent next-generation households. In addition, access to services has improved.

1 INTRODUCTION

May *et al.* (2000) argued the need for panel data for policy analysis in South Africa and described a panel data set that had become available. These data, known as the KwaZulu-Natal Income Dynamics Study (KIDS), were derived from households first surveyed in 1993 as a part of the national Project for Statistics on Living Standards and Development (PSLSD) and re-interviewed in 1998. KIDS has recently been extended by a further six years with a new wave of data collection conducted in 2004. The data provide a three period panel study that spans over a decade of South Africa's transition, the introduction of many policies intended to reduce poverty, and the era in which the impact of the rapid spread of the HIV epidemic has begun to be felt.

We review the earlier waves of the study before describing the 2004 wave providing a preliminary analysis of attrition in the panel. We go on to assess the demographic characteristics of the sample and review trends and transitions in economic well-being. Finally, we conclude by identifying important areas for further research that might explain these changes and provide information about public access to the KIDS 2004 data.

2 THE PSLSD AND KIDS 1998

The PSLSD was undertaken in the last half of 1993. It was the first nationally-representative household survey in South African to investigate poverty, inequality and socio-economic dynamics (PSLSD, 1994). The design of the study was based on that of the Living Standards Measurement Surveys (LSMS) undertaken with the sponsorship of the World Bank in more than 100 developing countries (Deaton 1997; Grosh and Munoz, 1996). The main instrument was a comprehensive household survey that collected an array of information on the socio-economic circumstances of households.

An important aspect of the design of PSLSD was the definition adopted of a household. A two-tiered definition of household members was formulated that distinguished between residents and non-residents. 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" (PSLSD, 1994). The household was also defined to include non-resident members who satisfied conditions (ii) and (iii) but needed only to have lived "…under this 'roof' or within the same compound/homestead/stand at least 15 days out of the past year."

The households visited by the PSLSD in KwaZulu-Natal province were re-surveyed from March to June, 1998 by KIDS. As the design of this study has been detailed elsewhere (May *et al*, 2000) and we will summarize only the features of the methods adopted pertinent to the most recent wave of data collection.

In KwaZulu-Natal, the 1993 PSLSD collected data on 1558 households of all races located in 73 sampling points or clusters. For KIDS in 1998, the 165 white and coloured households were excluded from the sampling frame as the PSLSD sample was stratified to ensure adequate representation of African and non-African households but not by the minority population groups individually. Efforts were made to trace households that had moved with 63 tracked to new locations. Four households had died out and further 218 households could not be located, some of which had probably died out too. For more than one third of them, information was obtained that verified that the household had moved but that was not detailed enough to allow tracking of it to a new residence. No trace was found of the remaining households: that is, no one approached in the community recognised the name of any household members when presented with the 1993 household roster. As Maluccio (2000) shows, while the loss of the former group may be regarded as attrition, the latter households have different characteristics and include a disproportionate number where the 1993 interview was poorly conducted or possibly fabricated.

Return visits to two clusters in 2001 revealed that the data on the 39 households in them probably were fabricated in both 1993 and 1998. An investigation was launched that interviewed members of the 1998 field team and 25 clusters in which fabrication might have occurred were visited and the panel households were located where possible. As a result of

this scan, 102 mostly African households in a further three clusters with high rates of attrition were temporarily removed from the data set as perhaps having been fabricated either in 1993 or 1998. In 2004, however, many households from the additional three clusters were successfully contacted and re-interviewed. Moreover, analysis has revealed that these clusters have characteristics or histories that might be expected to cause high rates of attrition. Since no evidence of fraud has been found in them and we are now certain that at least some of the interviews were genuine, these three clusters have been reincorporated in the dataset.

In 1993, respondents were asked to designate a single resident or a non-resident member of the household as the household's head. Analysis of these headship data reveals that the reported head is almost always the oldest resident of the dwelling. While the decision-making power and social status of these individuals is doubtlessly real, focussing subsequent waves of data collection solely on this group would overlook other important household decision-makers. To address such concerns, we decided to track and collect longitudinal data on a wider group of decision makers than reported household heads.

This was done in an *ex-ante* fashion through analysis of the 1993 data, and *ex-post* during the 1998 survey. Individuals who were likely to be key decision-makers were termed 'core' persons. The concept is an important feature of both the 1998 and 2004 waves of KIDS since it determines who or what is followed. As May *et al.* (2000) describe, a household member was designated ex-ante a core person if he/she satisfied any of the following criteria:

- A self-declared head of household (from 1993)
- Spouse/partner of self-declared head of household (from 1993)
- Lived in a three-generation household and <u>all</u> 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 the household

- Spouse/partner of person satisfying the previous criteria.

Thus, all heads of households and spouses of household heads are automatically designated core individuals, together with, in some three-generation households, adults in the next generation.

As an original 1993 household can include more than one core person, tracking them means that it is possible for original households to split and for both the households that result to remain in the sample. While the tracking procedures are somewhat involved, the crucial principle is that core persons have been followed if they have moved and are no longer household members.

3 KIDS 2004

3.1 Overview of methods

KIDS 2004 is the recently completed third wave of this panel study. Once again the collaborating institutions include the University of KwaZulu-Natal (UKZN), the University of Wisconsin-Madison and the International Food Policy Research Institute (IFPRI). In order to accommodate new areas of interest, the number of participating institutions was extended to include the London School of Hygiene & Tropical Medicine (LSHTM) and the Norwegian Institute of Urban and Regional Studies (NIBR). The study has also acquired a stronger policy focus and the 2004 wave was formally supported by the South African Government's Department of Social Development (DSD). In addition to the resources provided by each of the collaborating institutions, the study was funded by the Department for International Development (DFID), the National Research Foundation, the Norwegian Research Council,

USAID and the Mellon Foundation.

As before, the 2004 study is based upon the 1993 household socio-economic questionnaire and includes the collection of anthropometric data on children. New modules include the administration of a literacy test to children aged 7-9 years, a module on employment histories and one on the Child Support Grant (CSG). In addition, several existing modules have been expanded or amended, including those on deaths in the household, on health and caring, on social capital, and on children.

Using the same approach as in 1998, our strategy was to identify and survey the households of the core members of the original panel of households. Where core household members now live apart, all the households that they have joined or established were followed up wherever feasible. Moreover, it was decided to refresh the panel by designating the adult children of core household members who have established their own households and now have children of their own 'next-generation' cores and to survey their households as well. In addition, core members' children aged less than 18 years who are being cared for by other households were also tracked to increase the number of children on whom longitudinal information is available. In the survey documentation, these three groups of households are referred to as C (core), K (adult children of cores who have children and established their own households) and N (children aged less than 18 of cores who are being cared for by others). The questionnaires for the C and K households are similar, although the latter form includes questions about the parents of next-generation core members that had already been collected from the original core members in earlier waves. For the N group of households, the questionnaire covers only information on the household's composition and expenditure, the characteristics of the dwelling and the characteristics of the children themselves.

An improved community-level questionnaire was developed which collected information from key respondents by means of focus group discussions. To complement this information, geo-referenced data on the location of every household was collected using Global Position System (GPS) devices and a secondary data base compiled of existing geo-coded information, including access to services and facilities.

A scan of eligible households was undertaken prior to the main fieldwork. It involved visiting all households to be surveyed to gather information about household membership and migration, including households in the three clusters in which fabrication had been suspected. The 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, giving a provisional count of 1713 households to be interviewed. As already noted, the scan located a number of households from the three suspect clusters and interviews were conducted with them in 2004.

