CENTRE FOR
SOCIAL SCIENCE RESEARCH

Aids and Society Research Unit

# EXAMINING AIDS-RELATED ADULT MORTALITY IN THE KWAZULU-NATAL INCOME DYNAMICS SURVEYS: EMPLOYMENT, EARNINGS AND DIRECT MORTALITY COSTS

Alex Sienaert

Working Paper No. 206
Jointly published by the
CSSR, UCT, and HEARD, UKZN

November 2007

Centre for Social Science Research
A·S·R·U

HEALTH ECONOMICS AND HIV/AIDS RESEARCH DIVISION

Alex Sienaert is a DPhil in Economics candidate at the University of Oxford.

Comments are welcome at: alexis.sienaert@economics.ox.ac.uk

This paper is also available on the HEARD website: http://www.heard.org.za/

# Examining AIDS-related Adult Mortality in the KwaZulu-Natal Income Dynamics Surveys: Employment, Earnings and Direct Mortality Costs


## Abstract

In the KwaZulu-Natal Income Dynamics surveys (covering black households in the province in 3 surveys spanning 1993 to 2004), the mortality rate of 21-50 year-olds rises by 157% from 1993-1998 to 1998-2003. This paper compares this surge in mortality with the ASSA model of the HIV/AIDS epidemic and finds that, while the data appear broadly representative of likely, actual mortality, non-random attrition is a concern. Preliminary findings on the labour market profile of individuals suffering premature adult mortality in the 1998-2004 period, an estimate of foregone earnings, and the direct pre-death care and burial costs associated with their deaths, are presented.


## 1 Introduction

The scale of the HIV/AIDS epidemic in South Africa is staggering. UNAIDS estimates that about 5.5 million people (more than in any other country, and about 12% of the population) are HIV-positive, and that around 1.6 million have already died of AIDS. In the face of this human tragedy, attempts to quantify the economic costs of the epidemic can seem disingenuous; clearly, it causes unspeakable suffering, so why bother to calculate the price tag? There are in fact good reasons for doing so. While forecasting its future path is difficult, it is clear that the epidemic is only now "maturing". Unless a truly massive government treatment intervention takes place, the number of AIDS deaths will soar in the coming years. Quantifying the economic impact of this mortality helps policymakers gauge both the net financing implications of potential interventions, and the material welfare costs of inaction. Any macro-level analysis that attempts to shed light on the economic effects of the epidemic and the responses it elicits should be underpinned by micro-level evidence.

Existing micro-based, plausibly-generalisable evidence on the direct socio-economic effects of HIV and AIDS is limited. This is unsurprising for several reasons. Reliable estimation generally requires longitudinal data, which until recently have not existed (at least on any scale) in South Africa. AIDS-related mortality is particularly difficult to measure because opportunistic diseases, rather than the HI virus directly, ultimately kill (potentially masking the disease in surveys and mortality statistics). And there are major sensitivities – including persistent stigma and the ethics of soliciting information from the bereaved – which complicate the task of researchers in the field.

This paper examines the economic profile of adults who are likely to have died of AIDS, using data from KwaZulu-Natal spanning 1993 to 2004. Specifically, it asks where people whose deaths are probably AIDS-related fit in to the labour market – primarily in terms of their participation and employment rates, and earnings. It also examines expenses directly attributable to deaths (pre-death health care and burial costs), to arrive at an estimate of the magnitude of some direct economic losses associated with mortality. Lost earnings and pre-and post-death expenses are perhaps the most obvious, immediate ways in which the epidemic damages the economic welfare of families, and so need to be factored into evaluations of prevention, treatment and other support programmes.

A related, second motivation for asking this question is based on the usefulness of such quantifications in the broader discourse on HIV/AIDS in South Africa.[1] In this discourse, there is a persistent undercurrent that the epidemic's threat to the country as a whole may be exaggerated because the worst-affected part of the population is that at the bottom of the economic pile – the poor and unemployed. To the extent that the epidemic is concentrated there, so the argument goes, the labour market and wider economy may escape relatively unscathed. There is clearly a need for empirical evidence here. If the epidemic's direct impact on the labour market is and will be significant, then this should be understood as a consequence of inaction, and inform policy decisions. On the other hand, if the epidemic's direct impact on the labour market is and will be small, then it is appropriate for economists to focus on other vectors of economic impact (notably including public and private consumption and savings effects). Either way, this economic information has a role to play in a much deeper, normative debate on the ethics of managing an unfolding epidemic on this scale in a democratic, middle-income country.

[1] This discussion is based on Nattrass (2003).

# 2 Inputs: Data and the ASSA Model

## 2.1 The KIDS Surveys

### 2.1.1 Data Description

Data come from the KwaZulu-Natal Income Dynamics Survey (KIDS).[2] KIDS provides three waves of data from 1993, 1998 and 2004 on individuals in approximately 1000 households in KwaZulu-Natal (KZN). The 1993 wave comprises the 1354 household KZN sub-sample of the national Project for Statistics on Living Standards and Development Survey (PSLSD, conducted according to World Bank Living Standards Measurement survey guidelines) and is designed to be representative at the provincial level.

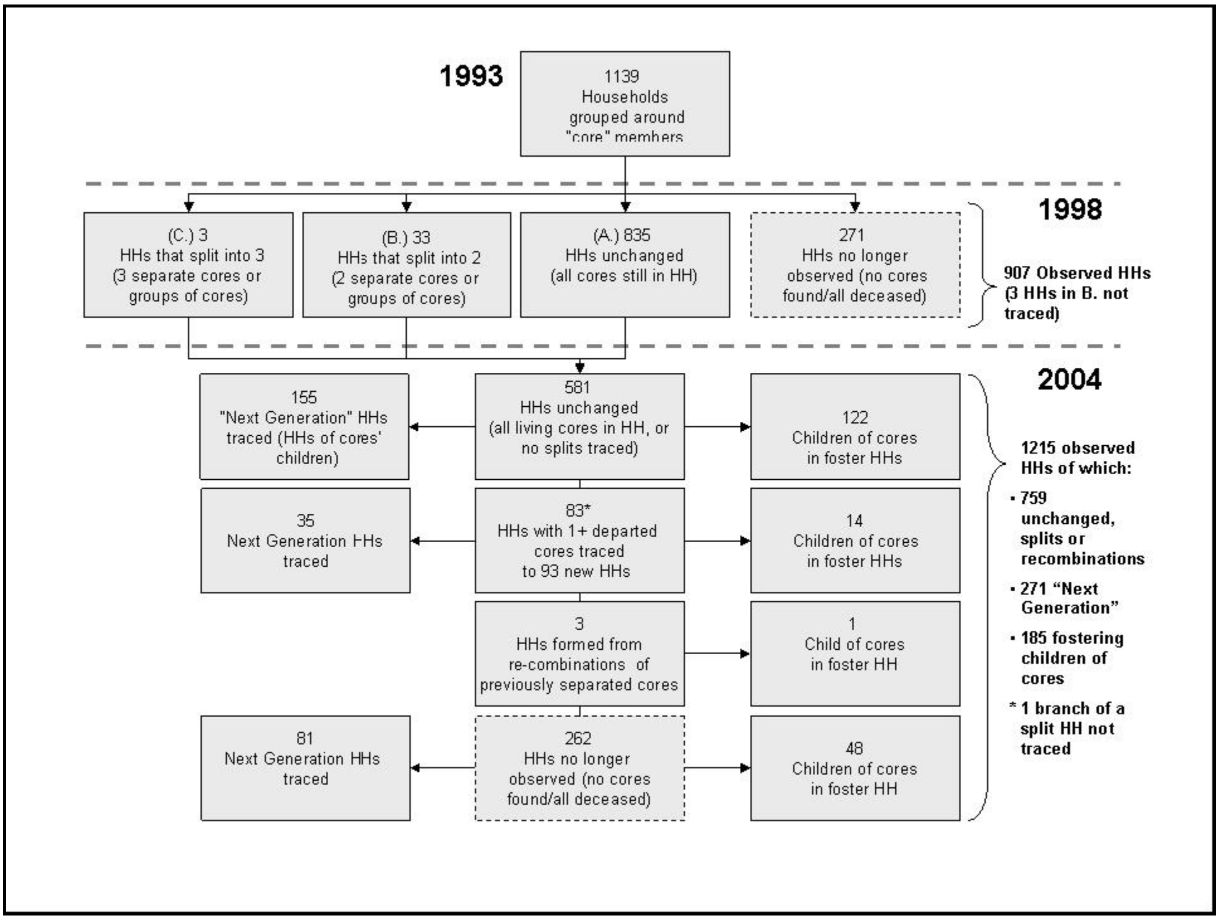
The fundamental sampling unit in the initial wave is the household, defined to comprise individuals who are resident for at least 15 days a year, "eat from the same pot" (when resident) and contribute in some way to a "common resource pool". The survey thus encompasses non-resident household members (very common in the sample), although data on these individuals are limited.

KZN households surveyed in the 1993 round which self-classify as African or Indian are revisited in 1998 and 2004. Because of the typically large size and potentially mutable nature of households, the sampling methodology of these surveys shifts to "core" individuals (adults deemed central to households sampled in 1993, essentially household heads and their immediate relatives).[3] Where core individuals, or groups of core individuals, have split, enumerators attempt to trace them. In addition, the 2004 wave is refreshed with new households formed by the adult children of core individuals, and households fostering core adults' minor children. Figure 1 provides a diagrammatic overview.

[2] KIDS was a collaborative project of the International Food Policy Research Institute, the University of Natal- Durban, the University of Wisconsin-Madison, and the Southern Africa Labour and Development Research Unit at the University of Cape Town. For detailed descriptions, see May, Carter, Haddad and Maluccio (2000) and May, Aguëro, Carter and Timæus (2006b).

[3] Specifically, core individuals are (1) household heads, (2) in three-generation households: children, children-inlaw, or nieces/nephews of heads, aged at least 30, with one or more children of their own in the household, and (3) partners of any of these (May et al. 2006b).

*Figure 1: The KwaZulu-Natal Income Dynamics Surveys: Scheme of Black Households*




Together, the three KIDS waves provide a representative sample of individuals in KZN in 1993, capturing the changing circumstances of the African and Indian sub-sample in two repeat surveys taken 5 and 11 years after that time. While a number of large household surveys have been conducted in post-apartheid South Africa, KIDS is unique in conferring the advantages of longitudinal data at the individual level on a large enough sample to support plausibly generalisable analytical results.

### 2.1.2 Panel Construction

Analysis is restricted to black African individuals. Indian households in KZN are overwhelmingly urban, are smaller and considerably wealthier than African households on average, and are much less likely to house HIV-positive members. Pooling across these groups is thus questionable, and the Indian sub-sample is too small to be independently useful. In addition, the foster households sampled in 2004 are excluded since the information gathered on them is incomplete.

The final panel thus comprises a provincially-representative sample of black African individuals in 1993, with repeat observations on them and their households (including new and nonresident members) in 1998 and 2004, including in the final wave the survey of "next generation" households formed by cores' adult children. The continuing representivity of the panel is dealt with in detail below but, in general, attrition has been found to be of limited importance in most studies using these data, at least up to 1998. KIDS ought to paint a *broadly* accurate picture of socio-economic conditions in the province amongst the black population over the 1993-2004 period.

## 2.2 ASSA Model

The most prominent, non-proprietary demographic and HIV/AIDS model in use in South Africa is the ASSA2003 AIDS and Demographic model of the Actuarial Society of South Africa ("ASSA model"). Released in November 2005, it builds on previous ASSA models and exploits the most recent, reliable data available at that time. The model is widely acknowledged to be the gold standard in modeling the epidemic in South Africa. From the user guide:

> The ASSA2003 model as disseminated has been calibrated to reproduce the patterns of past antenatal clinic survey data and the number of adult deaths. As such, the model represents the triangulation of data from the population census [2001], antenatal survey [2003] and registered deaths [2002/3] by some of the country's top actuaries, demographers and epidemiologists.

The model provides powerful tools for understanding and forecasting the demographic impact of the epidemic (and potential interventions), down to discrete race and province levels. I use it to uncover the likely representivity of adult mortality in the KIDS data, which yields important insights into the external validity of the results obtained. In all cases, I employ the model's default demographic, epidemiological and behavioural assumptions, applied to the black African population of KwaZulu-Natal (corresponding to the universe of the constructed KIDS panel). The model, along with detailed user information, is available at www.assa.org.za/aids; see also Dorrington, Johnson, Bradshaw and Daniel (2006).

# 3 Regional Context

KZN is South Africa's largest province by population (about 9.9 million, or 21% of the country total), and contains its second-largest city (Durban). Roughly 85% of the total provincial population is black, with an aggregate rural-urban divide of 54%-46%. The province's contribution to national GDP stood at 16.7% in 2004. The human development index (HDI) for the provincial black population was 0.5 in 2003, similar to the national (black) average. Income inequality is very high (with a GINI coefficient around 0.71), and about half the population is poor (and a quarter extremely poor) by conventional definitions.[4]

The province is the epicentre of the South African HIV epidemic. South African Department of Health statistics indicate a national HIV prevalence rate of around 11% in 2006, and 16% in KZN. Amongst antenatal clinic attendees (i.e. pregnant women visiting state health clinics) in KwaZulu-Natal, figures suggest annual increases in prevalence to an apparent peak of 40.7% in 2004. The number of AIDS-sick in the province is estimated to have increased from 1,571 in 1993, to 29,885 in 1998, and 167,515 in 2004. Estimated cumulative AIDS deaths in the province stood at 407,862 in 2004 and life expectancy at birth was 43.1 years.[5] Figure 2 summarises some key aggregate prevalence and mortality statistics produced by the ASSA model, for the total black population of KZN, and the sub-population aged 21-50.

Anti-retroviral (ARV) treatment in the public health sector to suppress full-blown AIDS and reduce mother-to-child transmission on any meaningful scale post-dates the 1993-2004 period examined here. [6]

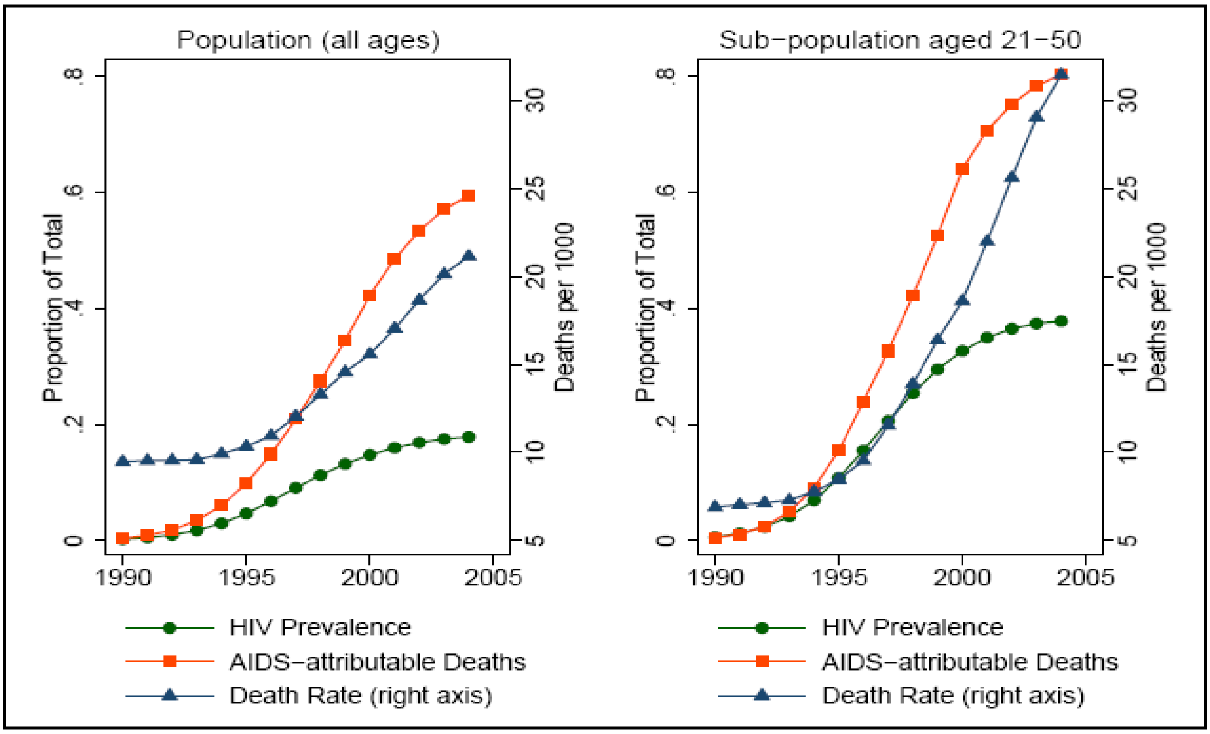
---

[4] Statistics from Statistics South Africa (2004) and Pauw (2005). A common (extreme) poverty line is the 40th (20th) percentile of per adult equivalent income, around 5000 (3000) rand a year (Pauw 2005).

[5] Statistics from Department of Health (2006) and Dorrington et al. (2006).

[6] The public sector rollout of free ARVs began in Gauteng province on 1 April 2004. However, UNAIDS estimates that at most 15% of South Africans who needed antiretroviral drugs were receiving them by mid-2005.

*Figure 2: Selected HIV/AIDS Statistics: ASSA Model (KZN, Black Population)*




# 4 Adult Morbidity & Mortality in KIDS

## 4.1 Morbidity

Health indicators in KIDS comprise antenatal and infant health information, child anthropometry, a record of recent illnesses and (for "core" individuals in 1998 and 2004) basic self-evaluations of health status. Trends in morbidity, while not the focus of this paper, ought to be good leading indicators of mortality. Some results for adults aged 21-50 over the survey period are:

- The proportion of core adults reporting that they feel "poor" or "very poor" (the lowest categories) rises from 5% in 1998 to 26% in 2004.
- In 1998, 17% of core adults who report their health status describe their health as "worse" or "much worse" than 4 years previously, compared to 36% who feel "better" or "much better". The equivalent figures for a similar comparison in 2004 are

almost reversed – 36% and 23%. [7]

- In 1998, 2.5% of the non-searching unemployed who provide data say that illness (distinct from permanent physical or mental disability) is the main reason they are not looking for work. 7.8% do so in 2004.

While these statistics certainly appear suggestive of a deteriorating adult health picture, they need to be treated very cautiously. First, as core members age over the eleven-year period of the surveys, some declines in average reported health indicators are natural even over the chosen 21-50 age range. Second, self-reported health is highly subjective; it is likely to be correlated with other aspects of subjective wellbeing, and prone to reference effects. For example, if the wealthy take basic good health for granted, they may focus more on minor ailments than the poor, obscuring true health gains for those with rising incomes over time. Lastly, inference of time trends from broad health categorisations is vulnerable to transitory (including seasonal) health events affecting some surveys but not others.

The real evidence of worsening health over the period lies in dramatic changes in recorded deaths between surveys.

## 4.2 Mortality

264 individuals sampled in 1993 are deceased in 1998; their mean age at death is 50. 480 individuals sampled in 1998 are deceased in 2004, with a mean age at death of 44. Taking into account the fact that there is one more year between the second and third surveys than between the first and second, this is a 52% rise in the average number of deaths per year, and a 6 year decline in the average age at death. It is important to note that, because of the survey methodology, these deaths exclude most infant mortality. [8]

Looking at these numbers, a Martian visitor might assume that they arise from substantial changes in the age distribution of the underlying samples between periods. If exactly the same individuals were observed in each wave, we would of course expected a uniform, 11-year rise in ages from 1993 to 2004. In fact, the mix of individuals observed changes as demographic change occurs –

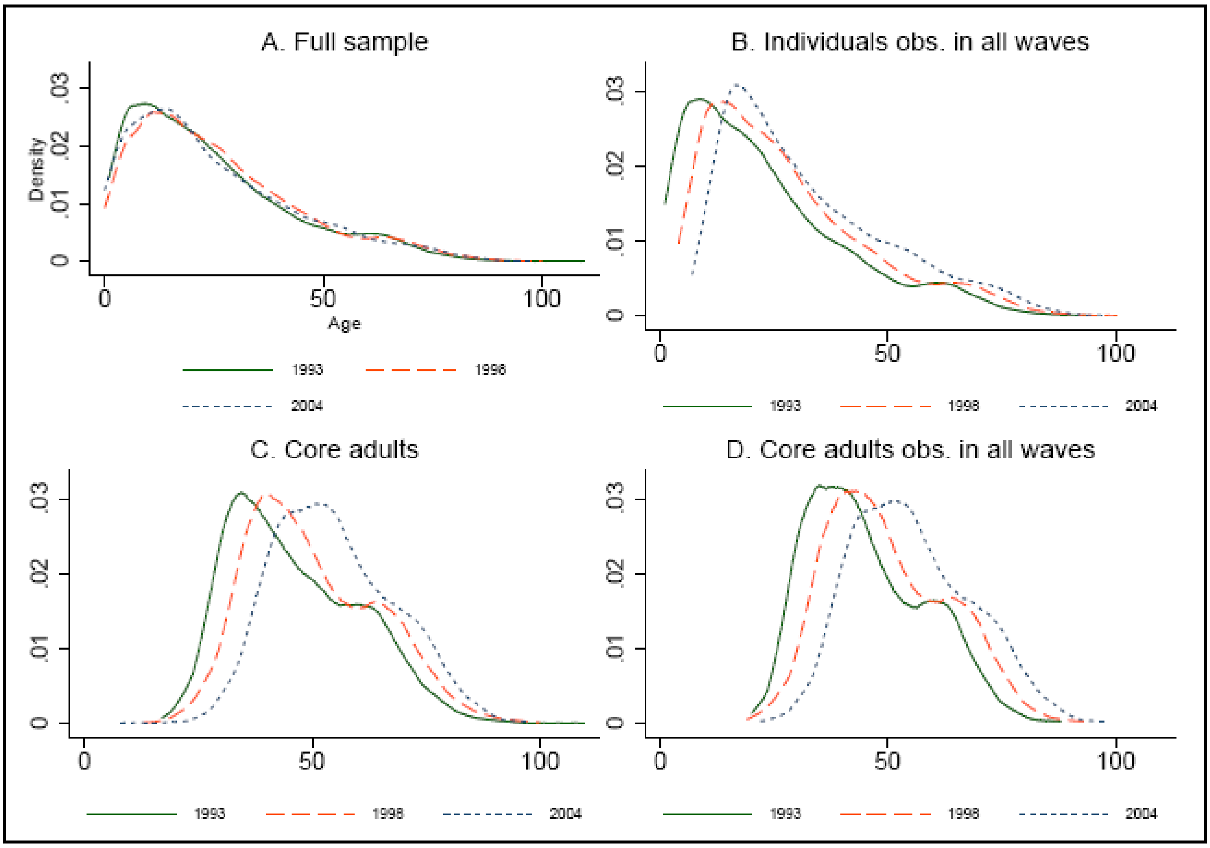
[7] To give respondents in 1998 and 2004 a reference point, they are asked to think back to the time of the previous national election (1994 and 1999, respectively). Consequently, the evaluation period is a year longer in 2004 than in 1998.

[8] Only the deaths of individuals present in the previous survey are canvassed in 1998 and 2004. Consequently, children who are born and die between surveys go unrecorded.

households add members (notably through births and marriages) and lose members (notably through deaths, separations, and departures of younger members to start their own families). In addition, sample attrition may not be random along the age dimension.

Figure 3 displays approximate (kernel) densities of the ages of living individuals in each survey. The densities for all sampled individuals (panel A) appear very similar. However, Kolmogorov Smirnov tests in fact reject the equality of any pair of distributions. Essentially, there is some ageing in the total samples from 1993 to 1998 (fewer young people), while 2004 sees some thinning out of the 25-45 population relative to 1998, and a rise in the proportion of young people (which is expected since the panel is refreshed with "next generation" individuals in this wave). Amongst individuals observed in all three periods (panel B) and core adults (panels C and D), we see the expected rightward-shifts in age densities over time. Such ageing might be expected to trigger a rise in mortality, but this is clearly not consistent with the observed decrease in the average age at death. In short, age density differences across periods certainly do not appear large enough to explain the large, observed mean changes in mortality and age at death between the 1993 to 1998, and 1998 to 2004, periods.

*Figure 3: Age Densities of Living Individuals by KIDS Survey Year (Epanechnikov Kernel Estimates, Default Bandwidths)*



Figures 4 and 5 report the ages in 1993 and 1998 of all individuals aged 10-60 in those years who die between 1993 and 1998, and 1998 and 2004, respectively (orange, dotted lines, right axes). Deaths from 1993 to 1998 (figure 4) in this age range appear fairly uniform, and the data look noisy because there are very few observations (0-7) per age. Deaths from 1998 to 2004 (figure 5) look very different. Deaths amongst 10-20 and 45-60 year-olds are comparable to those in figure 4, but there is a large bulge in the deaths of 21-45 year-olds. The difference is easy to see when the data are smoothed using a median spline function (cubic spline fitted to medians of 5-year age intervals, blue lines).