The fieldwork in 2004 was approved by the ethics committees of all three universities involved with regard to issues such as confidentiality, anonymity, the right of refusal and signed informed consent. Furthermore, respondents were asked in 1998 whether they were willing to be revisited and only those that agreed were approached in 2004. Respondents were given the opportunity to withdraw from the interview at any point or to refuse to answer specific questions. An incentive of household cleaning products and food was given to the respondent in each household irrespective of whether they participated, but always at the end of the interview. Respondents were also given a user-friendly, local language leaflet outlining the study and previous results.

Permission to work in the survey areas was obtained in advance from the relevant administrative authorities (municipal offices and/or traditional leaders). They were provided with information packs that contained summaries of the results from the previous waves of data collection and also included development-related material sourced from government and NGO's. The questionnaire and an informed consent form were translated into isiZulu and back-translated into English, to ensure consistency of interpretation, and administered in the language of the respondent, either English or isiZulu. Approval was obtained from their legal guardian(s) for the participation of children under 14 years of age before any data collection took place. Oral consent was sought from children that were old enough to understand the request that was being made.

The questionnaire was completed over two visits for more than three quarters of the households: the average time for the first visit was 2 hours and that for the second visit was 1.2 hours, with an average total contact time of 2.8 hours per household. Some 90 percent of the interviews were conducted between March and July, 2004, although data collection officially ended in January, 2005.

3.2 Attrition in the panel

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 panel study: the mobility of the target population, the success with which those who move are followed and interviewed, and the number of refusals. In practice, additional attrition may arise from other problems or errors in the fieldwork (both in earlier rounds and in the index one). A number of protocols were put in place to minimise attrition in the 1998 and 2004 re-surveys.

In 2004 tracking of the next generation was undertaken whether or not the parental core members were alive. Therefore, unlike in 1998, the household-level response rate in the third wave of KIDS incorporates 1993 "dynasties" where all the core members have disappeared or died but information was obtained on the households of children of core individuals who had established a new family (K group) or on core individuals' underage children who had been fostered out of the original household (N group).

Figure 1 shows the dynamics of the sampling of the families interviewed since 1993. The matched 1993 and 1998 waves of KIDS contained data on the 1171 households that had arisen by 1998 from 1132 of the 1354 eligible households interviewed in 1993. The third wave of the study interviewed 865 households containing core individuals. These core members originated in 760 of the households interviewed in 1993. The 865 core households represent 95 percent of such households initially identified as traceable in the scan. In 180 of the 760 dynasties that we traced, information was also collected on one or more next-generation households that had split off from the parental household. In addition, one or more households were surveyed containing children fostered out by 132 of the dynasties. In the case of these 760 dynasties for which we traced core households, interviewing next-generation households in 2004 (K and N groups) reduced the attrition of individuals but not of the dynasties themselves.

Although we failed to track any surviving core members, we obtained information on a further 81 of the 1132 dynasties contacted in 1998 by conducting an interview in one or more next-generation (K or N) households containing children of the core individuals. In almost 60 percent of these dynasties, the interviews established that all the core members of the original

household had died and provided information on these deaths. This sub-set of the original households have become extinct according to the study's definitions, rather than having been lost to follow up.

Figure 1 about here

In total, 2004 data exist for 74 percent of the dynasties contacted in 1998 and 62 percent of the eligible households interviewed in 1993. It is important to note that, although we only traced 841 dynasties, far more than 841 questionnaires were completed: in addition to the 105 additional core households generated by splits, the 2004 wave obtained data on 49 extinct core households, 319 next-generation households (68 percent of those identified by the scan) and 193 households containing children cared for by others (41 percent of those identified by the scan). The main reason for the high rate of attrition of these children is that they appear to be very mobile. The interviewers usually located the household identified in the scan as the child's home, but often found that the child either had never lived there or had already moved elsewhere.

The attrition rate of 26 percent of dynasties between the second and third waves of the study is higher than that between 1993 and 1998. Aging of the core members and the impact of AIDS related illness on adult mortality suggest that one reason why a lower proportion of dynasties were tracked was that more core households had ceased to exist because the last core member died than in between the first two waves. Also, increasing levels of internal and external migration in South Africa may be making households more difficult to track. For comparison, the LSMS Cote d'Ivoire panel survey in the late 1980s suffered more than 10 percent attrition in only one year (Grootaert and Kanbur, 1995) and the Peruvian (Lima)

LSMS lost track of more than 40 percent of the original sample after five years (Glewwe and Hall, 1998). On the other hand, the second wave of the Indonesian Family Life Survey successfully re-interviewed more than 93 percent of the sample after 4 years (Thomas, Frankenberg, and Smith 1999).

Table 1 compares the characteristics in 1998 of households that were interviewed in 2004 with those of households that were subject to attrition. Households that have disappeared from the panel were both smaller and less poor than those that were interviewed a third time. They also had lower dependency ratios on average. As May *et al.* (2000) and Maluccio (2000 discuss with regard to the 1993-98 panel, any response rate less than 100 percent implies some sample bias because households that do and do not remain in the sample may have different characteristics (that may be observed, unobserved or unobservable by the researcher). Equally, the implications of attrition for regression modelling of the data need not be great as such models typically stratify them according to many of the characteristics that influence attrition rates (Maluccio, 2000). Nevertheless, attrition of the sample, the limitations of the original 1993 sampling frame and the criteria used to define core members imply that the KIDS study is not exactly representative of Africans and Indians living in KwaZulu-Natal. Any panel study has this disadvantage, but we believe that it is overridden by the potential gains from the analysis of longitudinal data.

Table 1 about here

4 POPULATION STRUCTURE AND DYNAMICS

4.1 Demographic Structure

Although it can be rejuvenated by births or moves between households, the membership of a panel of households tends to age with time. The mean age of the resident members of the core KIDS households rose from 22.6 years in 1993 to 25.1 in 1998 and 26.0 in 2004. It was largely to ensure that the KIDS study continued to reflect the experience of recently formed households, especially those established by young people of childbearing age, that it was decided to track those children of core members of the 1993 households who had themselves had children. Figure 2 compares the age distribution of the resident members of core and next-generation (K) households in KIDS 2004 with that of the African and Indian population of KwaZulu-Natal as enumerated in the 2001 Census. Considered alone, the core households have more teenagers and people aged 50 or more among their members than the province as a whole. They contain many fewer young adults and children aged less than 10. However, the age distribution of the resident members of the core and next-generation households considered together matches that of the African and Indian population of KwaZulu-Natal more closely, although it continues to suffer from a deficit of members in their twenties and thirties and to have an excess of those aged 10-14. The sex ratio of those residing in KIDS households is 87 men per 100 women, which is the same as the ratio enumerated by the 2001 Census in KwaZulu-Natal. Moreover, the pattern of decline by age in the sex ratio of residents is almost identical in the two sets of data (not shown).

Figure 2 about here

Three different factors probably account for the under-representation of young adults in the panel. The first is that this age group contains a relatively high proportion of people living not in private households but in institutions such as hostels, army barracks, and prisons. The

institutional population is not represented in any household survey. Second, young adults without children frequently live alone or in small households with other young people. They may not spend much time at home and are difficult to contact and to recruit into a study. Domestic servants residing in single-person households on their employer's premises may be a particularly difficult group to contact. Therefore, even the population originally contacted in 1993 contained fewer young adults than one would expect based on the results of the 1996 Census. In addition, research priorities and cost considerations meant that the households established by the children of core members were tracked only if these children were reported to have children of their own. This has resulted inevitably in the under-representation of childless young adults who no longer reside in their parents' households.

4.2 Individual-level attrition

Attrition in the panel at the individual level is inevitably higher than the attrition of households because failure to trace households is compounded by the departure of individuals from the panel households who are neither core members themselves nor qualify for tracking as a adult child of a core member establishing a new household or an underage child of a core member. Table 2 presents attrition rates for individuals between 1998 and 2004. It reveals that, while South Africa is a society in which households usually endure, household membership is far more fluid. In addition to the 21 percent of individuals from 1998 who were members of households that could not be traced or were no longer eligible by 2004, a further 13 percent of individuals started off as members of households that were interviewed in 2004 but moved out of them during the 6 years between the waves. Thus, KIDS 2004 only collected information on 65 percent of 1998 household members. Moreover, many of these individuals were found in next-generation rather than core households. Of the 1998 members we have information on, 6 percent had died by 2004, 7 percent were non-resident members of

the household and 52 percent were resident. Slightly more of those household members who were resident in 1998 were traced and 55 percent of them were still residents in 2004.

The remaining rows of Table 2 shows the attrition rates by 2004 for different groups of 1998 household members. Demographic differentials in the proportion of individuals that have been followed longitudinally are modest. As one might anticipate from the design of the study, more core individuals than other household members were traced. Slightly more women were tracked than men. Unsurprisingly, young adults are more likely to have moved out of the panel households, and older people are less likely to have moved out, than other age groups.

Table 2 about here

4.3 Household sizes and dynamics

As one can calculate from Table 2, 17 percent of the members in 1998 of households that were traced in 2004 were no longer a member of the household. This outflow of people has been partly offset by other individuals joining these households. Overall, the mean size of households in the study rose from 6.8 in 1993 to 7.3 in 1988 before dropping to 6.9 in the 865 core households surveyed in 2004. The dynamics underlying these changes are complex. One factor pushing up household size is attrition from the sample. The average size in 1993 of those households that were not traced in 1998 was 5.2 and the average size in 1998 of those households that could not be traced in 2004 was 5.0, substantially less than the average sizes of the households surveyed successfully in the two later waves. Moreover, while approximately equal numbers of people moved into and out of the panel households during 1993-8, about 400 more births than deaths occurred in them, also leading to an increase in the

size of the households.[2] In contrast, between 1998 and 2004 only about 210 more births than deaths occurred in the households and about 870 more members left the core households than joined them. In addition, more households split between 1998 and 2004 than in the previous five years. Thus, the 1132 tracked households from 1993 became 1171 households in 1998 but the 773 tracked households from 1998 had split into 865 core households in 2004, a 12 percent increase. Overall, between 1993 and 1998, attrition and natural increase raised mean household size by 0.3 and 0.45 of a person respectively, while splits produced an offsetting reduction of 0.25 of a person. Household splits and the moves out of the households pushed the mean size of the households down by 0.9 and 0.6 of a person respectively between 1998 and 2004 but household size only fell by 0.4 because their impact was largely offset by higher attrition among small households.

Figure 3 about here

Although attrition in the panel was concentrated among small households, if we consider the core (C) and next-generation (K) households in KIDS 2004 together, the distribution of the sizes of the households has not changed greatly, although the proportion of households with 10 or more members did shrink between 1998 and 2004 (Figure 3). This suggests that many of the households that could not be traced were absorbed into larger households rather than continuing to exist as independent entities and that the dynamics of the panel households may be similar to those of households in the general population. While this is encouraging, Figure 3 also compares the sizes of the KIDS households with the equivalent distribution according to the 2001 Census, Although there are differences in the definitions used, this comparison

[2] As we do not know how many individuals who were born or died since the previous wave also moved between households, we cannot precisely disaggregate changes in membership into natural increase and changes in affiliation.

reveals that the KIDS panel has always contained too many large households and severely under-represented single person and two-person households. It is likely that the missing small households contain predominantly young adults.

Despite the substantial attrition occurring between successive waves of the KIDS study and the somewhat arbitrary decision rules used to revive the panel with newly-formed households, it continues in most of its important demographic characteristics to remain broadly representative of the population of the province.[3] Clearly, certain research questions should not be investigated using a panel of households that has been subject to substantial attrition between 1993 and 2004. The KIDS study under-represents the experience of small households and young adults – a group who are more mobile than older adults and a potentially important source of remittances but one that suffers from higher unemployment than more mature adults. For most purposes, however, the KIDS 2004 data seem unlikely to present a seriously biased picture of the welfare of the population of the province.

4.4 Mortality

Figure 4 portrays the proportion of members of the KIDS households dying between successive waves of the study by age in 1998 and sex. If mortality had been constant, we would expect about 35 percent more deaths in each age group between 1998 and 2004 than between 1993 and 1998 because the mean interval between interviews in the second and third waves was 6.15 years, compared to 4.56 years for the first and second waves. Except at ages

[3] Consider a simplified scenario in which all new households are established by a young couple who promptly have children: tracking all the children of cores who do this would lead to 100 percent over-representation of such households in KIDS. In practice, it is unlikely that the scan identified every next-generation household and only 68 percent of those it did identify were interviewed. Thus, the panel has remained broadly representative because of offsetting biases resulting from the design of the study and high attrition of next-generation (K) households.

20-44, this is more or less what occurred. At ages 20-44, however, the proportion of people dying between the second and third waves was nearly three times the proportion dying between the first two waves, suggesting that mortality in these age groups has more than doubled. Mortality in 1998 and 2004 reaches a local peak for men aged 30-34 in 1998 and women aged 25-29 in 1998 and then declines with age. Thus, 17 percent of male members of the households aged 30-39 in 1998 died by 2004 as did 10 percent of female household members aged 25-34. This unusual age pattern of mortality is typical of African populations in which HIV infection has become highly prevalent. Other recent mortality data for KwaZulu-Natal have found a similar hump in the age-specific mortality schedule affecting young adults and those that have collected data on causes of death have verified that it is attributable to deaths from AIDS (Hosegood *et al.*, 2004). The subsidiary peaks among women aged 50-54 in 1998 and 55-59 in 1998 can be accounted for by sampling error.

Figure 4 about here

In KIDS 2004, a module was added to the questionnaire asking about the circumstances surrounding deaths of panel members. This information was collected only for deaths occurring in 1998-2003 so as to avoid causing distress to the recently bereaved. The module includes questions about when the death occurred and about the costs associated with medical care during the dead person's terminal illness and with the funeral. While it did not attempt to collect detailed information on causes of death, the module does distinguish deaths from injuries from those from natural causes. This shows that only 29 percent of deaths of men at 20-44 in 1998-2003 resulted from injuries. Moreover, they accounted for only 6 percent of the equivalent group of deaths of women. As a high incidence of deaths of young men from accidents and violence is a long-established feature of South African mortality, it is clear that

injury deaths can not have made more than a small contribution to the higher mortality of the end of the 1990s and early part of this century.

Figure 5 about here

Simply comparing 1993-98 with 1998-2004 fails to convey the scale of the rise in mortality that had occurred in KwaZulu-Natal by 2004. Figure 5 shows the number of deaths occurring in the panel households each year from 1998 to 2004 by broad age group. They grew rapidly year-by-year, with more than 2.5 times the number of deaths occurring in 2004 as in 1998. Deaths of young adults aged 20-44 in 1988 underwent a fivefold increase during the six-year period between the second and third waves of the study.

5 INCOME DISTRIBUTION DYNAMICS AND CHANGES IN WELFARE, 1993-2004

Changes in income dynamics remain an important focus of KIDS and, using the new data, this section offers a first look at changing patterns of economic well-being over the first decade of South Africa's democracy. While interesting in their own right, these numbers are also provocative, demanding explanation for the patterns they portray.