*Figure 4: Total Deaths: ASSA Model & KIDS (1993-1998)*

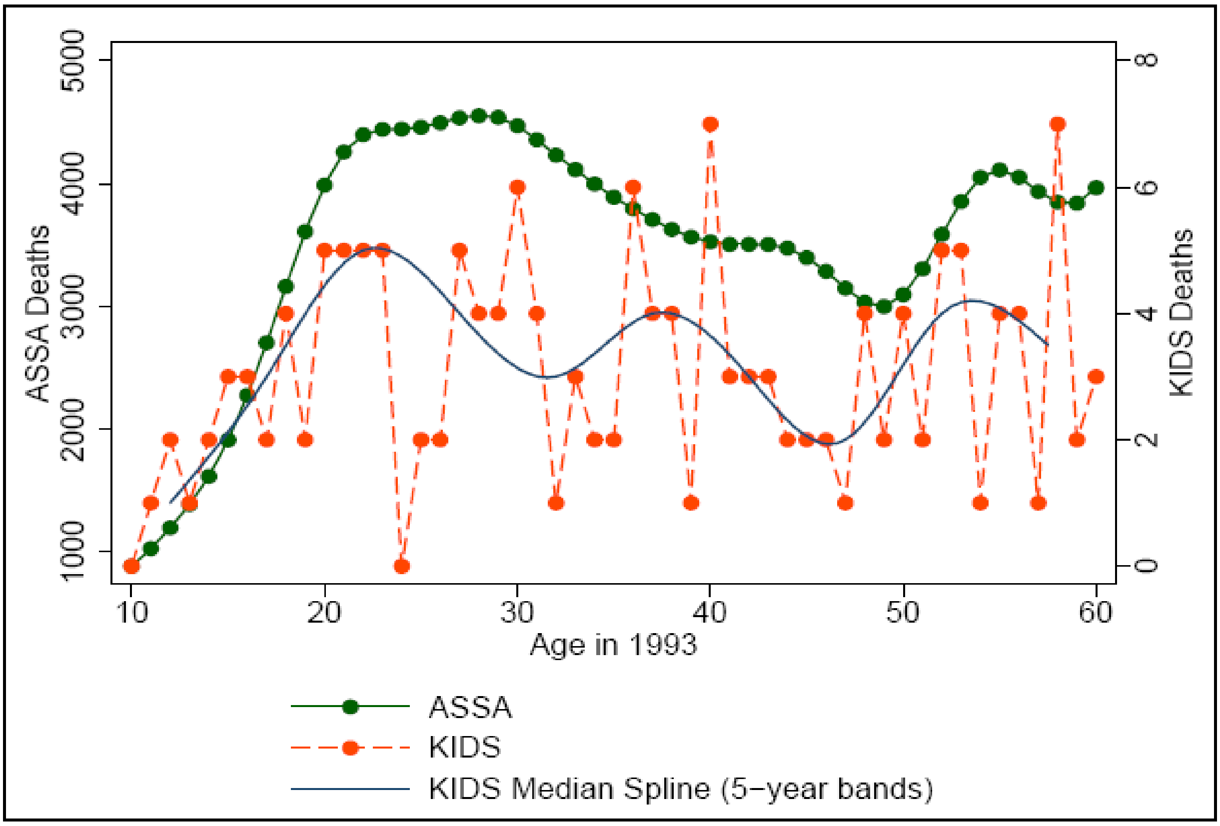


*Figure 5: Total Deaths: ASSA Model & KIDS (1998-2004)*

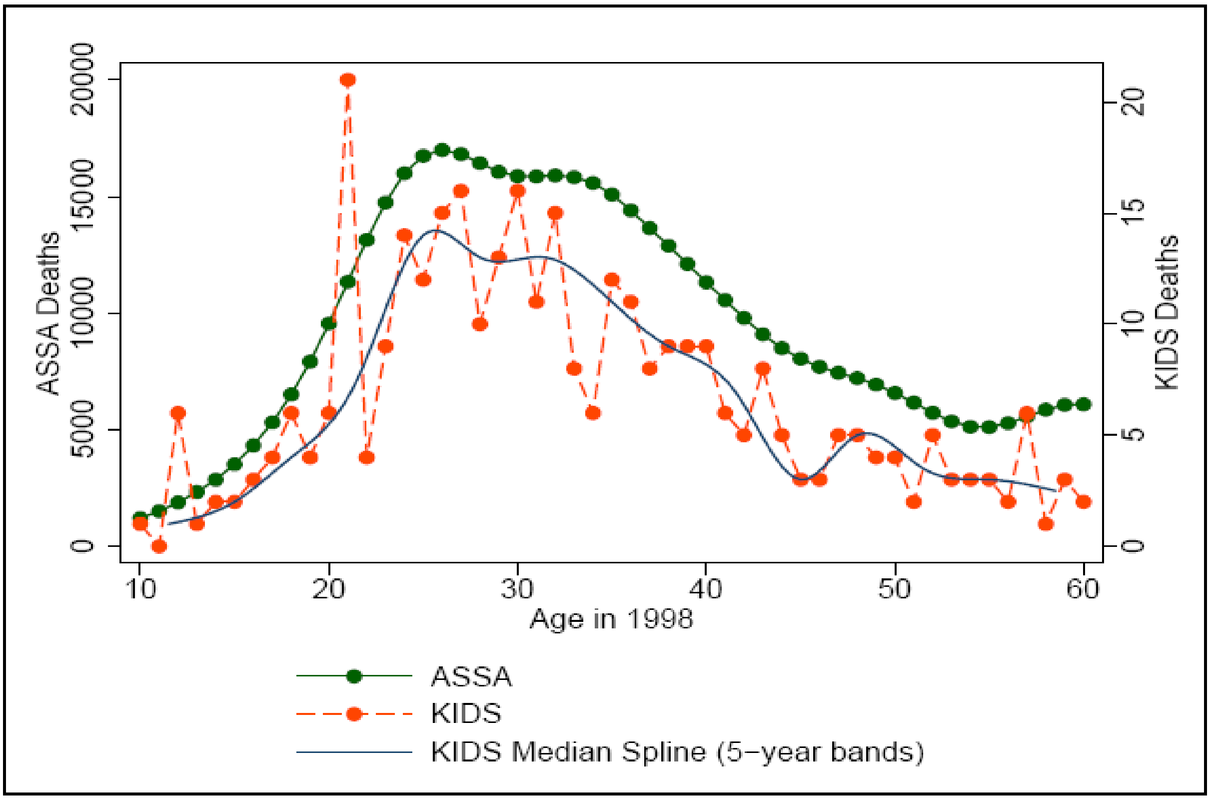
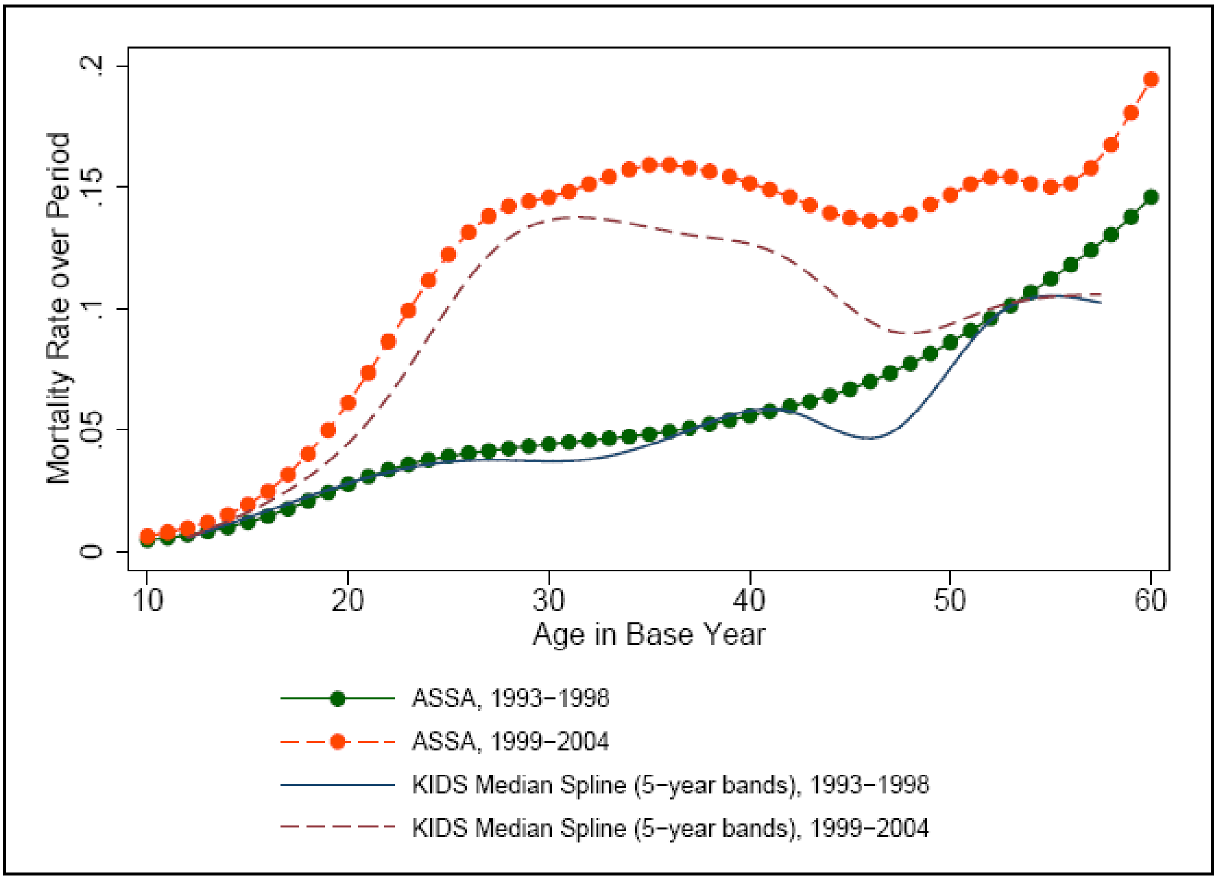



Figure 6 combines all the above information, standardising raw deaths by age against the age profiles of the underlying samples. By looking at the proportion of individuals at each age (observed alive or dead in both periods) who die between periods, we arrive at the mortality rate for each age over the period. Put another way, figure 6 reports the average probability that someone of a given age in 1993 and 1998 (respectively) dies before the next survey (1998 and 2004, respectively). Focusing first on KIDS, we are concerned with the dot-less lines, which (again, to draw out the pattern of the data) are smoothed using the medians of 5-year age buckets.[9] The 1993-1998 mortality curve looks exactly as we might expect under normal conditions: the risk of death in the subsequent 5 years rises gradually with age, appearing to accelerate somewhat towards 60. The 1998-2004 mortality curve, however, is striking. The "hump" in deaths of people aged 21-45 in 1998 seen in figure 5 is clearly visible after standardising on the underlying age distribution. For example, the probability of a 30-year old in 1993 dying before 1998 is about 4%. A 30-year old in 1998, however, has a 14% chance of death by 2004 (a 192% rise in mortality risk, accounting for the extra year from 1998 to 2004).

[9] For the raw data, see figure 14 (appendix).

*Figure 6: Aggregate Mortality Rates: ASSA Model & KIDS (1993-1998 & 1999-2004)*




## 4.3 External Validity

While this change in mortality is clearly not an artefact of changes in sample age profiles, it remains the case that only the initial, 1993 KIDS survey is known to be provincially representative. It is possible, therefore, that the rise in mortality is caused by some other idiosyncracy in the KIDS data. If so, results based these data will lack external validity, and could be misleading.

An obvious way to check the representivity of mortality in KIDS is to compare it to the best available view of actual demographic and mortality trends in the black population of KwaZulu-Natal. Figures 4, 5 and 6 therefore include equivalent information produced by the ASSA model (left axes). The age profile of observed deaths in KIDS clearly follows the same pattern as that of the ASSA model (note the differences in axis scales).

Figure 6 is key, since standardising against the relevant populations (the provincial population for ASSA, and the observed sample for KIDS), makes the

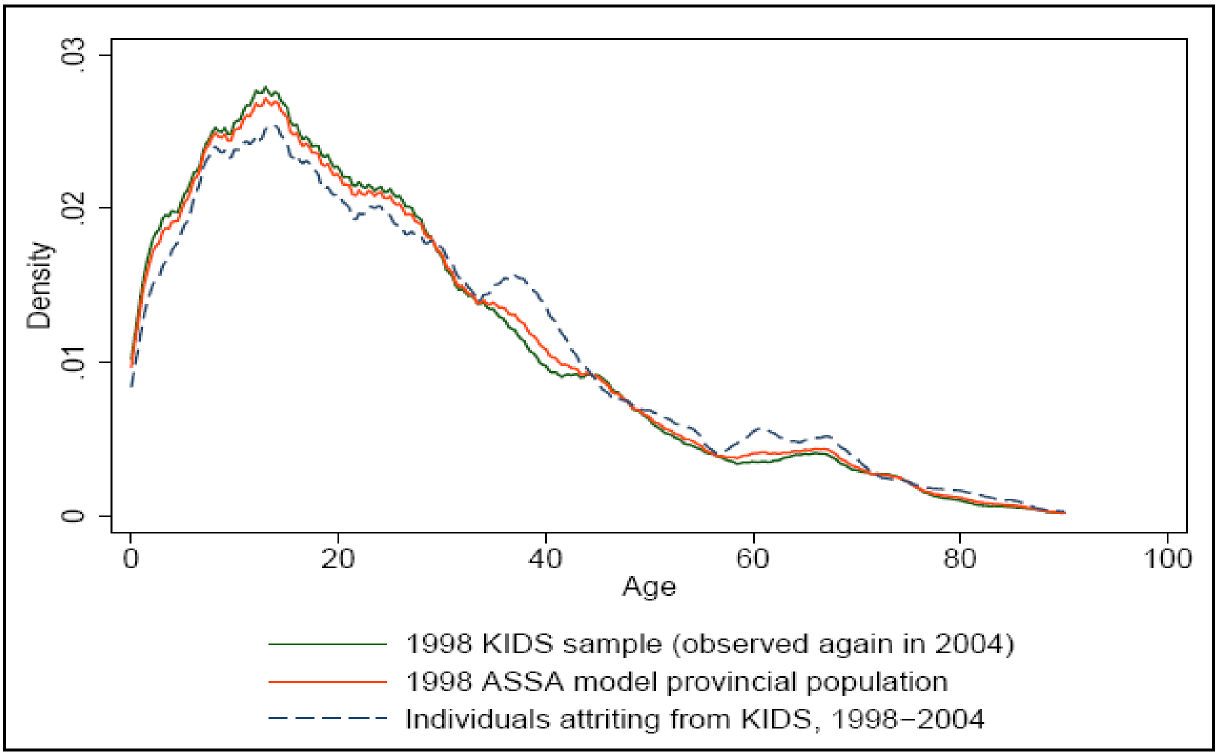
ASSA and KIDS mortality curves directly comparable. For the 1993-1998 period, KIDS follows the ASSA numbers very closely, although we observe an apparent under-estimation of mortality amongst 45-50 year-olds. This fluctuation, however, is not surprising given the limited number of observations over this period and is likely to simply be a blip. For the 1998-2004 period, the ASSA model also displays the hump in mortality amongst 21-45 year-olds that is so noticeable in KIDS. However, KIDS appears to systematically slightly underestimate mortality in the 20-30 age range, with the gap widening thereafter and becoming particularly notable in the 40+ age range.

A more precise estimate of the underestimation of mortality over the 1998-2004 period in KIDS is easy to obtain. In KIDS, 7.5% of the sample aged 10 to 58 in 1998 (and observed again in 2004) dies before the 2004 survey.[10] The ASSA model equivalent is 455,280 deaths out of a sub-population of 5,165,363 (8.8%). This suggests that KIDS underestimates mortality amongst 10 to 58 year-olds in 1998 over the 1998 to 2004 period by just over 15%. Figure 6 indicates that most of this underestimation occurs in the upper end of the age range.

To the extent that the ASSA model is accurate, the 1998 KIDS sample has either become demographically unrepresentative of the population of interest, or non-random attrition (related to mortality) has occurred from 1998-2004. Figure 7 unpacks this by comparing the age density of the sample used to evaluate mortality (all individuals alive in 1998 who are observed again [alive or dead] in 2004) with the age density of the actual population generated by the ASSA model, and the age density of individuals observed in 1998 who attrite from the survey before 2004. Kernel density estimates can be misleading, but these were carefully evaluated to obtain the optimal bandwidth (essentially, degree of smoothing). They suggest strongly that attrition, rather than a skewed 1998 sample, is behind the under-sampling of mortality in the 1998-2004 period. The age profile of the ASSA and KIDS samples in 1998 are very similar, *except* for some under-sampling around the 40 age range in KIDS, which (remembering that this sample comprises individuals observed again in 2004) is explained by the age profile of attriting individuals; notice the bulge around 40. In general, the density of attriters is shifted to the right: older people appear more likely to vanish from the survey between 1998 and 2004 than younger.

[10] The 10-58 age range is designed to capture almost all individuals who might be of working age at the time of death.

*Figure 7: Age Densities (Epanechnikov Kernel Estimates, 1.5 Bandwidth): KIDS (Non-Attriters & Attriters) & ASSA Figures (1998)*




Non-random attrition of this kind poses a potential threat to inference. Tables 1 (mean comparisons) and 2 (probit regression on the probability of 21-45 year-olds in 1998 attriting) examine the characteristics of attriting individuals more closely. Clear differences emerge, with attrition more likely amongst non-residents, and people living in smaller and relatively better-off households. Formally employed people are more likely to attrite, and the average monthly earnings of future attriters are (significantly) higher. The probit on table 2 confirms these patterns in partial correlation terms, but goodness of fit is an issue and (in general) regression models do not perform well in predicting future attrition.

*Table 1: Attriters (1998-2004) and non-Attriters, 1998 (Means, Std. Deviations in brackets)*

| | 1. All adults 21+ | | 2. Adults 21-45 | |
|---|---|---|---|---|
| | 1.1 Non-attriters | 1.2 Attriters | 2.1 Non-attriters | 2.2 Attriters |
| Demographics & HH Chars. | | | | |
| Age | 38.73 | 40.35** | 30.89 | 31.59** |
| | (15.28) | (15.74) | (6.99) | (6.96) |
| Proportion male | 0.43 | 0.42 | 0.44 | 0.45 |
| | (0.50) | (0.49) | (0.50) | (0.50) |
| Education (yrs.) | 5.52 | 5.38 | 6.52 | 6.59 |
| | (3.68) | (3.80) | (3.41) | (3.44) |
| Prop. resident | 0.68 | 0.51** | 0.62 | 0.49** |
| | (0.47) | (0.50) | (0.48) | (0.50) |
| HH size | 10.34 | 8.03** | 10.50 | 8.17** |
| | (5.26) | (4.72) | (5.27) | (4.80) |
| HH income (equiv.) | 376.23 | 457.61** | 373.99 | 457.94** |
| | (482.57) | (677.96) | (472.42) | (538.39) |
| Labour and Earnings | | | | |
| Labour force participant | 0.67 | 0.70* | 0.78 | 0.82** |
| | (0.47) | (0.46) | (0.42) | (0.38) |
| Employment rate | 0.56 | 0.63** | 0.52 | 0.61** |
| | (0.50) | (0.48) | (0.50) | (0.49) |
| Formal employment rate | 0.40 | 0.49** | 0.38 | 0.49** |
| | (0.49) | (0.50) | (0.49) | (0.50) |
| Informal employment rate | 0.15 | 0.14 | 0.14 | 0.12+ |
| | (0.36) | (0.35) | (0.35) | (0.33) |
| Monthly earnings | 374.68 | 470.78* | 419.30 | 528.79* |
| | (941.06) | (1060.53) | (1016.57) | (994.38) |
| Observations | 6182 | 2398 | 2318 | 936 |

+, *, ** denote one-sided t-test rejection of equality at the 10%, 5% and 1% levels.