5.1 Income distribution and poverty dynamics for core KIDS cohort

As described in section 2, the KIDS data come from repeated surveys of a 1993 cohort of core economic decision-makers. Using data only on the households of those core people who have been observed in all three time periods, we find that the headcount measure of poverty increased from 52 percent in 1993 to 57 percent in 1998, before falling to 47 percent in 2004. In calculating these measures, a household has been deemed poor if its per-capita

expenditures (inflated or deflated to 2000 prices) fell below the poverty line of R322 per month suggested for South Africa by Hoogeveen and Özler (2005).[4] While informative, consecutive snapshots of the poverty headcount reveal neither whether the same households have remained consistently poor, nor anything about changing patterns of income distribution among the non-poor.

Panel data like KIDS can help answer these questions. While the increasing availability of panel data has spawned new analytical methods and measures (see Carter and Barrett, 2006), transition matrices, which show how the fate of individual households evolves over time, continue to provide a compelling window into income distribution and poverty dynamics. Table 3 shows the transition over the full 1993 to 2004 period, whereas Table 4 shows the changes over the 1998 to 2004 sub-period. Both tables are based on a normalized real household expenditure measure, defined as total household expenditures, adjusted to 2000 prices, and divided by the Hoogeveen and Özler poverty line. Normalized expenditures equal to one thus indicate that household expenditures exactly equal the poverty line for the household; a measure of two indicates that household expenditures represent a level of material well-being that is twice the poverty line, and so on.

Table 3 about here

Each household in Table 3 is assigned to a row based on its 1993 normalized expenditure measure. Thus, the first row contains the 129 households whose 1993 level of well-being was less than half the poverty line. In the second row are the 218 households whose level of

[4]/ In order to derive this poverty line, Hoogeveen and Özler use a cost-of-calories approach in combination with the 2000 Income and Expenditure Survey undertaken by Statistics South Africa. Several options are suggested, and we have chosen to use their 'lower bound' estimate.

well-being was greater than half the poverty line but less than poverty line. The other rows are defined similarly using the well-being limits shown in the table.

The columns of Table 3 are defined using households' 2004 level of normalized well-being, and thus permit us to see the fate of each household over the 1993 to 2004 period. Looking across the first row, 38 percent of the households whose 1993 standard of living was less than half the poverty are just as poor in 2004. Another 34 percent of these households have modestly higher standards of living in 2004 (still below the poverty line, but above half of it). The remaining 28 percent of these households now enjoy standards of living in excess of the poverty line. In addition to the percentages of households in a specific transition category, each cell of the table reports the average 1993 and 2004 standardized expenditures for households in the cell. For example, the households that made the transition from less than half the poverty line to more than 2.5 times the poverty line had average expenditures equal to 40 percent of the poverty line in 1993, and equal to 3.6 times the poverty line in 2004. Finally, the main diagonal elements of the transition matrix are highlighted in bold and show the fraction of households in each row that have not changed their well-being category (for example, 73.4 percent of households that had living standards in excess of 2.5 times the poverty line in 1993 were still above that level in 2004.

Table 3 as a whole reveals several distinctive patterns of mobility with more than 60 percent of households that were poor in 1993 remaining poor in 2004. While there is some upward mobility amongst those who were initially poor, there is also substantial downward mobility (53 percent) amongst those just above the poverty line. These figures are consistent with the existence of a core group of persistently poor people, surrounded by a somewhat smaller group of sometimes poor who move in and out of poverty over time, an argument made

earlier by Carter and May (2000). The two expenditures groups just above the poverty line appear to be quite unstable. Roughly 40 to 45 percent of households that had expenditures between 1.0 and 1.5 times the poverty line in 1993 enjoyed expenditures more than 1.5 times the poverty line in 2004. Another 40 percent or so of these households had fallen below the poverty line in 2004, with the remaining 10 to 15 percent holding onto those middle positions. This pattern of apparent bifurcation (with some households slipping to a low level equilibrium and others rising toward a high level equilibrium) is consistent with that identified by Adato, Carter and May (2006) based on the 1993 to 1998 KIDS data supplemented with qualitative information from 2001.

Consistent with studies of the earlier rounds of the KIDS data, those households who were well above the poverty line in 1993 largely maintained their positions or moved ahead over time. On average, households with expenditures more than 2.5 times the poverty line in 1993 experienced a 61 percent income growth over the 11 years of the study. Nearly 40 percent of the households that had expenditures in 1993 between 1.5 and 2.5 times the poverty line moved ahead substantially over time and mean expenditure of this group grew by a massive 160 percent. Again consistent with the findings by Adato, Carter and May (2006), there is little downward mobility amongst these better-off groups.

Not surprisingly, the combined effect these mobility patterns is to increase income inequality, a finding consistent with those of Hoogeveen and Özler (2005) and many others analysing South Africa's income distribution since 1995. Among the KIDS households, the Gini coefficient measure of inequality in the distribution of household expenditures has risen steadily from 0.42 in 1993, to 0.50 in 1998 and a remarkably high 0.57 by 2004. As discussed by Carter and May (2001), this increase in income inequality is neither surprising nor an

unambiguously bad thing, although at some point high levels of inequality may impact upon economic growth rates and social stability.

While Table 3 presents an overall picture of the evolution of economic well-being for a cohort of KwaZulu-Natal households over the first decade of the post-apartheid economy, the KIDS data permit us a closer look at that history. As analyzed elsewhere (May *et al.*, 2000; Carter and May, 2001), the 1993 to 1998 period saw substantial increases in poverty and slippages at the bottom end of the income distribution, with substantially more improvement at the top end of the income distribution. The 1998 to 2004 period saw some moderation in this trend. Table 4 displays a transition matrix using the core KIDS households for the 1998 to 2004 period. This matrix is constructed identically to Table 3, although only the percentages of households in each cell are reported.

Table 4 about here

In addition to these basic percentages, the table also includes a simple coding scheme that indicates how the 1998 to 2004 transitions differ from the 1993 to 1998 transitions. As can be seen, much less downward mobility occurs among the poorest households in the later period than during 1993 to 1998 period. For example, while 24.7 percent of households that were just above the poverty line in 1998 had fallen below the poverty line in 2004, this fraction is more than 10 percentage points lower the corresponding transition figure over the 1993 to 1998 period (the figure from the 1993 to 1998 transition matrix, not shown here, is 38 percent). More generally, downward mobility and chronic poverty rates for the three lowest well-being categories (while still high) are not as unfavourable as they were for the earlier sub-period of the KIDS study.

The fourth well-being category (those households whose 1998 expenditures were between 125 percent and 150 percent of the poverty line) shows a mixed pattern, with some elements of downward mobility increasing and others decreasing. Finally, the pattern of upward mobility amongst the relatively well-off (already apparent in the 1993 to 1998 period) has become even more pronounced during the more recent period.

5.2 Household services and basic human needs

Improvements in the delivery of services have been identified by Statistics South Africa as an important achievement of the post-Apartheid government (Stats SA, 2004). As is shown in Table 5, positive changes have taken place in household's access to the set of basic needs indicators gathered by all three waves of KIDS.