*Table 2: PROBIT REGRESSION: ATTRITION PROBABILITY (21-45 YEAR-OLDS, 1998-2004)*

| | |
|---|---|
| Age | 0.000 |
| | (0.001) |
| Gender (male=1) | -0.025 |
| | (0.016) |
| Education (yrs.) | -0.001 |
| | (0.003) |
| Resident dummy | -0.083** |
| | (0.021) |
| HH Size | -0.018** |
| | (0.003) |
| Unemployed | 0.019 |
| | (0.025) |
| Formal sector employed | 0.086** |
| | (0.028) |
| Informal sector employed | 0.005 |
| | (0.037) |
| N | 3237 |
| Log Likelihood | -1837.6 |
| Pseudo $R2$ | 0.05 |
| Wald test statistic | 60.6 |

Dep. var. is 1 if individual attrites 1998-2004, 0 otherwise (including if individual has died but is observed). Coefficients report marginal effects at the mean (or effect of 0 to 1 change in case of binary variables). Omitted labour category is out of the labour force. Std. errors are heteroskedasticity-robust and corrected for within-cluster correlation. $^{+}$, *, ** denote significance at 10%, 5% and 1% levels.

In summary, adult mortality in the KIDS panel changes dramatically between waves, with a sharp rise in the mortality of people of young-to mid-working age. This mirrors the results of the ASSA model based on the best available modeling of the AIDS epidemic quite closely, except that there is some under-sampling (on the order of 15%) of the mortality of working-age adults, concentrated amongst adults aged above 40. The most likely explanation is that attrition is related to mortality, and especially the mortality of relatively older "prime age" adults. A plausible hypothesis is that households which experience adult deaths are more likely to migrate or break up than average (and as a result are less likely to have been traced in 2004). We might expect that working age adults over 40 are especially central to households (e.g. main breadwinners), so that their loss makes household dissolution especially likely.

### 4.3.1 Existing, Related Work using KIDS

The rise in young adult mortality reported above is described in May, Aguëro and Carter (2006a), who highlight it as an area for future research. Carter, May, Aguëro and Ravindranath (2007) examine the economic welfare impact of premature adult mortality over the 1998-2004 period, using a household *per capita* expenditure approach adapted from Beegle, De Weerdt and Dercon (2006) for Tanzania. On this basis, the adult mortality impacts on household welfare appear most pronounced for relatively better-off families (above the poverty line). The authors acknowledge that the possibility that household size is endogenous to mortality qualifies these results. In addition, the robustness of welfare estimates based on extrapolating a single-period trend may be questionable, and if families have recourse to short-term ways of smoothing consumption (e.g. sale of assets), consumption data gathered relatively soon after a mortality shock may substantially underestimate permanent welfare costs (Beegle et al. 2006). In attempting to understand material welfare impacts, then, scope clearly remains for estimating the earnings lost and direct costs associated with death as likely, major contributors to welfare losses.

Yamauchi, Buthelezi and Velia (2006) examine the effects of adult mortality (1998-2004) on female labour supply, and adolescents' school to labour market transitions. They find tangible, positive rises in female labour supply in response to adult deaths, and evidence of accelerated school-leaving by older children, consistent with increased care and cash earnings demands in affected households. As the paper notes, survival strategies such as these may have long-term welfare consequences (including reduced human capital accumulation), and are examples of other, important costs of mortality which add to the consequences that are the focus of the present paper.

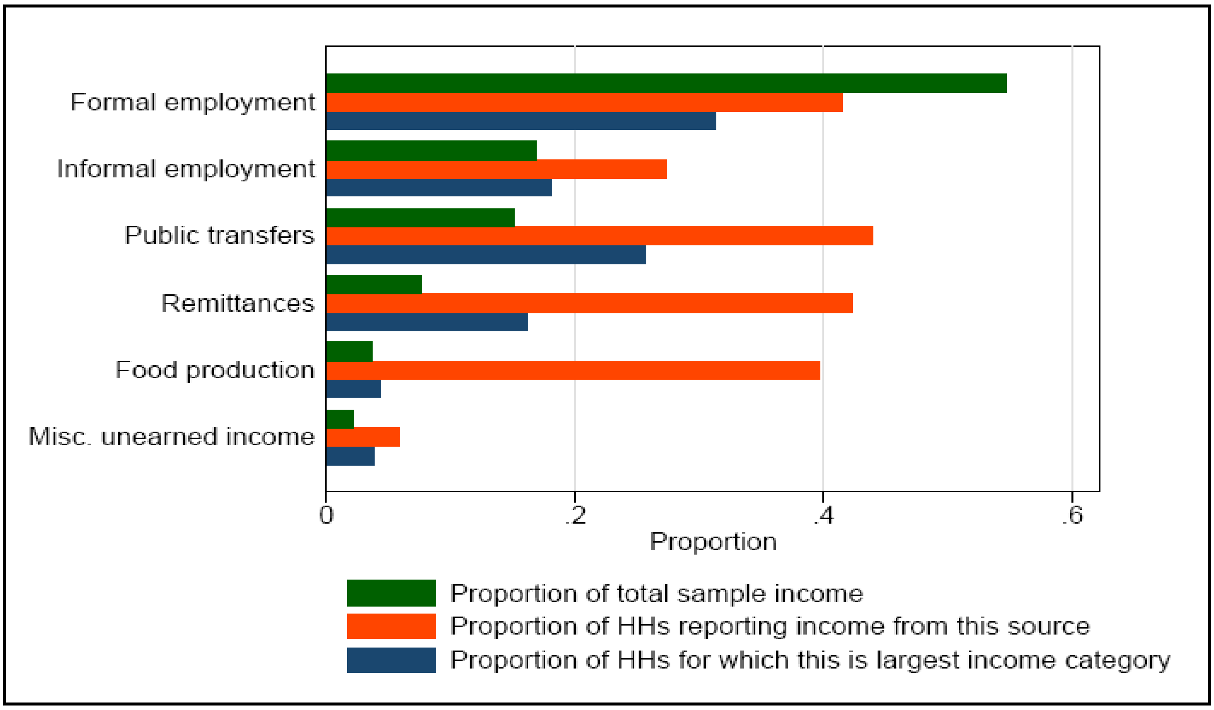
# 5 Mortality in the Labour Market, 1998-2004

I now turn to the selected characteristics of individuals likely to have died of AIDS. After very briefly contextualising employment and earnings in the data, the construction of the evaluation and comparison groups is described, followed by results. Labour marker definitions aim to follow Cichello, Fields and Leibbrandt (2005), and are outlined (along with key assumptions and imputations) in the methodological appendix (section B). Briefly, labour force participation is broadly defined (to include the unemployed who want work but are not actively searching), and employment differentiates between formal and informal sector employment. These definitions reflect existing analysis of labour dynamics in the KIDS data (1993-1998), and the predominant view in the empirical, South African labour literature that labour force participation should be defined broadly, and that important differences exist between formal and informal sector employment.

## 5.1 Context -Labour and Income Indicators

Figure 8 displays income by source in the KIDS 1998 survey. Formal employment earnings account for the bulk of observed income, followed by informal employment earnings, public transfers and remittances. Food production generates a remarkably small proportion of income considering that half the sample (at the household level) is rural (though this is a well-known feature of income composition in South Africa and KZN).[11]

*Figure 8: Income Source Decomposition: KIDS Sample (1998)*




Examining the proportion of households reporting particular kinds of income, and the proportion for which a source is the most important income-generator, may yield a better indication of the income mix than aggregate income proportions. These measures suggest a bigger role for public transfers and remittances in sustaining livelihoods (and demonstrate that many households are

[11] Food production includes the estimated net value of subsistence production. Since many imputations are involved, these figures should be treated as indicative.

involved in at least some food production even if its net value and relative importance is very small), and suggest a considerably more diverse mix of income sources. Nonetheless, formal employment earnings are clearly of predominant importance, followed by informal employment earnings and public transfers in roughly equal measure. Access to the labour market, and formal sector employment in particular, is the major determinant of material welfare in the KIDS survey (Cichello et al. 2005), mirroring the picture for the South African economy as a whole (see amongst others Bhorat, Leibbrandt and Woolard (2001) and Bhorat (2004)).

For detailed examinations of poverty, earnings dynamics and inequality in KwaZulu-Natal (all based on the 1993 and 1998 KIDS surveys), see Carter and May (2001), Keswell (2004) and Cichello et al. (2005).

## 5.2 Mortality & Comparison Groups

### 5.2.1 AIDS-related mortality sub-sample

We are interested in the labour market characteristics of individuals who are likely to have died of AIDS, over the period 1993-2004.

The 1998-2004 period is pivotal since this is when we observe the sharp rise in adult mortality which is likely attributable to AIDS. Furthermore, the 2004 survey (in which these deaths are recorded) is unique in allowing deaths from illness to be differentiated from deaths due to accident or injury. Retaining deaths from illness between 1998 and 2004 leaves us with 381 deaths. Since the focus is on working age, AIDS-related deaths, I further restrict deaths under consideration to individuals aged 21-45 in 1998, allowing ages at death to range from 21 to 51 – individuals in prime working life. Individuals who are AIDS-sick at the time of the 1998 survey will clearly be disadvantaged in the labour market (see, for example, Rosen, Feeley, Connelly and Simon (2007) and Fox, Rosen, MacLeod, Wasunna, Bii, Foglia and Simon (2004) for evidence from other African contexts), and their inclusion will contaminate the baseline for answering the question at hand. I therefore exclude people dying in 1998 (11 cases), and the 2 individuals who die in 1999 who are recorded to have been unable to perform their usual activities due to illness for longer than 6 months. One further check is possible: respondents identify household members who have been ill in the 15 days before the 1998 survey, and record whether they are still ill. To minimise the risk that the 1998 labour markers of the mortality group already reflect the impact of AIDS, I also exclude these individuals if they die before 2001 (4 cases).[12] This leaves 203 deaths which, following Carter et al. (2007), are categorised as the "premature adult mortality" (PAM) cohort.

While we know only that these deaths are caused by illness rather than accident or injury, the genders, ages and years of death of the PAM cohort can be plugged into the ASSA model to produce an estimate of the probability that AIDS is the cause of death. On this basis, the model predicts that 79% of these deaths are attributable to AIDS.[13]

### 5.2.2 Comparison cohorts

PAM cohort characteristics need to be contextualised, necessitating careful consideration of appropriate comparison groups. The least restrictive option is to examine all adults of working age in the 1998 sample. But the PAM cohort represents a non-random selection of deaths within the working age sample, chosen to focus on likely AIDS-related mortality. Consequently, comparing them to all others is not appropriate, and we are not interested in the labour market close to life-cycle transition points such as retirement. The following comparison groups are constructed:

1. All individuals in the same age group as PAMs (21-45 in 1998) who remain alive from 1998-2004. For methodological consistency, individuals who are reported sick at the time of the 1998 survey are excluded. A risk is that individuals in this comparison group, while still alive in 2004, are AIDS-ill by that time (this matters because, illustratively, someone who dies the day after the 2004 survey should in principle be in the PAM group). To minimise this risk, albeit crudely, individuals who are reported sick at the time of the 2004 survey are also excluded. This group thus comprises all individuals of the same age as PAMs who, as far as it is possible to tell, are healthy in both the

[12] This leaves 20 deaths in 1999, with a maximum period of debilitating illness prior to death of 5 months. The estimated, median period of serious illness preceding death from non-ARV-treated, full-blown AIDS is 18 months. In practice, dropping them from the sample has no material impact on results. Nor does retention of the sick at the time of the 1998 survey.

[13] The ASSA model suggests that 68% of all KIDS deaths of 21-45 year-olds in 1998 over the 1999-2004 period (including from accident or injury) are AIDS-related. The probability that any one of these deaths is from illness is 85%. Since all AIDS deaths are deaths from illness, the (average) probability that a death from illness (the PAMcases) is caused by AIDS is $\frac{0.68}{0.85} = 0.79$: Equivalently, by Bayes' rule (with everything conditioned on death): $P(AIDS|illness) = \frac{P(illness|AIDS) x P(AIDS)}{P(illness)} = \frac{1x0{:}68}{0{:}85}$

1998 and 2004 surveys. Since this is the least restrictive comparison group, this is termed the "broad" cohort.

2 Individuals in the broad cohort who live in families which experience PAM are directly affected by the epidemic. Excluding them yields a more restrictive comparison group, designed to approach a less AIDS-affected sub-sample. This restriction can be strengthened by also excluding families which have experienced a prime age adult death in 1993-1998. Such households may be at the forefront of the AIDS epidemic, making their inclusion in the less-AIDS-affected group inappropriate. Group 2, then, contains people of the same age as PAMs, who appear healthy in the surveys, and live in families which record no working age adult deaths between 1993 and 2004. Since this is a sub-set of the broad cohort, this is termed the "reduced" cohort.

3 A final comparison of interest is essentially the opposite of the reduced cohort: considering working age people in households experiencing PAM helps illuminate what economic roles PAMs play within their own families. In this case, the relevant comparison is all non-PAMs of working age in these households. I retain a margin around strict working age (16-59 and 16-64 for women and men respectively, using state old age pension eligibility as the upper bound) in order to avoid life-cycle driven transitions into and out of the labour market. This comparison group thus comprises adults aged 21-58 in 1998, living in PAM-affected households. This is the "family" cohort.

The appropriateness of all of these cohorts is open to debate. It is important, therefore, to be clear about our aims.

The first is to examine the economic profile of people who die of AIDS, before they become AIDS-sick. Consequently, the PAM cohort concentrates on adult deaths that are statistically likely to be AIDS-related over the period (21-45 year-olds in the base year who die from illness), and tries to avoid including individuals who are already AIDS-sick when surveyed. Looking at this group alone allows us to say something about the direct earnings losses (and other immediate death costs) associated with AIDS over the period.

A second aim is to contextualise this economic profile, which involves comparing the PAM cohort to something. This is a challenge because, in the terminology of the programme evaluation literature, AIDS is a non-random “treatment” (in particular, it is likely to be related to largely unobserved behavioural and other risk factors). Nonetheless, we can in principle compare the PAM cohort to all other working age adults. In practise, we know that age plays an important role in driving labour force outcomes, and so standardising through comparison with people of the same age is desirable. In addition, inclusion of AIDS-sick people in the comparison group will dilute any real differences between AIDS sufferers and “the rest”, and this too is avoided as far as possible (this is the broad cohort). The reduced cohort restricts the comparison further, to a group that is likely to be less directly affected (particularly in 2004 labour outcome terms). While the broad cohort is of primary interest for comparisons with the pre-AIDS morbidity economic profile of the PAM cohort, the reduced cohort will play an important role in the foregone earnings estimates to follow (section 6.2). The final (family) cohort allows for a different comparison, between likely AIDS victims and their own household members.

In the author’s view, these groups are a sensible way to contextualise the experience of future AIDS sufferers in 1998. But it should be acknowledged that they involve a trade-off between the need to adjust the sample to yield useful comparisons, and the risk of introducing unanticipated biases through selection. Crucially, these comparison groups should *not* be viewed as counterfactuals. For example, no claim is being made that people in the PAM cohort are “the same” as individuals in, say, the reduced cohort, except that they have the misfortune to die (probably from AIDS) before 2004.

## 5.3 Descriptive Statistics

Basic demographic descriptive statistics are presented in table 3. Notable features include the fact that the strict comparison group picks up proportionally more (self-declared) household heads than the PAM cohort, and that PAM individuals are a bit less likely to have partners in households than individuals in the comparison groups. The high proportion of non-resident working age people in households should be noted (including 39% of PAM individuals). There is little doubt that the inclusion of non-residents as active family members is appropriate, given the continuing importance of migrant labour. At under 10%, the proportion of people in the PAM cohort who are new to their survey households since 1993 appears no different to that for comparison individuals (one source of household-level indicator endogeneity stems from the possibility of the AIDS-affected moving into new households for care).

On the basis of self-reported primary activity, individuals in the PAM cohort are as likely to be employed, in full-time education and unemployed (except compared to the reduced cohort, against which they are 6% more likely to describe themselves as unemployed [significant at the 10% level]).

*Table 3: Basic Indicators: PAMs & Comparison Cohorts, 1998 (Means, Std. Deviations in brackets)*

| | 1. PAM | 2. Broad | 3. Reduced | 4. Family |
|---|---|---|---|---|
| **Demographics** | | | | |
| Age | 31.60 | 30.55* | 30.54* | 32.97 |
| | (6.77) | (6.92) | (6.94) | (9.21) |
| Proportion male | 0.44 | 0.44 | 0.45 | 0.45 |
| | (0.50) | (0.50) | (0.50) | (0.50) |
| Education (yrs.) | 6.70 | 6.61 | 6.64 | 6.31+ |
| | (3.29) | (3.41) | (3.39) | (3.42) |
| **HH Role & Links** | | | | |
| HH head/partner | 0.12 | 0.14 | 0.19* | 0.14 |
| | (0.33) | (0.35) | (0.39) | (0.35) |
| Child/child-in-law of head | 0.62 | 0.61 | 0.62 | 0.55+ |
| | (0.49) | (0.49) | (0.48) | (0.50) |
| Has partner in HH | 0.16 | 0.21+ | 0.22* | 0.21* |
| | (0.37) | (0.40) | (0.41) | (0.41) |
| Has minor children in HH | 0.43 | 0.40 | 0.42 | 0.38 |
| | (0.50) | (0.49) | (0.49) | (0.49) |
| Resident in HH | 0.61 | 0.61 | 0.59 | 0.60 |
| | (0.49) | (0.49) | (0.49) | (0.49) |
| New to HH since 1993 | 0.09 | 0.08 | . | 0.07 |
| | (0.28) | (0.27) | . | (0.25) |
| Non-member with links | 0.01 | 0.02 | 0.00 | 0.03 |
| | (0.12) | (0.14) | (0.00) | (0.18) |
| **Self-declared Main Activity** | | | | |
| Employment | 0.40 | 0.41 | 0.43 | 0.41 |
| | (0.49) | (0.49) | (0.50) | (0.49) |
| Unemployment | 0.44 | 0.40 | 0.38+ | 0.41 |
| | (0.50) | (0.49) | (0.49) | (0.49) |
| Education | 0.10 | 0.11 | 0.11 | 0.10 |
| | (0.31) | (0.32) | (0.31) | (0.30) |
| Observations | 203 | 1884 | 1115 | 551 |

+, *, ** denote one-sided t-test rejection of equality with PAM cohort mean at the 10%, 5% and 1% levels.

Table 4 turns to labour market and earnings markers, differentiating between comparisons involving residents (panel A) and all household members, including non-residents (panel B). The distinction is important because earnings data are unavailable for the latter. We see that PAM residents are more likely to be labour force participants, though the difference with the broad cohort is significant at only the 10% level, and insignificant compared to PAM household members. PAM residents are somewhat less likely to be employed than reduced and broad cohort members, which is attributable entirely to lower formal sector (not informal sector) employment. Almost all these differences vanish when non-residents are included in the comparisons. That is, while the employment outcomes of resident PAMs appear somewhat worse than non-PAMs (excluding their own family members), the employment outcomes of the 39% of PAMs who are non-resident are sufficiently good to overcome this and establish statistical equality with comparison groups. The one exception is that PAMs are 6% less likely to be formally employed than individuals in the reduced cohort, a difference which is mirrored (though without statistical significance) in the difference in overall employment rates between these groups.

The proportion of resident PAMs with observed, non-zero earnings is a low 29%, but this is only different to individuals in other groups in the case of reduced cohort members (9 point difference). There are no statistically-significant differences in earned income amongst groups (both including and excluding zero-earners); small positive-earnings sample sizes make this unsurprising (note the very large standard errors).

The simple mean comparisons of table 4 are suggestive in a number of respects. PAM individuals are if anything slightly more likely to be labour force participants than non-PAMs, but their labour outcomes as measured by employment probability look a bit worse amongst residents. Taking non-resident labour into account, however, neutralises most of these differences, with the exception of formal employment probability compared to the reduced cohort. In general, the enormous importance of taking non-residents' contributions to household labour supply into account is clearly in evidence.[14]

[14] This has been pointed out in the context of evaluating state old age pension labour supply effects by Posel, Fairburn and Lund (2004). See Posel and Casale (2003) for a comparatively recent description of internal labour migration in South Africa.

## Table 4: Labour Indicators & Earnings: PAMs & Comparison Groups, 1998 (Means, Std. Deviations in brackets)

| | A. Residents | | | | B. All (including non-residents) | | | |
|---|---|---|---|---|---|---|---|---|
| | 1. PAM | 2. Broad | 3. Reduced | 4. Family | 1. PAM | 2. Broad | 3. Reduced | 4. Family |
| *Participation & employment* | | | | | | | | |
| *Labour force participation rate* | 0.79 | 0.72 + | 0.72 * | 0.75 | 0.82 | 0.78 | 0.79 | 0.8 |
| | -0.41 | -0.45 | -0.45 | -0.43 | -0.39 | -0.41 | -0.41 | -0.4 |
| *Employment rate* | 0.39 | 0.46 + | 0.51 * | 0.44 | 0.49 | 0.52 | 0.55 | 0.52 |
| | -0.49 | -0.5 | -0.5 | -0.5 | -0.5 | -0.5 | -0.5 | -0.5 |
| *Formal employment rate* | 0.22 | 0.30 + | 0.33 * | 0.28 | 0.34 | 0.39 | 0.40 + | 0.38 |
| | -0.42 | -0.46 | -0.47 | -0.45 | -0.48 | -0.49 | -0.49 | -0.49 |
| *Informal employment rate* | 0.16 | 0.16 | 0.18 | 0.16 | 0.15 | 0.14 | 0.15 | 0.14 |
| | -0.37 | -0.37 | -0.38 | -0.36 | -0.36 | -0.34 | -0.35 | -0.35 |
| *Prop. of employed in formal sector* | 0.58 | 0.66 | 0.65 | 0.64 | 0.7 | 0.74 | 0.73 | 0.73 |
| | -0.5 | -0.48 | -0.48 | -0.48 | -0.46 | -0.44 | -0.44 | -0.45 |
| *Sole breadwinner in HH* | 0.12 | 0.1 | 0.12 | 0.1 | 0.1 | 0.06 * | 0.07 + | 0.04 ** |
| | -0.33 | -0.3 | -0.33 | -0.3 | -0.3 | -0.24 | -0.25 | -0.2 |
| *Earnings* | | | | | | | | |
| *Proportion with observed non-zero earnings* | 0.29 | 0.34 | 0.38 * | 0.32 | . | . | . | . |
| | -0.46 | -0.47 | -0.49 | -0.47 | (.) | (.) | (.) | (.) |
| *Total earned monthly income* | 398.73 | 442.26 | 498.12 | 426.69 | . | . | . | . |
| | -1102.88 | -1050.74 | -1094.67 | -1078.17 | (.) | (.) | (.) | (.) |
| *Total earned monthly income (earners only)* | 1507.71 | 1365.78 | 1381.05 | 1400.75 | . | . | . | . |
| | -1726.99 | -1466.64 | -1451.51 | -1569.98 | (.) | (.) | (.) | (.) |
| *Observations* | 124 | 1143 | 663 | 330 | 203 | 1884 | 1115 | 551 |

+, * , ** denote one-sided t-test rejection of equality with PAM cohort mean at the 10%, 5% and 1% levels.
Earnings data unavailable for non-resident household members.
"Sole breadwinner" is 1 if individual is the only employed (A) resident or (B) household member, 0 otherwise.

## 5.4 Regression Analysis

The above mean comparisons can be subjected to more careful scrutiny using regression analysis. We know that age is a strong correlate of labour market outcomes, and since the comparison groups under examination comprise individuals spanning a good proportion (21-45) of typical working life, controlling for this is likely to be important.

We begin with non-parametric, lowess estimation of the correlation between age and labour force participation (LFP) and employment, including a formal/informal employment breakdown. Figures 9 and 10 present results (solid lines for PAMs). Estimated LFP and employment age profiles have the expected, parabolic shapes. The higher average LFP of PAM individuals is revealed to arise mainly amongst relatively older participants (30-40), but differences amongst categories are small and appear unlikely to be statistically significant. The lower employment rate mean for PAMs commented on above holds throughout all but the youngest part of the age range, compared to the strict and broad cohorts, but there is essentially no difference between employment rates between PAMs and their family members.

*Figure 9: Lowess Regressions: Labour Force Participation & Employment Rates on Age (1998)*

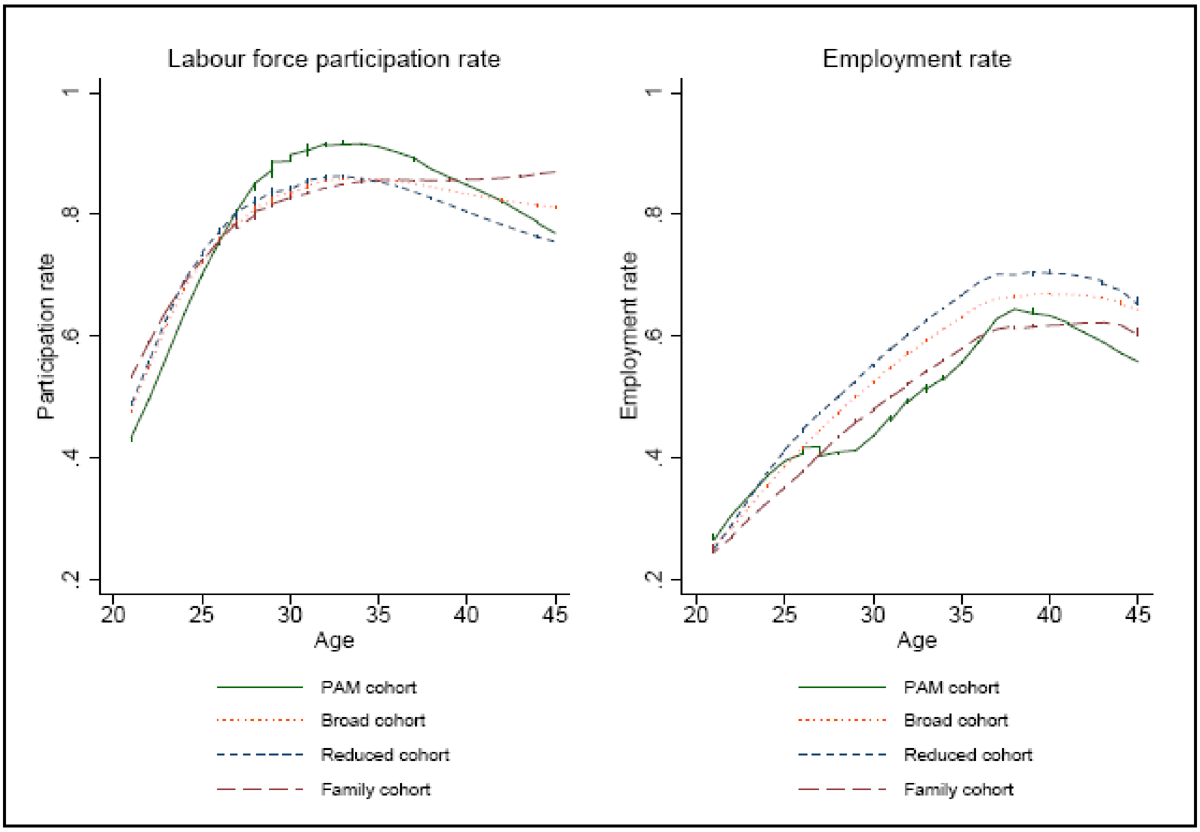
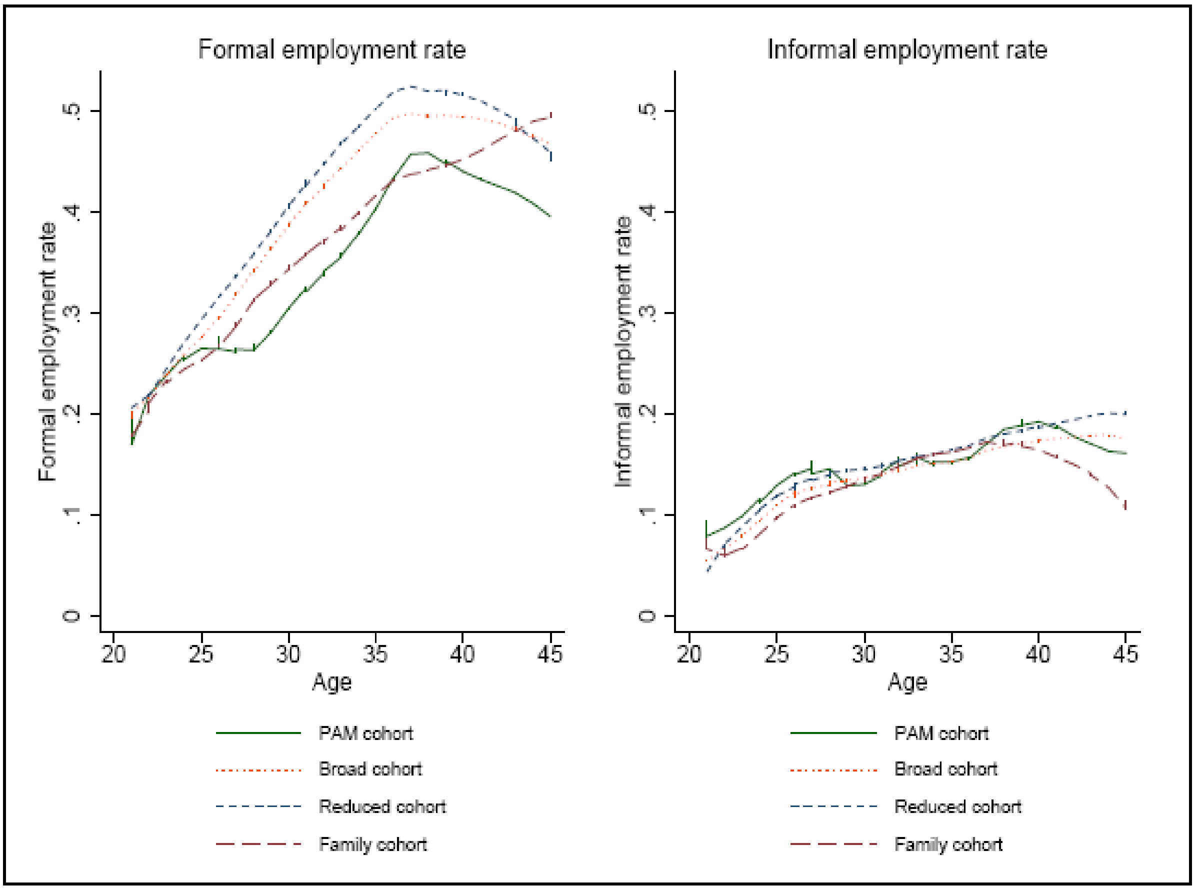


*Figure 10: Lowess Regressions: Formal & Informal Employment Rates on Age (1998)*




Comparison with the LFP picture suggests that employment rate differences are likely to arise partly from the higher probability that PAMs are labour force participants (since the rate of employment will be lower for PAMs than comparison groups for any given absolute proportion working). The obvious interpretation is that PAM individuals are as or more likely to be willing labour market participants than anyone else, but on average face worse employment prospects (though the differences are by no means dramatic). It remains to understand whether these differences are driven by observable (e.g. locational) or unobservable (e.g. behavioural) factors. The fact that PAM family members' employment outcomes also look slightly inferior to those for the more general groups might be interpreted as evidence for the former, since the implication is that the households as a whole containing PAMs fare somewhat worse in the labour market than others.