Table 5 about here

The most notable progress is in electricity connections, which improves by from 43 percent of the sample to 75 percent between 1993 and 2004,. This is followed by the percentage of the sampled households who live in formal housing, which increases from 67 percent to 85 percent of the sample. Improvements in access to piped water has been more modest, while the percentage of households with access to a toilet in the dwelling or on the stand inexplicably declines between 1993 and 1998, perhaps due to definitional changes during fieldwork, before increasing to 86 percent of the sample. Finally, some progress has also been made in the percentage of households who report owning their house and in room density, measured as the median number of people per room. Some caution is needed in interpreting these roles since information was not gathered concerning disconnections, while

home ownership need not imply that the building can be sold or used as collateral arising from the form of tenure rights that are involved. Nonetheless, the picture suggested by the KIDS data concerning services is one of progress in a number of the key goals outlined by the Reconstruction and Development Programme in 1994.

5.3 KIDS: the next generation

Cumulative distribution functions (CDF) are a convenient way of comparing the income levels of different samples. Figure 6 provides the CDF of a sub-sample of the KIDS data. In this figure we depict the 2004 income distribution of the adult children of the original KIDS sample who have established their own homes with their own children (K households) whose parents were still alive in 2004. The figure compares this distribution to that of the parental homes from which these K households came for 1993, 1998 and 2004. The intention is to map the progress of the next generation as compared with that of their parents and, once again, we employ the normalized real household expenditure measure described earlier.[5]

Figure 6 about here

While the initial impression suggested by Figure 6 is that the independent adult children are doing noticeably better than their parents (their CDF lies below each of the CDF's for the parental homes at all per-capita expenditures), the story is more complex.

First, the sample of independent adult children may not be representative of all children of the

[5]/ The maximum of the scale has been cut at 7 times the poverty line in order to better show the data at the lower end of the income distribution.

KIDS households who are of a similar age since not all adult children have been able to graduate to having their own homes. It may well be the more successful, perhaps better educated and higher earning children who have been able to successfully complete the transition to adulthood by establishing their own households. Poorer, less-educated children may have remained in the parental household having been unable to afford the costs of establishing an independent household such as *lobola* (brideprice), the cost of obtaining and setting up a home or the cost of starting a family.

Supporting this idea, the situation of the sub-samples of adult children and their parental homes is quite different from that of all core households in the KIDS sample. The parental group is currently somewhat wealthier (43 percent are below the poverty line compared to 47 percent for the total sample), and they have been so historically. Despite this generally favourable history though, the poorer households in parental group have not shared in the progress made by the KIDS sample as a whole, and there has been no decrease in the percentage falling below the poverty line in 2004, compared with 1998, as was experienced by the core KIDS sample. Nonetheless, parental households above the poverty line have improved their position, again indicating an increase in inequality among this group.

6 AN AGENDA FOR FUTURE RESEARCH

The collection of data from the same households or individuals over time to form a panel study is becoming an increasingly popular option for research concerned with the analysis of trends and transitions. Although complex and subject to their own forms of bias, such data can better reveal the nature of population and socio-economic dynamics than more conventional cross-sectional data. In the context of South Africa, panel data provide an important window on the impact of a period of profound demographic and socio economic

change arising from political reform and more recently the impact of the HIV epidemic. While limited to only one of South Africa's provinces, the three waves of KIDS have gathered information from a 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.

Analysis of the KIDS data reveals some disturbing trends. The proportion of people at ages 20-44 dying between the second and third waves was nearly three times the proportion dying between the first two waves, suggesting that mortality in these age groups has more than doubled. This age pattern of mortality is typical of African populations in which HIV infection has become highly prevalent and may be indicative of trends unfolding elsewhere. Further, the pattern of income distribution among the KIDS cohort is one of increasing poverty and inequality between 1993 and 1998, a result that has also been found in national surveys of income and expenditure. That said, the partial reversal of these trends in the 1998 to 2004 period is somewhat hopeful as are signs of relative prosperity among those that successfully established independent next-generation households. In addition, access to electricity, reticulated water and housing has generally improved for households in the KIDS sample, another finding supported by national surveys.

Several alternative hypotheses suggest themselves as possible explanations for the socio-economic trends observed in the KIDS data and will require more detailed investigation:

1. The incident and impact of multiple deaths preceded by a long illness is likely to have become an important factor determining the well-being of households. Although testing for HIV status has not been undertaken by KIDS, simple cause-of-death data have been gathered to identify deaths from violence or accidents. This permits a deeper investigation

of the extent to which the adverse effects on the household of adult deaths result from increased poverty or reflect other processes (e.g. substitution of child for adult domestic labour, reduced time inputs to parenting). In addition, identifying the effects themselves will be important in such analysis as these might include the ill-health of other members, withdrawal from the labour marker, poor achievement at school or teenage pregnancy.

2. Government expenditure on transfer payments increased throughout the post-Apartheid period, while the introduction of the Child Support Grant has dramatically increased the number of recipients of such grants. This is partly reflected in the KIDS data which show that the amount of transfers per-household have doubled, while the incidence of grants other than the Old Age Pensions increases five-fold between 1998 and 2004. Measuring the impact of the government transfers on the well-being the KIDS sample is thus another fruitful area for analysis. Due to the panel nature of the KIDS data, quasi-experimental approaches may be possible that will better identify the link between receipt of the grant and the well-being of children and other household members.

3. Another area for analysis is on the operation of the labour market, and the expansion or contraction of employment and change in wages. Mobility and its causes, such as improved education and better functioning of the labour market represent a related theme that is worth exploring. In particular, the continued presence of poverty traps should be explored.

4. Finally, the impact of improved service provision warrants further investigation. This may have enabled other improvements in well-being as time is freed for other productive or reproductive work such as child care or the care of the ill.

Although it was never intended to become a panel survey, the effort put into the design of the original PSLSD has enabled an important and unique data resource to be developed for socio-economic and demographic analysis that remains relevant a decade after the data were first gathered. The updated three-wave KIDS data have been placed into the public domain and can be downloaded from http://sds.ukzn.ac.za/. 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 2008/2009, which should provide further insight into inter-generational dynamics as well as longer-run economic and demographic change. However the imminent introduction of a national income dynamics study may be a more useful direction for future panel data collection in South Africa.

REFERENCES

ADATO, M, CARTER, M.R & MAY, J, 2006. Exploring Poverty Traps and Social Exclusion in South Africa using Qualitative and Quantitative Data, *Journal of Development Studies*, 42(2): 226-247.

CARTER, M.R, & BARRETT, C, 2006. The Economics of Poverty Traps and Persistent Poverty: An Asset-Based Approach, *Journal of Development Studies.* 42(2): 178-199.

CARTER, MR & MAY, J, 2001. One kind of freedom: poverty dynamics in post-Apartheid South Africa. *World Development*, 29(12): 198-2006.

DEATON, A.., 1997. *The Analysis of Household Surveys*, The Johns Hopkins University Press, London.

GLEWWE, P & HALL, G 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.

GROSH, ME & MUNOZ J. 1996. *A Manual for Planning and Implementing the LSMS Survey*. Living Standards Measurement Study, Working Paper No. 126, Washington, D.C.: The World Bank.

HOOGEVEEN, JG & ÖZLER B, 2005.*Not Separate, Not Equal: Poverty and inequality in post-apartheid South Africa* William Davidson Institute Working Paper No. 739. Ann Arbor: University of Michigan.

HOSEGOOD, V, VANNESTE, A-M & TIMÆUS, IM, 2004. Levels and causes of adult mortality in rural South Africa. *AIDS,* 18(4): 663-671.