Figure 10 is rather dramatic general evidence of the over-riding importance of formal sector relative to informal employment. In addition, we see that the differences in employment outcomes stem almost entirely from formal sector employment differences. Again, there is little if any difference between PAMs and their family members.

These regressions examine the two-way relationship between selected labour outcomes and age, allowing for flexibility in the summarising functional form. In doing so, they shed additional light on an important dimension obscured by simple mean comparisons, but many other correlates are of interest, and standard error calculation is cumbersome. Furthermore, the indication is that quadratic age terms should yield good fits for LFP and employment. Consequently, we turn (lastly) to two sets of parametric, nonlinear regressions.

The regressions presented in table 5 ask: controlling for standard, Mincerian and relevant household-level characteristics, does death over the subsequent survey period predict employment in 1998? Three probit regressions on binary employment (1 if employed, 0 otherwise for all labour force participants) are presented, identical except for the comparisons performed by the reduced (1 if PAM individual, 0 if in the reduced cohort), broad (0 if in the broad cohort) and family (0 if in PAM family) group dummies. In general, these regressions appear to do a reasonable job of modeling key correlates of employment probability (Wald test statistics are all statistically significant at the 1% level and pseudo-$R^2$s are above 0.12). While all the comparison group dummies have the negative signs we would expect based on the preceding analysis, all their confidence intervals include zero at any reasonable level. These regressions indicate that, controlling for the traditional, observable characteristics that drive labour outcomes, future death (probably from AIDS) in 1999-2004 has no systematic relationship with the probability of being employed in 1998. This finding is robust across all the (considerable number of) alternative specifications considered (not shown).

*Table 5: PROBIT REGRESSIONS: EMPLOYMENT PROBABILITY*

| | 1. Broad cohort | 2. Reduced cohort | 3. Family cohort |
|---|---|---|---|
| Age | 0.140∗∗ | 0.155∗∗ | 0.065∗∗ |
| | (0.017) | (0.022) | (0.017) |
| Age squared | −0.002∗∗ | −0.002∗∗ | −0.001∗∗ |
| | (0.000) | (0.000) | (0.000) |
| Gender (male=1) | 0.151∗∗ | 0.156∗∗ | 0.112∗ |
| | (0.025) | (0.030) | (0.045) |
| Education (yrs.) | 0.017∗∗ | 0.019∗∗ | 0.015∗ |
| | (0.004) | (0.005) | (0.007) |
| Urban dummy | −0.007 | 0.020 | −0.030 |
| | (0.033) | (0.042) | (0.056) |
| Resident dummy | −0.201∗∗ | −0.204∗∗ | −0.209∗∗ |
| | (0.027) | (0.033) | (0.046) |
| Land access (hct., equiv.) | −0.021 | −0.033 | 0.009 |
| | (0.037) | (0.041) | (0.069) |
| HH OAP receipt dummy | −0.071∗∗ | −0.069∗ | −0.069+ |
| | (0.025) | (0.032) | (0.039) |
| Broad cohort dummy (PAM=1) | −0.036 | | |
| | (0.038) | | |
| Reduced cohort dummy (PAM=1) | | −0.057 | |
| | | (0.041) | |
| Family cohort dummy (PAM=1) | | | −0.011 |
| | | | (0.043) |
| N | 2077 | 1311 | 752 |
| Log Likelihood | -1224.1 | -778.1 | -448.8 |
| Pseudo $R^2$ | 0.13 | 0.13 | 0.12 |
| Wald statistic | 244.8 | 175.2 | 83.8 |

Dependent variable is 1 if individual is employed, 0 otherwise. Equiv. denotes variable in adult-equivalent terms. Coefficients report marginal effects at the mean (or effect of 0 to 1 change in case of binary variables). Std. errors (in parentheses) are heteroskedasticity-robust and corrected for within-cluster correlation. $^+$, *, ** denote significance at 10%, 5% and 1% levels.

Table 6 presents multinomial probit regressions, using the same regressors as the above probits, but breaking employment down into formal and informal sector components. Conditional on being a labour force participant, unemployment (the base category) is compared with formal and informal sector employment probabilities.[15] Results over comparison groups 1 and 2 are shown (1.1 and 1.2, and 2.1 and 2.2, on the table, respectively). While several interesting differences between formal and informal sector employment relative to the unemployed base categories emerge, discussion focuses on the role of the comparison group dummies.

Amongst the broad cohort, future death has no discernable impact on the probability of being in any of the 3 examined categories. Amongst the reduced cohort, future death is associated with an 8.3% reduction in the probability of being formally employed rather than unemployed at the mean (and has no association with informal employment). This effect is statistically significant at the 10% level. Compared to something like being male, which raises this probability by over 18% at the mean (and is significant at the 1% level), this is rather weak. But the result does appear to bear out what we might expect on the basis of mean comparisons and lowess regressions: the labour markers of PAM individuals are basically indistinguishable from those of all comparable individuals, though PAM individuals are slightly less likely to be formally employed than individuals in the comparison group designed to be less directly AIDS-affected on the basis of what we see from 1993 to 2004. Informal employment is quantitatively less important in the sample, but displays no sensitivity to future death, and when employment as a whole is considered, differences are not statistically significant amongst any of the comparison groups.

---

[15] Multinomial logit and probit regressions are typically based on choice models, whereas unemployment, informal and formal sector employment are clearly successively preferable outcome categories for labour participants. Nonetheless, while atheoretical in this case, multinominial analysis is an illuminating way of breaking down differences in employment correlates, provided dependent variable categories are mutually exclusive. A multinomial probit (rather than logit) is chosen since this allows the assumption of the irrelevance of independent alternatives (distinctly problematic in this case) to be relaxed. The obvious alternative – an ordered probit or logit – yields qualitatively identical results (not shown).

*Table 6: MULTINOMIAL PROBIT REGRESSIONS: UNEMPLOYMENT, FORMAL & INFORMAL EMPLOYMENT*

| | 1. Broad cohort | | 2. Reduced cohort | |
|---|---|---|---|---|
| | 1.1 IE | 1.2 FE | 2.1 IE | 2.2 FE |
| Age | 0.024+ | 0.072∗∗ | 0.018 | 0.084∗∗ |
| | -0.013 | -0.021 | -0.017 | -0.026 |
| Age squared | −0.000 | −0.001∗∗ | −0.000 | −0.001∗ |
| | *0* | *0* | *0* | *0* |
| Gender (male=1) | −0.051∗∗ | 0.151∗∗ | −0.085∗∗ | 0.184∗∗ |
| | *-0.018* | -0.027 | *-0.022* | -0.032 |
| Education (yrs.) | −0.002 | 0.020∗∗ | −0.000 | 0.021∗∗ |
| | *-0.003* | -0.005 | *-0.004* | -0.006 |
| Urban dummy | 0.004 | −0.017 | 0.003 | 0.014 |
| | -0.025 | *-0.037* | -0.033 | -0.046 |
| Resident dummy | 0.036+ | *−0.196∗∗* | 0.052∗ | −0.202∗∗ |
| | -0.019 | *-0.029* | -0.023 | -0.036 |
| Land access (hct., equiv.) | −0.036 | 0.038 | −0.051 | 0.043 |
| | *-0.036* | -0.042 | *-0.049* | -0.048 |
| HH OAP receipt dummy | −0.067∗∗ | −0.028 | −0.073∗∗ | −0.025 |
| | *-0.02* | *-0.03* | *-0.025* | *-0.036* |
| Broad cohort dummy (PAM=1) | 0.008 | −0.063 | | |
| | -0.03 | -0.04 | | |
| Reduced cohort dummy (PAM=1) | | | 0.001 | −0.083+ |
| | | | -0.032 | -0.044 |
| N | 1637 | | 1038 | |
| Log Likelihood | -1487.5 | | -945.7 | |
| Wald Test Statistic | 198.7 | | 168.6 | |

Base category: unemployed labour force participant. Comparison categories: IE -employed in informal sector, FE employed in formal sector. Equiv. denotes variable in adult-equivalent terms. Coefficients report marginal effects at the mean (or effect of 0 to 1 change in case of binary variables). Std. errors are heteroskedasticity-robust and corrected for within-cluster correlation. $^{+}$, *, ** denote significance at 10%, 5% and 1% levels.

# 6 An Estimate of Foregone Earnings & Direct Mortality Costs

I now turn to estimates of some direct costs associated with premature adult mortality, focusing first on pre-death care, and burial, costs, and then on earnings.

## 6.1 Pre-death Care and Burial Costs

Information on pre-death care and burial costs comes directly from the 2004 KIDS survey. Respondents are asked "how much in total was spent on health care for x during the course of their illness/injury?", and "how much in total did this household spend in order to bury x and on memorial service?". Data are also gathered on how these expenses were met, based on pre-determined categories. Out of respect for respondents mourning a recent loss, these questions are not posed for deaths occurring in 2004.

Health care costs cover expenses related to medicine, travel, health care providers, hospital/clinic accommodation and "other". I aggregate these and term them pre-death costs. For burials and memorials, I net out any costs recorded as being paid for by burial insurance or stokvels (informal insurance pools), and term these burial costs.

These data should be treated as rough estimates of direct pre-death and death costs only. Pre-death cost data are often incomplete, with amounts frequently recorded in only a sub-set of categories with others missing due to respondents refusing to answer or not knowing. Consequently, the aggregation of these costs should be viewed as the absolute lower bound, comprising just those costs which survey respondents are able to quantify. Furthermore, respondents are recalling costs associated with deaths as long as 5 years prior to the survey for the PAM deaths under consideration, so recall bias poses a substantial risk.

Figure 12, a box plot, summarises the data available for PAM deaths. To contextualise these, costs are compared to the monthly earnings of labour force participants in the sample, and to total monthly household income per adult equivalent (all in 2007 rand). Mean pre-death costs are 2090.60 and burial costs 4795.50 rand, though these mask considerable variation (particularly in burial costs), and medians (perhaps a better measure of central tendency in this case) are substantially lower – 447.45 and 3579.60 rand, respectively. As a proportion of average earnings and total household income per adult equivalent, these costs are very high; this is especially true of burials, but measurement difficulties affect pre-death costs disproportionately as explained above, so these are likely to be underestimates. The combined pre-death and burial cost of a PAM death is 6886.11 and 4027.05 rand at the mean and median respectively; roughly, this is 4 months' average earnings for the employed.