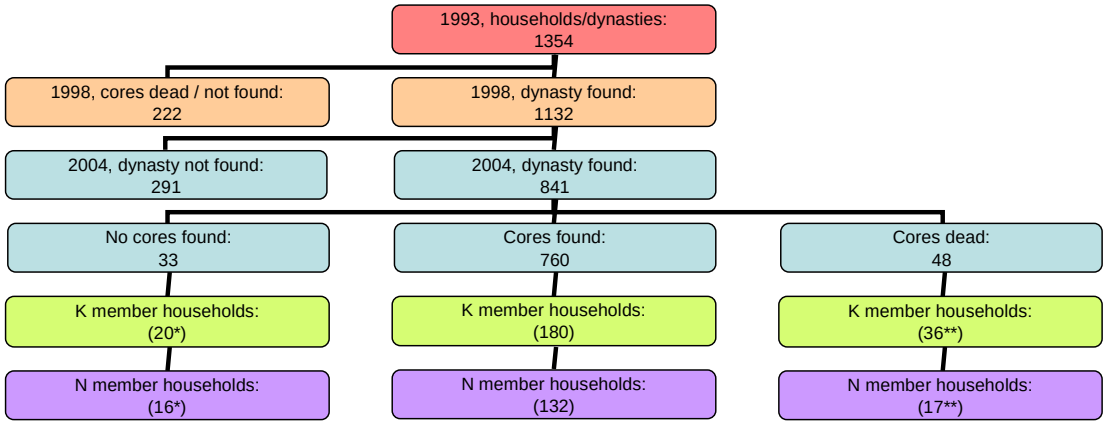
MALUCCIO, JA, 2000. *Attrition in the KwaZulu-Natal Income Dynamics Study 1993-1998*. Food Consumption and Nutrition Division Discussion Paper No. 95. Washington D.C.: International Food Policy Research Institute.

MAY, J, CARTER, MR, HADDAD, L & MALUCCIO, J, 2000. KwaZulu-Natal Income Dynamics Study (KIDS) 1993-1998: A longitudinal household data set for South African policy analysis, *Development Southern Africa***,** 17(4): 567-581.

PSLSD, 1994. *Project for Statistics on Living Standards and Development: South Africans rich and poor: Baseline household statistics.* Cape Town, South Africa: South African Labour and Development Research Unit, University of Cape Town.

THOMAS, D, FRANKENBERG, E & SMITH, J, 1999. *Lost but not Forgotten: Attrition in the Indonesian Family Life Survey*. RAND Labour and Population Program Working Paper No. 99-01. Santa Monica, California: RAND.

Figure 2. Dynamics of the sample from 1993 to 2004



* 3 dynasties where none of the original core members were found are represented by both K and N households in 2004
** 5 dynasties where all the original core members have died are represented by both K and N households in 2004

Table 1: Characteristics of 1998 households by whether interviewed in 2004

Characteristic (mean of households' values)	Core household(s) interviewed in 2004	Other household(s) interviewed in 2004	No households interviewed in 2004	All 1998 households
Number of household members	8.7	7.8	4.9	7.6
Number of resident members	7.0	6.4	4.3	6.2
Per capita expenditure (rand)	479	501	661	528
% below poverty line	56	56	44	53
% of residents of working age	50	52	62	53

Figure 2. Age distribution of resident members of core (C) and next-generation (K) households compared with the 2001 Census population (Africans and Indians in KwaZulu-Natal)

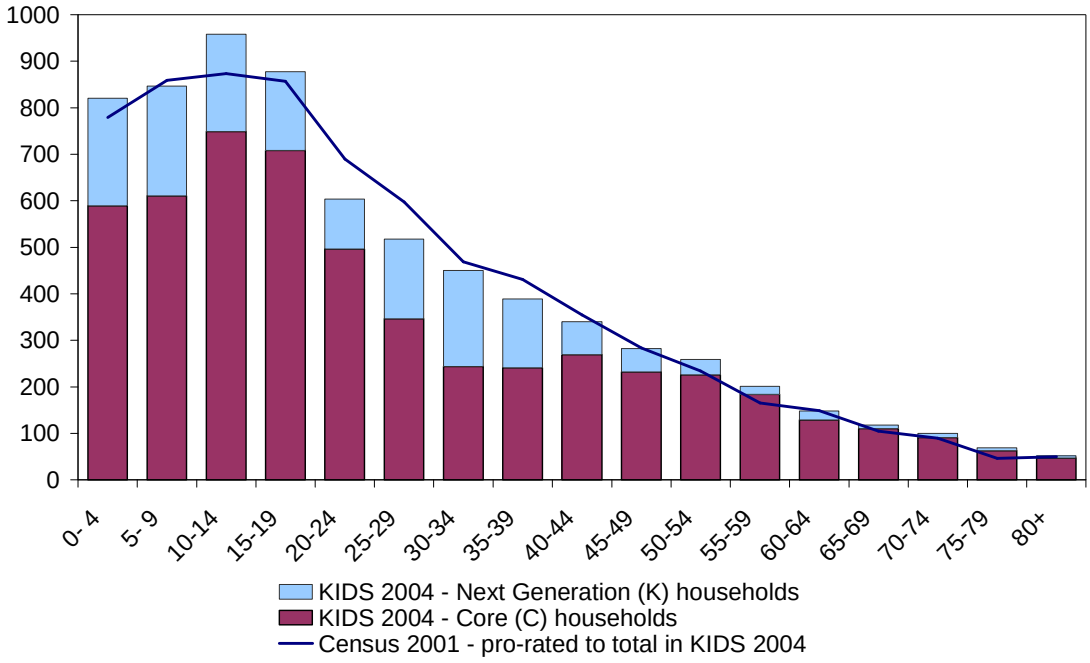
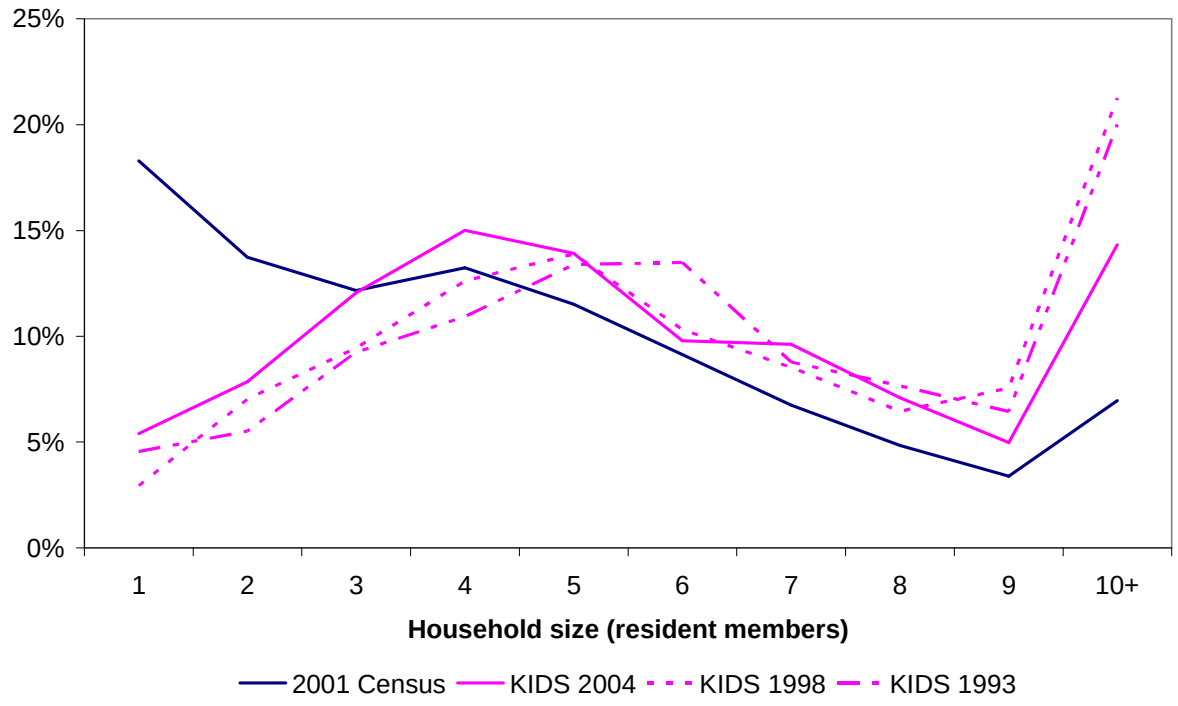


Table 2. Attrition by 2004 of household members in 1998 (%)

6.1		*6.2*	*Members in 1998*	*6.3*	*Resident in 1998*
6.4	*Household not traced*	*6.5*	*17.3*	*6.6*	*17.9*
6.7	*Household not eligible (no core members)*	*6.8*	*4.0*	*6.9*	*3.8*
6.10	*Individual no longer member in 2004*	*6.11*	*13.4*	*6.12*	*11.6*
6.13	*Data collected traced in 2004, comprising:*	*6.14*	*65.2*	*6.15*	*66.7*
6.16	*Died between waves*	*6.17*	*6.0*	*6.18*	*5.9*
6.19	*Non-resident member*	*6.20*	*7.0*	*6.21*	*5.7*
6.22	*Resident member*	*6.23*	*52.2*	*6.24*	*55.2*
6.25	*Attrition in different sub-groups:*	*6.26*		*6.27*	
6.28	*Core individuals*	*6.29*	*29.6*	*6.30*	*29.5*
6.31	*Other individuals*	*6.32*	*36.5*	*6.33*	*34.6*
6.34	*Men*	*6.35*	*36.2*	*6.36*	*34.5*
6.37	*Women*	*6.38*	*33.6*	*6.39*	*32.3*
6.40	*Aged 0-19*	*6.41*	*33.4*	*6.42*	*32.3*
6.43	*Aged 20-39*	*6.44*	*40.2*	*6.45*	*38.5*
6.46	*Aged 40-59*	*6.47*	*32.2*	*6.48*	*31.6*
6.49	*Aged 60+*	*6.50*	*24.6*	*6.51*	*23.5*
6.52	*Number of individuals*	*6.53*	*8547*	*6.54*	*7304*

Figure 3. Distribution of households according to size, core (C) and next-generation (K) households compared with the 2001 Census population (Africans and Indians in KwaZulu-Natal)



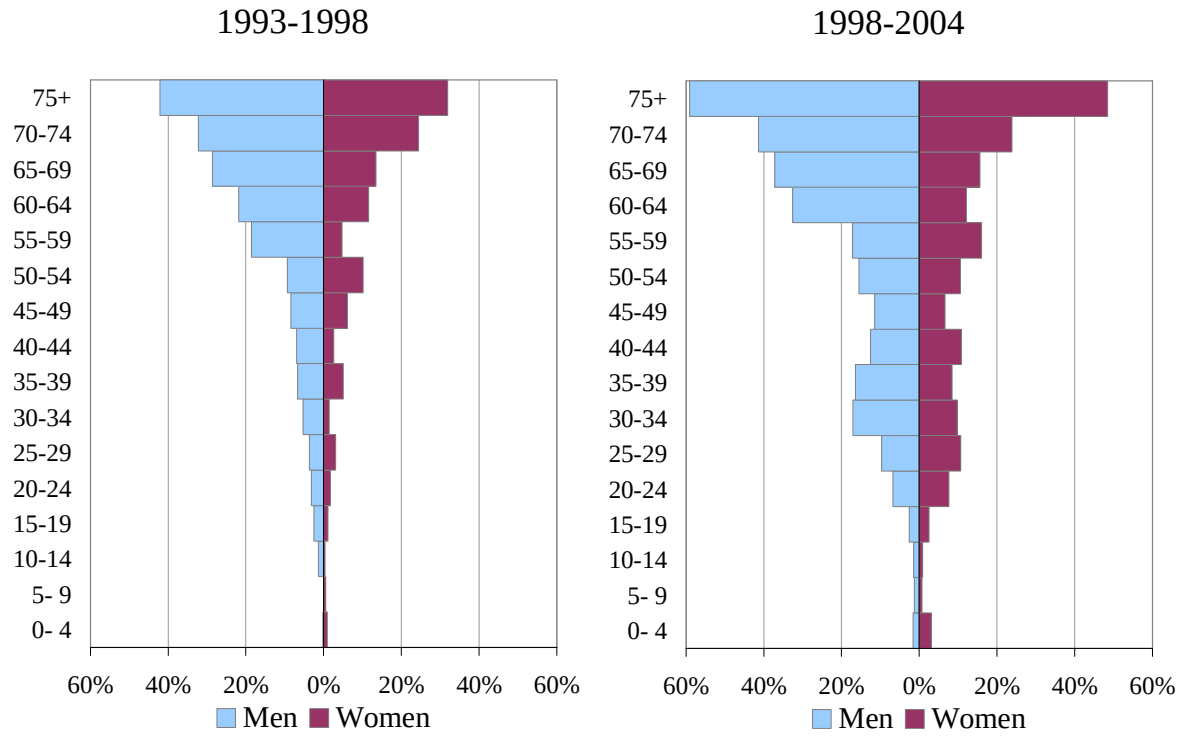


Figure 4. Percentage of traced household members dying between waves by age at first wave and sex (ABOVE)

Figure 5. Deaths in the panel households by year according to broad age group in 1998

(*Deaths in 1998 and 2004 have been multiplied up to correspond to a whole year of exposure using dates of interview)