*Figure 12: PAM Pre-death Care and Burial Costs in Context (2007 rand)*

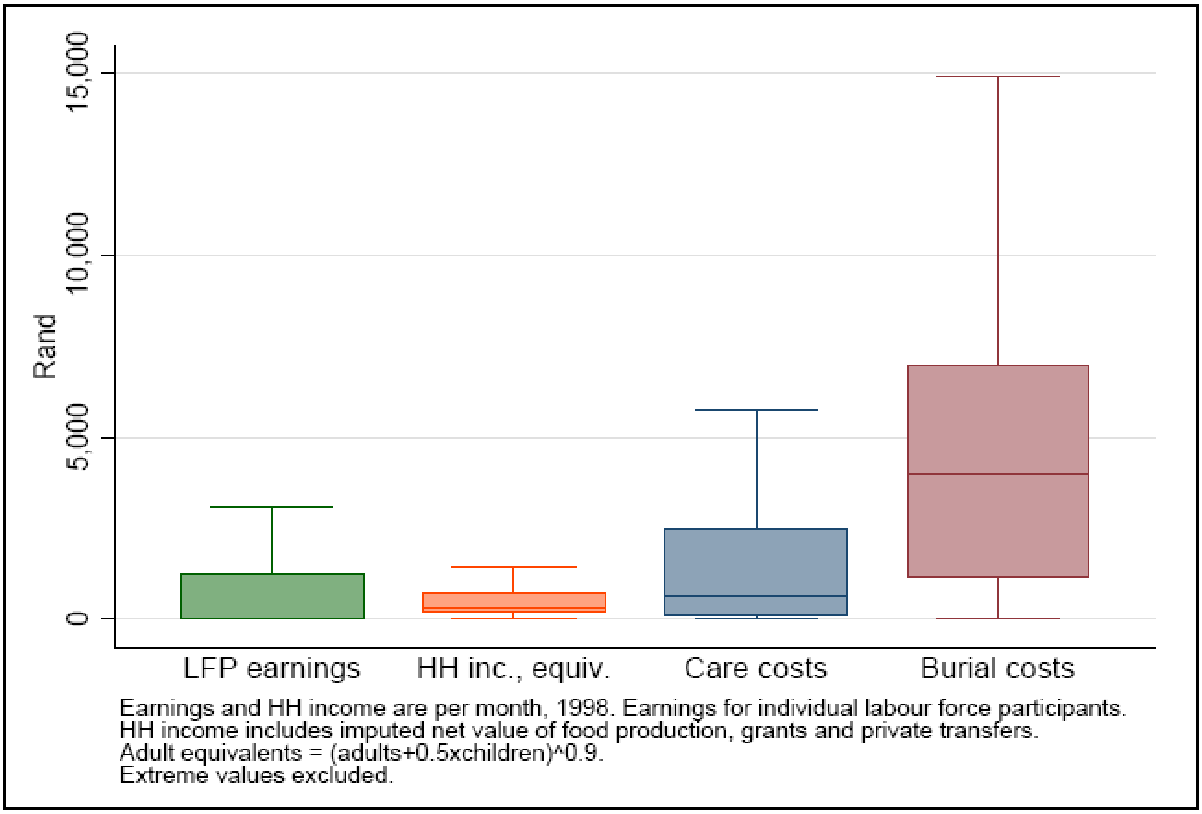
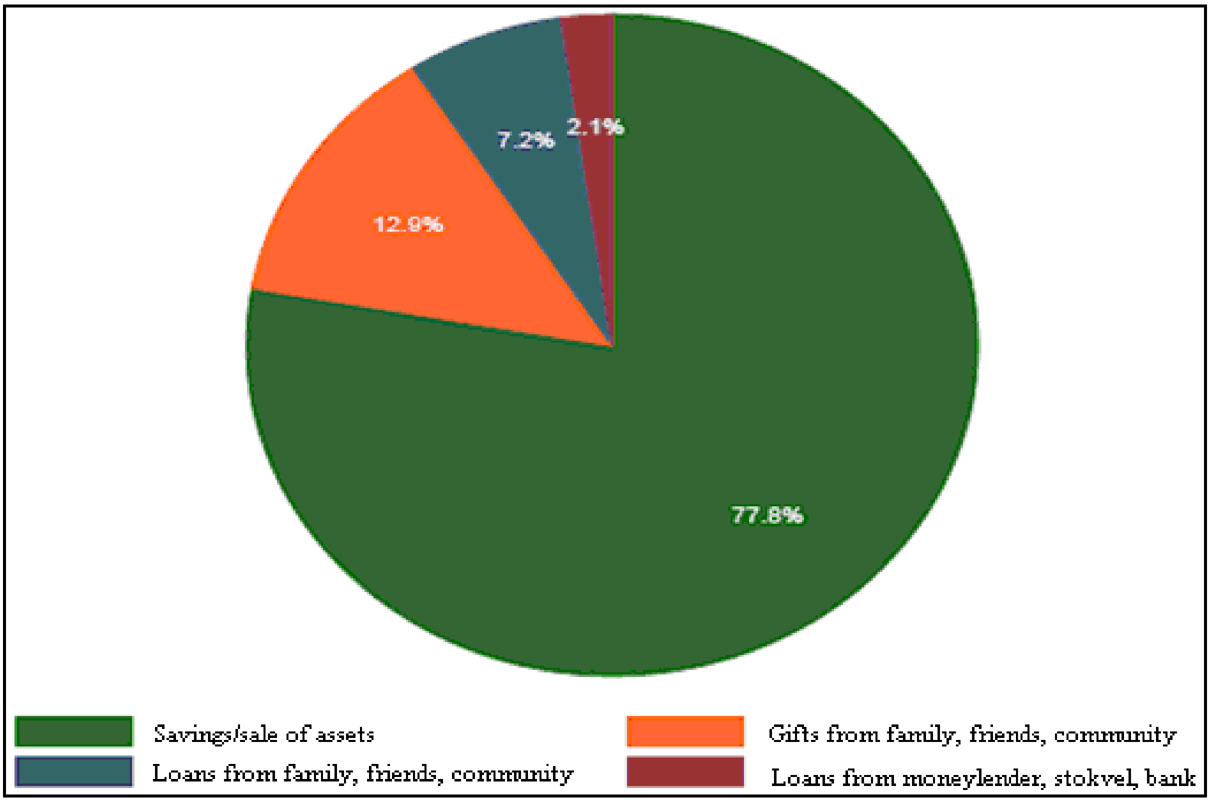



Figure 13, a pie chart, summarises data on how PAM expenses are funded. Again, measurement error is a concern here, and it should be noted that about a quarter of observed mortality costs (24%) are not attributed to any of the available sources. One likely candidate which is excluded by the survey design is funding for ongoing, pre-death costs out of contemporaneous earnings (which, for most households, is likely to mean a reduction in consumption rather than savings). The 13% of costs reported to be covered by gifts is one example of the many economic ripples radiating out from PAM into the networks of directly affected individuals and families.

Figure 13: Sources of PAM Death Cost Funding




## 6.2 Earnings

### 6.2.1 Methods

The question is: had PAM individuals remained well, what *would their earnings have been* around the time they (in fact) died? The present value of this stream of income projected into the future is an estimate of the foregone earnings component of the total cost of their deaths. Answering this question is basically impossible, since we do not have access to an alternative universe in which these same individuals do not fall sick and pass away. Nonetheless, with the caveat that a large number of assumptions and estimates need to be made (all of which are spelt out below), this section attempts an answer.

The last time we observe the earnings (if any) of PAM individuals (who die between 1999 and 2004) is in 1998. Further, a substantial number of PAMs are employed non-residents in 1998, and for all but 3 of these individuals (as for all but a handful of non-residents in the sample) we do not observe earnings. However, we know the year in which individuals die, and of course we observe the earnings of living, resident individuals in 2004.

An obvious starting point is to assume that PAMs' earnings in 1998 would have changed little for the remainder of their working lives. This is used as a baseline for what follows, but because we know that there is a lot of churning in the labour market, and that life-cycle effects are likely to be at work driving employment outcomes and earnings, this is probably a very bad estimate of foregone earnings at the time of death. An alternative is to adopt a matching approach. The (very simple) idea is as follows: find an individual who closely resembles a PAM individual but remains alive (and apparently well) in 2004. Observe this individual's earnings in 2004, compare to 1998, and estimate their earnings at the time of the PAM's death. Relate these earnings to PAMs to establish their estimated foregone earnings.

For this to work, matched individuals' earnings need to be a good indicator of the desired counterfactual. There are two major concerns. First, the matching process cannot capture any systematic, unobserved differences between PAMs and non-PAMs that would have resulted in a difference in their earnings trajectories, even in the absence of morbidity. Because HIV infection is a non-random event, unobserved behavioural or other risk factors which are correlated with labour market outcomes are an important concern. On the other hand, the general similarity of PAMs and non-PAMs through the lens of the labour market in the comparisons conducted above provides some reassurance that PAMs and non-PAMs are similar enough to make a matching-based estimate meaningful. Second, AIDS may have "macro", or general equilibrium, effects on the regional labour market, and hence affect the employment outcomes and earnings of even the directly unaffected. Increased mortality on this scale represents a shock to both labour supply and demand (itself on both the cost [firm employment costs] and price sides [potentially shifting consumption patterns]). The net effect on labour force participants is ambiguous *a priori*, and very unlikely to be homogenous. Conceptually, this means we are actually answering the (odd) question: had PAM individuals remained well, what would their earnings have been around the time they (in fact) died, *given* a contemporaneous regional upsurge in adult mortality; we are assuming away abnormal individual, but not contextual, mortality. The possibility that unobserved differences and macro effects matter is an important and unavoidable qualification of the estimates to follow.[16]

[16] For another example of a matching approach to examine mortality costs (parental death on school enrollment), see Gertler, Levine and Ames (2004).

### 6.2.2 Implementation

The first step is to estimate non-residents' earnings in 1998. I assume that non-employed nonresidents' earnings are zero. For employed non-residents, a standard Mincerian earnings function based on individual demographic characteristics is fitted (not shown). Dummy variables rather than a continuous education term capture possible non-linearities in the effect of educational attainment (primary, grade 10, grade 12 and higher education), and an additional dummy allows for a formal sector earnings premium (which appears large in the data). Mean monthly earnings amongst non-residents (including zero-earners) are estimated at 671 rand for the 79 non-resident PAMs without observed earnings, compared to 625 for all non-PAMs in the broad cohort (the difference is not statistically significant and arises from the greater proportion of non-resident PAMs who are classified as employed; considering only employed non-residents, the equivalent figures are 1181 and 1205 rand).
After applying these estimates, we "observe" positive earnings for 79 PAM individuals (39%), while 124 individuals have zero earnings.[17] It is worth noting that zero earnings does not imply zero contribution to household labour supply, but only to no cash earnings from regular, self or casual employment. Zero earners thus include people involved in subsistence food production, child-rearing and any other non-cash earning domestic production. This labour supply is not quantifiable in the data and its loss as a result of PAM, while possibly significant, cannot be included here.

To perform the matching exercise, it remains to model the probability that someone is in the PAM cohort versus not being in the PAM cohort, based on observables. Potential matches are drawn from the reduced cohort since, while the broad cohort is less restrictive, the latter contains individuals in families experiencing PAMs, raising the risks of spillovers from adult deaths into matched earnings. For example, the earnings of a PAM individual's partner may be reduced as a result of an increased child care burden after the partner's death. The estimation yields a predicted probability of selection into the "treatment" (PAM) and "control" groups (non-PAMs in the reduced cohort), and one-to-one matching with no replacement is performed on this basis, using the Stata PSMATCH2 module (Leuven and Sianesi 2003).[18] Demographic and labour marker means for PAM individuals and their matches are presented in table 7. The match process is successful in generating a sample that closely resembles the PAM sample on average, with no statistically significant differences found

---

[17] For the concerned, attentive reader: it is coincidental that the number of non-zero earner PAMs is equal to the number of non-resident PAMs.

[18] Prediction is based on demographic and labour market characteristics, as well as location (community-level dummies).

(at anything close to the weakest traditional [10%] cut-off). We thus have a match group who "look like" PAMs in terms of their observable, labour-relevant, demographic and locational characteristics.

*Table 7: Demographic & Labour Characteristics: PAM & Match Cohorts, 1998 (Means, Std. Deviations in brackets)*

| | PAM cohort | Matched cohort |
|---|---|---|
| Demographics | | |
| Age | 32.27 | 31.60 |
| | (7.22) | (6.77) |
| Gender (male=1) | 0.40 | 0.44 |
| | (0.49) | (0.50) |
| Education (yrs.) | 6.67 | 6.70 |
| | (3.40) | (3.29) |
| Urban dummy | 0.18 | 0.20 |
| | (0.39) | (0.40) |
| Resident dummy | 0.55 | 0.61 |
| | (0.50) | (0.49) |
| Labour markers | | |
| Labour force participant | 0.85 | 0.82 |
| | (0.36) | (0.39) |
| Employment rate | 0.50 | 0.49 |
| | (0.50) | (0.50) |
| Formal employment rate | 0.38 | 0.34 |
| | (0.49) | (0.48) |
| Informal employment rate | 0.12 | 0.15 |
| | (0.33) | (0.36) |
| Total earned monthly income | 460.69 | 495.41 |
| | (820.71) | (1019.46) |
| Total earned monthly income (earners only) | 1393.13 | 1552.93 |
| | (1177.24) | (1719.53) |
| Observations | 203 | 203 |

T-tests find no statistically significant mean differences.

The next step is to establish hypothetical earnings. This is done as follows:

1. Where the match has positive earnings in both 1998 and 2004, earnings at the time of the matched PAM death are estimated using the average compound growth rate (positive or negative) of earnings between 1998 and 2004. [59 cases]
2. Where the match has zero earnings in both periods, zero earnings at death are assumed. [77 cases]
3. Where the match has zero earnings in 1998 and positive earnings in 2004 (i.e. one or more jobs have been acquired), permanent job acquisition is assumed and the positive earnings amount is assigned. [36 cases]
4. Where the match has positive earnings in 1998 and zero earnings in 2004 (i.e. there has been job loss), permanent job loss is assumed and zero earnings are assigned. [31 cases]

In all cases, earnings at the time of death are converted to 1998-equivalent rand using the national consumer price index for the relevant year.[19] While estimation rules 3 and 4 may appear heroic, it is important to remember that the object of the exercise is to obtain an indicator of the costs of mortality across cases (not a precise estimate for every case); since there are a roughly similar number of individuals involved, adding two straightforward if extreme assumptions influencing predicted earnings in opposite directions was deemed superior to implementing a more arbitrary decision rule (such as randomising year of job acquisition or loss to determine whether to attach positive or zero earnings). Note too that job acquisition will include individuals at the bottom of the age range moving out of full-time education into the workforce, so the assumption of permanently positive earnings thereafter may be the best basic approximation.

### 6.2.3 Results

Figure 11 summarises results, displaying the cumulative density functions of annual earnings foregone at the time of death for PAMs in 1998 rand, as estimated by applying the earnings of matched individuals, and (alternatively) extrapolating 1998 earnings.[20] The two estimates are very similar; if anything, matching yields a slightly more conservative estimate of earnings at the time of death.

[19] This was considered preferable to the provincial consumer price index available for the period from Statistics South Africa, which excluded rural areas.

[20] One outlier (earnings of 99,600 rand in 1998 compared to the next-highest at 57,248), is removed in the latter case.

*Figure 11: Annual Earnings Foregone at Time of Death, PAM cohort (1998 rand)*


Finally, we convert the instantaneous loss of earnings at the time of death to a stock expressing the total value of these losses. This is conceptually identical to the loss of an annuity paying the amount earned, over the expected working life of the individual. Its present value (PV) is given by the standard formula:

$$\mathrm{PV} = \frac{\text{Amount per period} \times \left(1 - \left(\frac{1}{1-i}\right)^{t}\right)}{i}$$

PV estimates are very sensitive to any changes in projected future earnings, the number of periods ($t$), and the rate at which future earnings are discounted ($i$). In principle, we could try to characterise individuals' lifetime earnings profiles and their mortality risk in any given year, to establish precise individual-specific formulations. In practise, given the data at hand, this is unlikely to add much value. Consequently, I simply calculate $t$ as remaining years of ordinary working life on the assumption of retirement at state pension eligibility age (60 for women, 65 for men), and apply a real discount rate of 5% to future earnings. The latter is based on the rounded average difference between the nominal lending rate and CPI, 1998-2004. Most heroically, I assume constant real earnings for the remainder of working life. While we expect earnings to rise with experience in general, employment uncertainty makes building in consistent increases problematic. Further, over the 1998-2004 period, real earnings growth for the broad cohort is zero statistically. On balance, this is likely erring towards the overly-conservative.

On this basis, the present value in 2007 rand of foregone earnings as a result of the premature adult deaths considered averages 117,527 rand when estimated by matching, and 120,281 rand on the basis of extrapolated 1998 earnings. In aggregate, the 203 deaths in the sample carry estimated foregone earnings costs of just under 24 million rand.

Table 8 combines these foregone earnings with the direct, pre-death healthcare and burial costs estimated in section 6.1.

*Table 8: PAM FOREGONE EARNINGS AND DIRECT COSTS (MEANS)*

| | PAM Cohort | Matched Cohort |
|---|---|---|
| Earnings (1998 rand) | | |
| Annual earnings in year of death | 5709.07 | 5275.48 |
| Years to OAP eligibility (retirement) at death | 25.7 | – |
| PV of earnings, earnings as at 1998 | 76,611.87 | – |
| PV of earnings, matched estimate | – | 74,857.74 |
| Expenditure (1998 rand) | | |
| Pre-death health costs | 1313.6 | – |
| Burial costs | 3235.24 | – |
| Total expenditure | 4641.62 | – |
| Combined Cost | | |
| Total (1998 rand) | 81,253.49 | 79,499.36 |
| Total (2007 rand) | 127,567.97 | 124,813.99 |

# 7 Summary of Findings

This paper has explored aspects of premature adult mortality in the KIDS data, focusing on characterising changes in adult mortality, the labour market profile in 1998 of individuals likely to have died of AIDS ("PAM" individuals) from 1999 to 2004, and quantifying the earnings losses, pre-death care, and burial costs associated with their deaths. Mortality amongst working-age adults surges between 1998 and 2004, replicating the pattern in our best view of actual mortality for the black African population of KZN over the period. I find few differences between the labour market markers of PAMs compared to others of the same age group. Seen through the lens of the labour market in 1998, premature adult death in the subsequent 6 years appears quite indiscriminate, though there is fairly weak evidence that the unobservable factors driving future death make unemployment about 8% more likely than formal employment compared to being in a sub-sample designed to be minimally "directly" PAM-affected (the reduced cohort). Conditional on employment, there are no significant differences in earnings.

Estimating earnings lost from PAM is fraught with difficulty, but can be done by extrapolating last-observed earnings, or matching PAM individuals to non-PAMs with similar observable, labour-relevant attributes. These produce broadly similar estimates of foregone earnings, with the latter (preferred) method suggesting a ballpark estimate of nearly 120,000 rand on average. In combination with (shorter-term) pre-death health care and burial costs, this generates a narrow, direct cost estimate of nearly 125,000 rand for the average PAM death. Relative to household income, these figures are high and constitute stark evidence of the enormous economic impact wreaked by the epidemic. In principle, they could serve as inputs to a more comprehensive cost-benefit analysis of possible interventions, feeding into a wider debate on responses to the unfolding tragedy of HIV/AIDS in South Africa.

These results need to be contextualised carefully. The findings that discrepancies in the labour market status of PAMs and non-PAMs are not marked, and that foregone earnings from PAMs are substantial on average, do not mean that individuals suffering premature adult mortality are, in general, in any sense materially well off. Pre-morbidity PAM labour market outcomes are not much worse than anyone else's, *in a labour market* characterised by roughly 45% unemployment, where mean monthly earnings for the employed stood at only 1258 rand (about $230 at the time), and where almost a third of households report public transfers rather than labour earnings as their main income source. The economic vulnerability of adults dying prematurely in the sample is apparent in the match-based predictions of their position at the time of death in the (counterfactual) no-illness case: while 81% would have been labour force participants, only 47% would be taking home any earnings.

# 8 Qualifications

A number of crucial qualifications need to be made (or revisited) concerning this analysis, most of which suggest that results represent an absolute lower bound of the examined economic impacts of premature adult death at the micro and macro levels.

First, there is evidence of non-random attrition, related to mortality, between the 1998 and 2004 surveys. While the age profile of the 1998 sample broadly matches the estimate of the actual age profile of the black population of KZN (suggesting that the 1998 sample is still representative), attrition by 2004 is concentrated amongst individuals of upper working age. This corresponds closely to the under-sampling of adult mortality which we find in KIDS from 1998-2004 (compared to the ASSA model figures), which is surely unlikely to be a coincidence. As explained in section 4.3 (and tables 1 and 2), there is also evidence that relatively better-off individuals are more likely to attrite. Taken together, these observations suggest strongly that when individuals with relatively good labour market outcomes die, their families are more likely to disappear from the data (leaving their deaths unrecorded). This is very plausible: death of a primary breadwinner is an economic catastrophe that raises the likelihood of household dissolution (or migration) and attrition, wherever the household lies on the income spectrum. This will induce downward bias in estimates of the economic profile of people suffering future death (PAMs). This possibility qualifies the external validity of the finding that PAMs in the sample may be slightly worse off in the labour market, and the estimates of the average PV of foregone earnings from PAM. The weak evidence of lower employment outcomes for future PAMs, in other words, could well be an artefact of the disappearance of PAMs with relatively strong labour market profiles from the data. Future work should attempt to account for this possibility explicitly (though despite observable mean differences between attriters and non-attriters, modeling attrition probability is not straightforward). *Any* work relating to adult mortality in these data needs to recognise the risk to inference that arises from the relationship between the undersampling of mortality probably arising from attrition, and outcomes of interest.

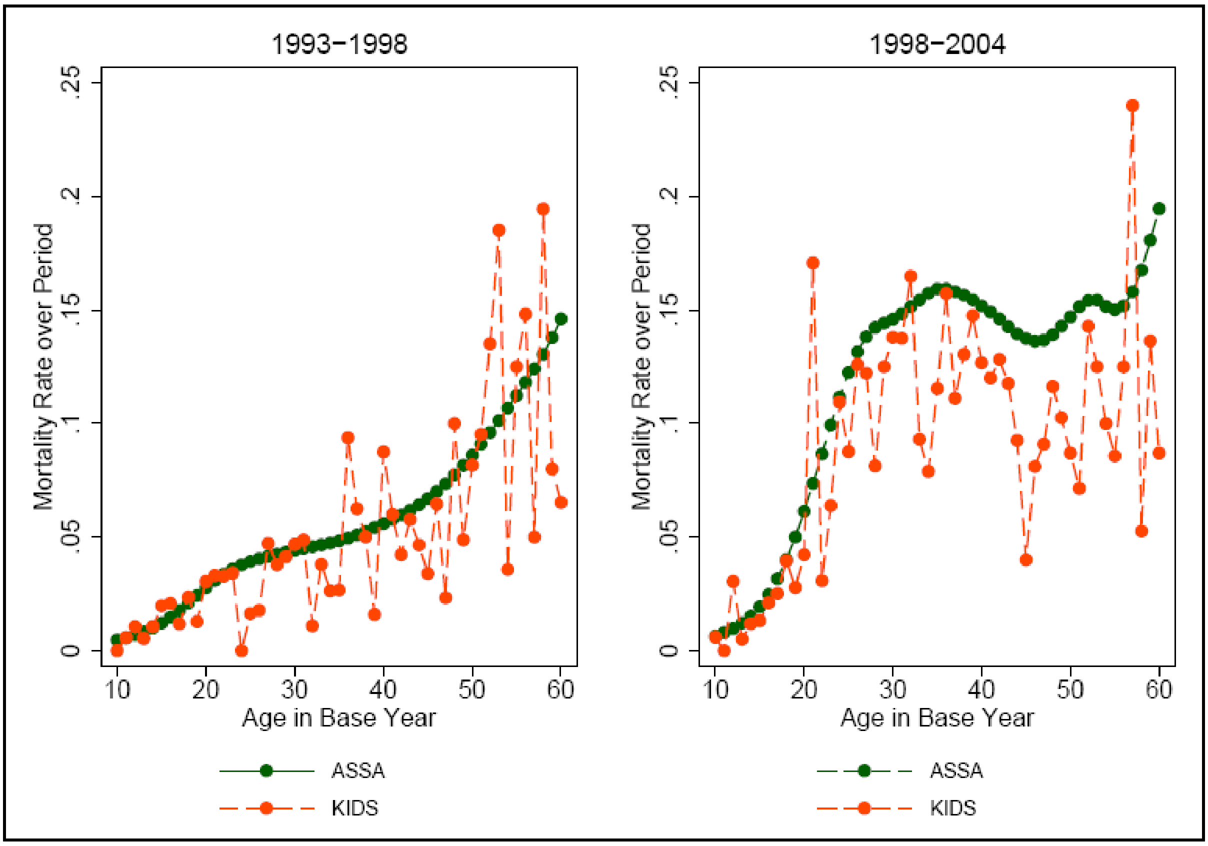
With regard to loss estimates, the limitations of the matching approach – unobservables and spillover effects – were discussed in section 6.2.1, and present value calculations are notoriously sensitive to variable inputs.

While obvious, it is also important to recognise that although foregone earnings and direct death costs are important components of the burden of the epidemic, they represent just a sliver of the total cost. A holistic measure might add (to name just a few of the most obvious): the intrinsic value of life; the emotional suffering of bereaved relatives; the foregone earnings and earnings trajectory differences of survivors generated by *ex ante* and *ex post* care burdens; costs imposed by the broader social network (e.g. supporting mortality-related costs of other families); human capital accumulation effects (especially on children); general equilibrium effects (various kinds of externalities and social network disruptions).

Finally, the distinction between general prime age adult mortality and AIDS death needs to be borne in mind. KIDS makes adult mortality visible more than is usually the case in widely-representative, socio-economic household survey data, in a time and place at the centre of South Africa's AIDS epidemic. But, strictly speaking, we observe only adult deaths due to illness, with an estimated 79% probability of being AIDS-related, for the sample examined. The homogenous treatment of non-AIDS and AIDS deaths adds noise and may certainly impact the interpretation of results as indicative of the impact of (specifically) AIDS-related mortality; addressing this awaits fresh household survey evidence with a more explicit focus on the epidemic.

# A Raw Mortality Data by Age

*Figure 14: Unsmoothed Aggregate Mortality Rates: ASSA & KIDS, 1993-1998 & 1999-2004*




# B Methodological Appendix

## B.1 Definitions

Labour force participants comprise all individuals unless:

- Aged less than 15 or older than 59 (women) or 64 (men);
- Not working and not looking for work due to illness, disability, listed as housewife or child-rearing, in formal education, retired or "other". I.e. this is "broad" labour force participation, including the discouraged (non-actively searching) unemployed.

The employed comprise all labour force participants who:

- Earn wages from regular or casual work of any kind, or earn income from self-employment (that is, anyone with labour earnings);
- For non-residents (for whom earnings are unavailable), list any form of employment as their primary activity.

Formal *versus* informal employment:

- Employed individuals are assumed to be employed in the informal sector, unless (in which case they are assigned to formal employment) they:
  - receive wages from regular employment;
  - are self-employed and business type is listed as "professional".

## B.2 Key Assumptions

Earnings:

- Residents who are not employed and have no recorded earnings are assumed to have zero earnings;
- Non-residents are assumed to have zero earnings if unemployed.

## B.3 Key Imputations

- A handful of missing net wages are imputed from available gross wages, using a 30% deduction rate;
- Net food production value is taken from KIDS imputation programmes (see KIDS metadata);
- A handful of missing non-labour incomes are imputed using median values based on grant category (not likely to be problematic since variation of grant amounts is low);
- Self-employment income is available at only the household level, though respondents list up to three people who are involved in the business. Following Cichello et al. (2005), I assign net profits to labour earnings (few businesses are capital-intensive). Unlike Cichello et al, I divide net business earnings evenly amongst participating individuals. Since these numbers are based on income and expenses in the previous month, their reliability as indicative of long-term net income flows from self-employment are open to doubt. A handful of implausible, negative values in each survey are removed by applying the relevant, mean profit margin (profit/sales), for the sample excluding these businesses, to these business' sales.

# References

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon, "Adult Mortality and Consumption Growth in the Age of HIV/AIDS," World Bank Policy Research Working Paper 4082 December 2006.

Bhorat, Haroon, "Labour Market Challenges in the Post-Apartheid South Africa," *South African Journal of Economics*, December 2004, *72* (5), 940–977.

___ , Murray Leibbrandt, and Ingrid Woolard, "Household Inequality and the Labour Market in South Africa," *Contemporary Economic Policy*, 2001, *19* (1), 73–86.

Carter, Michael and Julian May, "One Kind of Freedom: Poverty Dynamics in Post-Apartheid South Africa," *World Development*, 2001, *29* (12), 1987–2002.

___ , ___ , Jorge Aguëro, and Sonya Ravindranath, "The Economic Impact of Premature
Adult Mortality: Panel Data Evidence from KwaZulu-Natal, South Africa," *Submitted t*o *Special Edition of AIDS, 20 April 2007*, 2007.

Cichello, Paul, Gary Fields, and Murray Leibbrandt, "Earnings and Employment Dynamics for Africans in Post-apartheid South Africa: a Panel Study of KwaZulu-Natal," *Journal of African Economies*, 2005, *14*, 143–190.

Department of Health, "National Hiv and Syphilis Prevalence Survey -South Africa 2005," Survey Findings, South African National Department of Health 2006.

Dorrington, Rob, Leigh Johnson, Debbie Bradshaw, and Timothy-John Daniel, "The Demographic Impact of HIV/AIDS in South Africa: National and Provincial Indicators for 2006," Centre for Actuarial Research, South African Medical Research Council and Actuarial Society of South Africa 2006. Cape Town.

Fox, Matthew, Sydney Rosen, William MacLeod, Monique Wasunna, Margaret Bii, Ginamarie Foglia, and Jonathon Simon, "The impact of HIV/AIDS on labour productivity in Kenya," *Tropical Medicine and International Health*, March 2004, *9* (3), 318–324.

Gertler, Paul, David Levine, and Minnie Ames, "Schooling And Parental Death," *Review of Economics and Statistics*, February 2004, *86* (1), 211–225.

Keswell, Malcolm, "Non-Linear Earnings Dynamics in Post-Apartheid South Africa," *South African Journal of Economics*, December 2004, *72* (5), 913–939.

Leuven, E. and B Sianesi, "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.," http://ideas.repec.org/c/boc/bocode/s432001.html. 2003. Version 3.0.0.

May, Julian, Jorge Aguëro, and Ian Carter Michael Timæus, "The KwaZulu-Natal Income Dynamics Study (KIDS) 3rd Wave: Methods, first findings and an agenda for future research," School of Development Studies Working Paper No. 43. School of Development Studies, University of KwaZulu-Natal 2006.

___ , ___ , Michael Carter, and Ian Timæus, "The KwaZulu-Natal Income Dynamics Study (KIDS) 3rd wave: methods, first findings and an agenda for future research," 2006. University of Wisconsin, unpublished manuscript.

___ , Michael Carter, Lawrence Haddad, and John Maluccio, "KwaZulu-Natal Income Dynamics Study (KIDS) 1993-1998: a longitudinal household database for South African policy analysis," *Development Southern Africa*, October 2000, *17* (4), 567–581.

Nattrass, Nicoli, *The Moral Economy of AIDS in South Africa*, Cambridge: Cambridge Africa Collection, November 2003.

Pauw, Kalie, "A Profile of KwaZulu-Natal: Demographics, Poverty, Inequality and Unemployment," Technical Report, Provincial Decision-Making Enabling (PROVIDE) Project 2005.

Posel, Dorrit and Daniela Casale, "What has been Happening to Internal Labour Migration in South Africa, 1993-1999?," *South African Journal of Economics*, September 2003, *71* (3), 455–479.

___ , James A. Fairburn, and Frances Lund, "Labour Migration and Households: A Reconsideration of the Effects of the Social Pension on Labour Supply in South Africa," 2004. Presented at Ninth Annual Conference on Econometric Modelling for Africa, June-July 2004.

Rosen, Sydney, Frank Feeley, Patrick Connelly, and Jonathon Simon, "The private sector and HIV/AIDS in Africa: taking stock of 6 years of applied research," *AIDS*, 2007, *21* (supp. 3), S41–S51.

Statistics South Africa, “Report No. 00-91-05 Provincial Profile 2004: KwaZulu-Natal,” Technical Report, Statistics South Africa 2004.

Yamauchi, Futoshi, Thabani Buthelezi, and Myriam Velia, “Gender, Labor, and Prime-Age Adult Mortality: Evidence from South Africa,” *International Food Policy Research Institute, FCND Discussion Paper 208*, 2006.