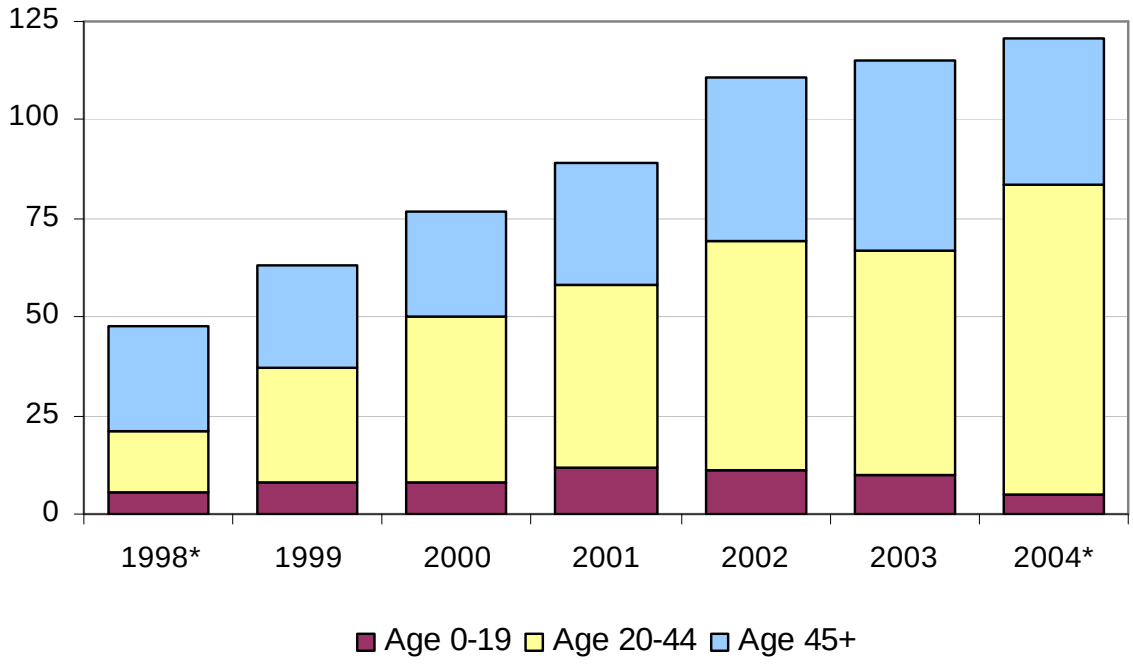


Table 3: 1993 to 2004 Transition Matrix

		< 0.5 PL	< 1PL	<1.25 PL	<1.5PL	<2.5PL	>2.5PL
<0.5PL (n=129)	% of Row	**38.0**	34.1	10.9	3.1	8.5	5.4
	1993 NE*	0.4	0.4	0.4	0.4	0.4	0.4
	2004 NE	0.3	0.7	1.1	1.4	1.8	3.6
< 1PL (n=218)	% of Row	26.8	**31.9**	11.4	10.4	11	8.5
	1993 NE	0.7	0.7	0.7	0.7	0.7	0.8
	2004 NE	0.3	0.7	1.1	1.4	1.9	4.2
<1.25 PL (n=111)	% of Row	26.1	27	**12.6**	1.8	16.2	16.2
	1993 NE	1.1	1.1	1.1	1.1	1.1	1.2
	2004 NE	0.3	0.7	1.1	1.3	2	5.1
< 1.5 PL (n=69)	% of Row	11.6	20.3	5.8	**7.2**	18.8	36.2
	1993 NE	1.4	1.3	1.3	1.3	1.4	1.4
	2004 NE	0.4	0.7	1.1	1.3	1.9	5.3
< 2.5 PL (n=161)	% of Row	10.6	15.5	8.1	4.3	**23.6**	37.9
	1993 NE	1.9	1.8	1.8	1.7	1.9	2
	2004 NE	0.4	0.8	1.1	1.4	2	5.2
>2.5 PL (n=79)	% of Row	1.3	10.1	3.8	3.8	7.6	**73.4**
	1993 NE	4.8	3.5	3	3	3.4	5.2
	2004 NE	0.3	0.7	1.1	1.4	2	8.4

- NE indicates real per-capita expenditures that have been normalized by the Hoogeveen and Özler (2005) poverty line.

Table 4: 1998 to 2004 Transition Matrix*

	< 0.5 PL	< 1PL	<1.25 PL	<1.5PL	<2.5PL	>2.5PL
<0.5 (n=221)	**40.7**--	37.6	4.5	3.2-	8.6+	5.4++
<1 (n=275)	25.5---	**32.1**-	11.7	8.4	13.5++	8.8++
<1.25 (n=82)	17.3	24.7---	**9.9**	11.1-	22.2+++	14.8+
<1.5 (n=68)	16.2++	16.2---	25+++	**8.8**	13.2	20.6+
<2.5 (n=109)	3.7-	12.8---	11.9	5.5-	**22.9**-	43.1+++
>2.5 (n=113)	0	5.3	3.5-	2.7	11.5--	**77+++**

- Changes from the 1993 to 1998 transition matrix are indicated by '+'s' and '-'s.' An element which is between 2.5 and 5 percentage points lower is indicated by a '-'; between 5 and 10 percentage points lower by a '--'; and, greater than 10 percentage point decrease by a '---'. Similarly, '+' means between 2.5 and 5 percentage higher; '++' between 5 and 10 percentage points higher; and '+++' more than 10 percentage points higher

Table 5: Basic needs indicators for core KIDS households

Year	Have electricity connection	Live in formal house	Piped water supply	Toilet on stand	Own house	Median people/ room
1993	42.5	67.3	37.2	76.6	87.0	1.4
1998	64.8	n/a	41.5	67.2	88.7	1.2
2004	74.5	85.2	50.7	86.3	90.4	1.0

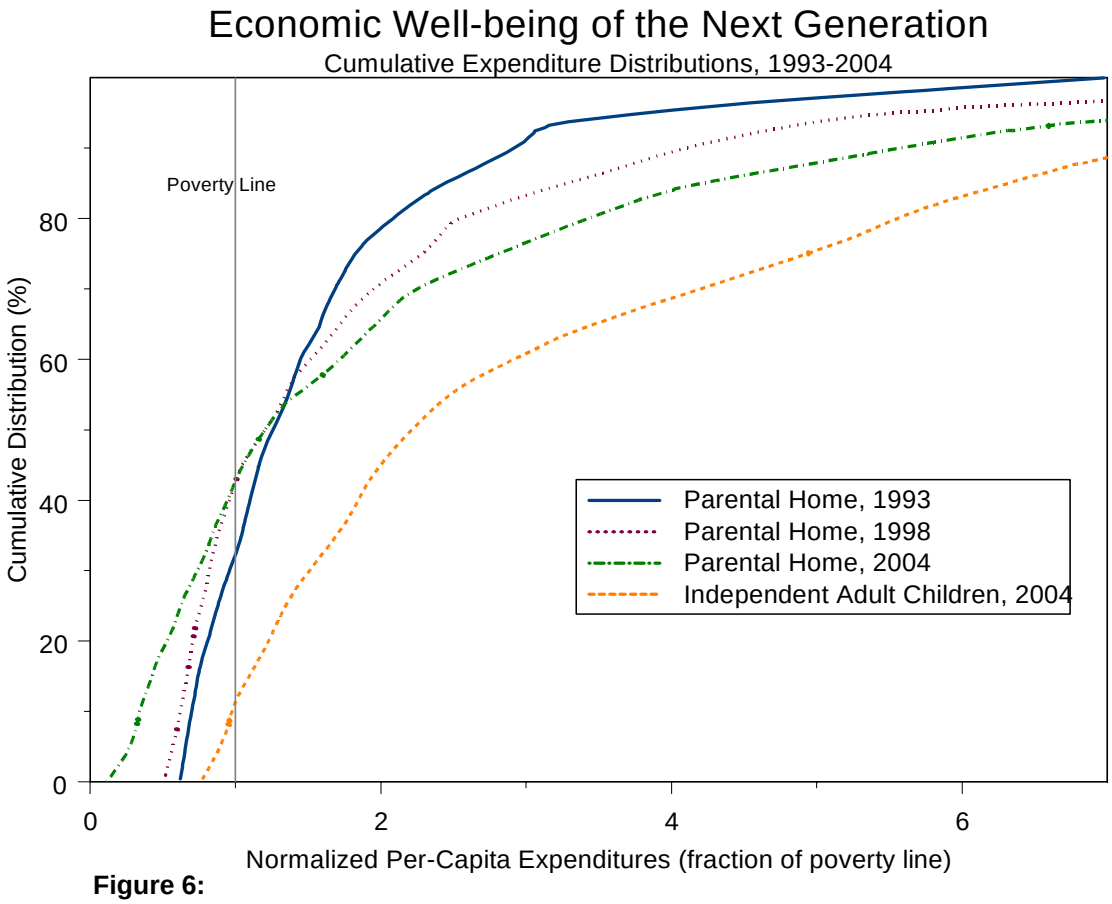

Economic Well-being of the Next Generation
Cumulative Expenditure Distributions, 1993-2004
Poverty Line
Cumulative Distribution (%)
0
20
40
60
80
0
2
4
6
Normalized Per-Capita Expenditures (fraction of poverty line)
Parental Home, 1993
Parental Home, 1998
Parental Home, 2004
Independent Adult Children, 2004

Figure 6: