Poverty and Labour Market Markers of HIV+ Households: An Exploratory Methodological Analysis

Haroon Bhorat and Najma Shaikh[1]

Draft 1: Please do not quote or circulate without permission of authors

---

[1] Director, Development Policy Research Unit (DPRU), School of Economics, University of Cape Town and Epidemiology Consultant, Provincial Administration of the Western Cape (P.A.W.C) respectively.

Introduction 3
The Methodological Approach 3
Aim 3
Data Sources 3
The Antenatal Clinic Survey, 1998 .... *4*
The October Household Survey,1999 .... *4*
The Matching Process 4
The Imputation Exercise 6
Key ANC98 Data Results 7
Results from the Imputation Exercise 10
Socio-Economic Correlates of Imputed HIV-Status 13
Household Characteristics by Imputed HIV Status 13
Measures of Absolute and Relative Poverty Differences *15*
Household Size and Dependency Ratios *16*
Attachment to Wage Earners and the Quality of Support *18*
Attachment to the Unemployed and Non-Wage Earners *23*
Individual Characteristics by Imputed HIV Status 26
Labour Market Characteristics by Imputed HIV Status *26*
Marital Status and Relationship to Head of Household *31*
Conclusion 32
References 35
Appendix 36

Introduction

While HIV epidemic continues to mature in South Africa, showing no signs of plateauing, it is clear that South Africa(SA) stands at the brink of an AIDS crisis. Some of the salient features of the AIDS pandemic are understood in terms of the demographic profile namely; HIV/AIDS tends to strike young, heterosexual, sexually active adults (DOH, 1998; DOH, 2000). However, the economic implications of the pandemic remain murky at best, with a limited understanding on how HIV related mortality and morbidity trends affect productivity, factor returns, employment, income distribution, savings rates, consumption patterns and other economic variables.

In addition, there is very little empirically robust information on the social and welfare correlates of the pandemic. Understanding these economic implications is clearly critical in the post-*apartheid* period. The core asset of a country is its skilled human resources and it is the very nub of this component that is being threatened by the pandemic. Whilst there has been debate about the impact of HIV/AIDS on poverty, the link between HIV/AIDS prevalence and the various determinants of poverty and labour market outcomes has not been rigorously dealt with in the South African literature. One of the key obstacles to understanding of these issues is the absence of hard data on HIV status and its association with poverty, inequality and labour markets.

The purpose of this proposed research initiative is therefore to develop a methodology to analyse the features of the HIV/AIDS epidemic at the household level through the use of existing household survey data. This would allow us to get a snapshot of the features of the HIV/AIDS epidemic at the household level. This combined with HIV antenatal data would provide a tentative window into the households that a sub-set of HIV/AIDS sufferers and their dependents emanate from.

The Methodological Approach

Aim

The research initiative is designed to develop a method of analyzing the household characteristics associated with HIV status using currently available data.

Data Sources

Currently, there is a dearth of robust population based, household level data linking HIV status to socio-economic and demographic factors. Household surveys generally, are known to be expensive and labour intensive and are particularly subject to bias with respect to ascertaining HIV status, given factors associated with stigma and confidentiality.

The October Household Survey of 1999 (OHS99), a nationally representative household survey produced by Statistics South Africa (SSA), in conjunction with the published statistics from the Antenatal Clinic (ANC) data of the National Department of Health were the two data sets used for this study. The HIV antenatal survey provides the primary source of information for measuring the epidemic in the heterosexual, sexually active, adult population (Shaikh & Abdullah, 2002).

The Antenatal Clinic Survey, 1998

The annual national HIV antenatal sentinel surveys which are conducted by the National Department of Health remains the primary source of data in SA reflecting the trends and magnitude of HIV infection amongst young, sexually active, heterosexual adults. The data derived from the ANC surveys form the primary source data for modelling the epidemic in terms of growth and its socio-demographic impact as well as for planning, monitoring and evaluating HIV strategies and programmes (DOH, 2000; WHO, 2000).

Each year, during the month of October, women attending public sector clinics are anonymously examined for HIV. This group is examined because they are considered to be representative of the sexually active, heterosexual members of the general population and are accessible to blood testing as this forms an integral part of the routine antenatal clinical care at the first visit. The primary aim of the survey is to estimate HIV sero-prevalence at provincial and National level. A two-stage random cluster sample using the Probability Proportional to size (PPS) method is used (DOH, 1999). This allows for a self-weighted sample, where each attender has an equal chance of being selected for survey. A total of 15089 specimens were collected country wide. The same sites (sentinel sites) are chosen each year for the survey as this allows one to track the epidemic over time.
After the routine pregnancy education session and following verbal consent, blood specimens are collected, transported and processed according to standard operational procedures outlined in the national protocol and according to the WHO guidelines for unlinked sentinel surveys (WHO, 2000; Nicoll *et al,* 2000).

The October Household Survey, 1999

For the OHS99, and for all the previous national household surveys, a sample of 30 000 households was drawn in 3 000 Enumerator Areas (EAs) (that is 10 households per enumerator area). A two-stage sampling procedure was applied and the sample was stratified, clustered and selected to meet the requirements of probability sampling. The sample was based on the 1996 Population Census enumerator areas and the estimated number of households from the 1996 Population Census. Within each explicit stratum the EAs were stratified by simply arranging them in geographical order by District Council, Magisterial District and, within the magisterial district, by average household income (for formal urban areas and hostels) or EA. The allocated number of EAs was systematically selected with probability proportional to size in each stratum. The measure of size was the estimated number of households in each EA. A systematic sample of 10 households was drawn.

The questionnaire has seven sections, which cover a range of issues ranging from standard person and household characteristics, details of births and deaths within the household, household income and expenditure, and fairly detailed labour market information. It is these particular questions that we explore through the rest of this paper.

The Matching Process

The broad approach in this exercise involved linking the data from the ANC survey of 1998 to the OHS99 data set, by randomly allocating[2] proportions to a sub-sample of women from

[2] Given that the random process is subject to instability, 35 replications of the random assignment according to the ANC rates were done and the mean of the variable under consideration was reported .

the OHS99 who used public sector clinics and were pregnant in the previous year. Through this sample of pregnant women in the OHS99, we attempted to determine the labour market and poverty correlates of HIV/AIDS sufferers and the households they reside in. We turn now to a more detailed discussion of this 'matching process' that will hopefully elucidate on the procedure used to derived imputed HIV status within the household survey data set.

Since one cannot assume that all pregnant women derived in the OHS99 sample were HIV+, the ANC98 published results were critical for providing these prevalence figures. The distribution of HIV+ pregnant women were provided by province, age and education levels. Hence, in theory, one is able to deriver from the ANC98 data a three-dimensional matrix containing the proportion of HIV+ women as a share of the total women in the sample, for that particular combination of categories. Hence, we can define the matrix for example as ρ_{ijk} where ρ is the share of women in the ANC98 sample who are reported as HIV+ according to province *i*, age *j* and educational category *k*. Assume then that we define four age categories: 15-24; 25-34; 35-44 and 45-49, and four education categories: primary schooling or less; incomplete secondary education; matric and tertiary education.

The ANC98 data was aligned in a manner where each of the individuals in the sample was coded according to the above procedure. Hence, we arrived at values of ρ_{ijk} for each of the women in the sample. Hence, with the ANC98 data, derived estimates of ρ_{ijk} that would for example for province 1 ($i = 1$) yield estimates of the share of women who are HIV+ according to a 4x4 matrix for all combination of the categories *j* and *k*.

The values for ρ_{ijk} were used to impute the HIV status of pregnant women in the OHS99 dataset, given that we had the values for ρ_{ijk} for each possible combination of province, age and education categories from the ANC98 data. In terms of the OHS99 data then, we used the sample of women who reported giving birth in the last year, and attended a public clinic, and re-categorised the age and education variables, to reflect, as with the ANC98 data two educational categories and three age cohorts. We present here, a fictitious example of what the OHS99 data, combined with the ρ_{ijk} values from the ANC98 data, may look like in the data set. Table 1 thus illustrates that from the OHS99 sample of all women who reported giving birth in the previous year and attending a public health clinic, they would be sorted according to the relevant province, age and education categories.

Table 1: Dummy Table of Matching Process between OHS99 and ANC98

Pcode	**Province (*i*)**	**Age (*j*)**	**Education (*k*)**	**ρ_{ijk} (from ANC98)**
1	1	1	1	0.56
2	3	2	4	0.22
3	5	3	3	0.33
4	1	1	1	0.56

The term 'pcode' represents the unique identifier for each woman in our sample. She was categorised into her provincial, age and educational categories. The ρ_{ijk} values are those we take from the ANC98 dataset and then *imputed* into the OHS99 sample. If the ANC98 data suggested that of those women in province 1, age category 1 and education cohort 1, 56% were HIV+, this is the proportion we impute into OHS99, should an individual fit into that category. The procedure from this point was as follows:

A binary variable was created, 0 for HIV- and 1 for HIV+ on the basis of the ρ_{ijk} values. Hence, taking the $\rho_{ijk} = 0.56$ as an example, we randomly assigned an HIV+ status to 56% of all the women in the sample who were in the first cell of the matrix, where $(i,j,k)=(1,1,1)$. So, if we had 10 women in this cell, then approximately 6 were randomly assigned an HIV+ status and the 4 an HIV- status. In terms of Table 1 shown above, we created an additional binary variable (0 and 1 values only). The binary variable represents imputed HIV status.

The Imputation Exercise

As indicated above, the imputation exercise involved at the outset, the alignment of the OHS99 dataset so that it conformed with the covariates contained in the Ante-Natal Clinic data for 1998(ANC98). The ANC98 survey data is composed of the sample of women, aged 15 to 49, who were pregnant in 1998 and attended a public sector clinic during their current pregnancy. The ANC survey runs in October, the same month as the OHS, and hence seasonal variation between the two data sets was not a problem. The OHS99 has data on women in this age range, that gave birth in the last year, hence, would have had a high likelihood of including women who were pregnant in 1998[3]. This sample of 1998 pregnant women from the OHS99, would therefore provide the first initial match to the ANC98 data.

The OHS99 data contain information on whether individuals in the household were private medical aid members or not. This would allow us to derive the sample of women who had been pregnant in 1998, aged 15 and 49, and were not on a private medical aid.

Table 2 below therefore provides an indication of how this sample of pregnant women was derived within the OHS99 data set. Hence, we began with a total (unweighted) sample of about 107 000 individuals. Of these, about 90 000 reported attending a public clinic in the last month. In terms of live births in the last 12 months, the survey yielded an estimate of 1954 women, who reported giving birth over this period.

Table 2: Sampling Procedure in OHS99 Data Set

Category	Number of Individuals	Share of Total
Total Sample of Individuals	106650	100
Did not attend public clinic	16738	15.69
Attended Public Clinic	89912	84.31
Live Birth in Last 12 months	1954	1.83
Pregnant & Attended Public Health Clinic	1659	1.55

[3] The match though is an inexact one: from the ANC data in October 1998, it implies that a woman would have become pregnant in the period February 1998 to September 1998, and hence would have given birth in the period November 1998 to June 1999. While there is an overlap, the period of birth from about July 1999 to September 1999 is not captured in the ANC data set for 1998.

We then had to ensure that these women, many of whom would have been pregnant in 1998, were also public health clinic attendees. In addition, this sample needed to be cleaned of any obvious anomalies[4]. One of the key additions required here was to ensure that the age profile of these women matched with those in the ANC data. Hence, we retained all pregnant women between the ages of 12 and 49 who reported attending a public sector clinic in the last year. Ultimately then, the final sample to be utilized for our matching process was 1659 pregnant women between the ages of 12 and 49 who reported attending a public sector clinic.

This remains, as is obvious, an extremely small sample. This is a function primarily of the initial sample size of the OHS, where 30 000 household and just over 100 000 individuals were surveyed. Given this, the results that we will report throughout this paper need to be treated with caution. On the one hand, the results should be viewed strictly as indicative only. On the other hand, the strength of the paper lies in the fact that we propose here a fairly rigorous methodological approach for accessing socio-economic information on the pandemic, in the absence of direct survey data probing these issues. It is hoped therefore that this paper is seen in this light, that while we report on the labour market and poverty markers of HIV/AIDS, we accept the limitations inherent in the reported results due to the small sample size[5].

Prior to presenting the specific results that were obtained from the imputation exercise, it may be useful to firstly provide an overview of some of the key results that emanate from the ANC survey data.

Key ANC98 Data Results

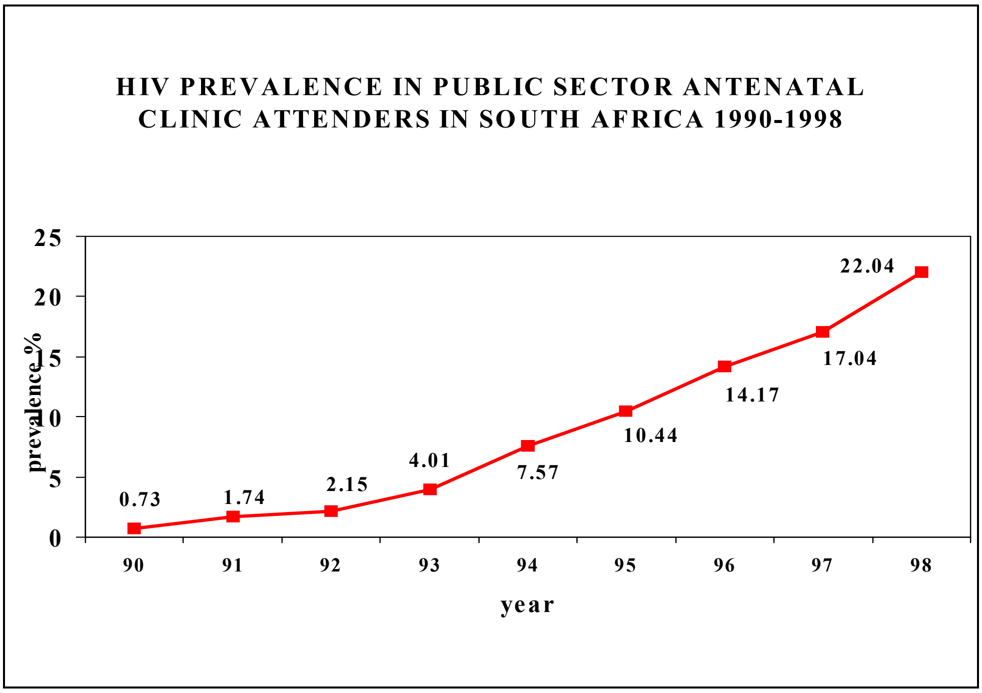
Before proceeding to a more detailed analysis of some of the key results emanating from the ANC98 data, it is useful to provide at least an initial snapshot of the trend in the epidemic since 1990, as measured by the ANC data. Given, as noted above, that this data is the only 'hard' information on HIV trends in South Africa, the estimates obtained take on an added significance. Simply put, all other projections and population-wide estimates are essentially modeled around these prevalence figures from the National HIV Ante-natal Survey results.

Figure 1 below then, conforms the general picture of an epidemic that has been steadily rising at the national level. From a base level of 0.73% in 1990, the epidemic has spread at a tremendous rate over the 1990s.

[4] For example, the sample of women who reported bearing a child in the last year, contained at least 10 women over the age of 60 and one man!

[5] One obvious way around this problem would have been to use Census data. However, using the Census 1996 data is problematic given that it is accepted that the ANC for 1995 is not reliable. In turn, the Census 2001 results were not available at the time of writing.

Figure 1. Trends of HIV Prevalence for Public Sector Antenatal Clinic Attenders. 1990-1998.



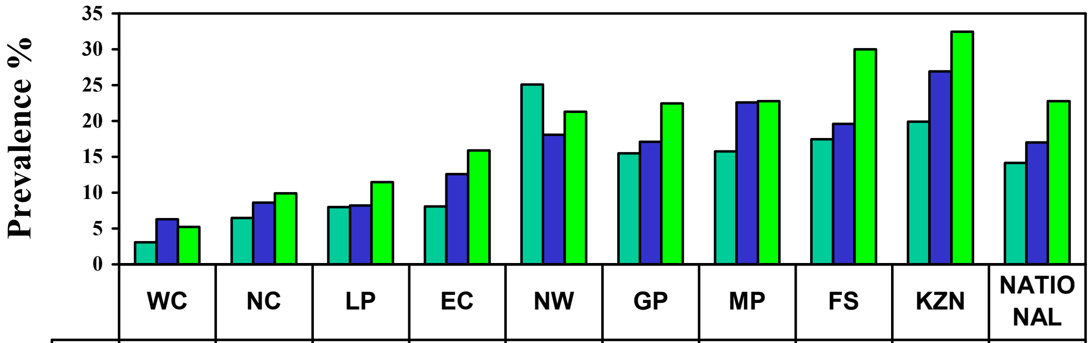
Trends show that for the period 1994-1998, there was a 14.5 percentage point rise. The rapid increase during this period is indicative of the early phase (exponential phase) of the HIV epidemic whereby the doubling period is approximately two years. Unlike the mature phase of the epidemic, whereby the number of new cases balance out the number of deaths, the early phase of the epidemic is characterized by a larger number of individuals recently infected and fewer deaths.

Analysing the data by provinces, we see that there is wide variation between the nine provinces with Western Cape on the lower end of the scale at 5% and Kwa- Zulu Natal the highest (33%)(Table 3).

Table 3: Estimated HIV Prevalence, by Province (ANC98)

Province	Estimate	Std. Err.	95% Confidence Interval	
Eastern Cape	0.16	0.009	0.1418	0.1784
Free State	0.23	0.012	0.2045	0.2516
Gauteng	0.23	0.009	0.2124	0.2478
Kwazulu-Natal	0.33	0.008	0.3096	0.3407
Mpumulanga	0.30	0.013	0.2743	0.3264
Northern Cape	0.10	0.013	0.0759	0.1255
Northern Province	0.12	0.008	0.1016	0.1334
North West	0.21	0.011	0.1920	0.2339
Western Cape	0.05	0.005	0.0420	0.0623

Figure 2. HIV Prevalence by Provinces
South Africa 1996-1998



	WC	NC	LP	EC	NW	GP	MP	FS	KZN	NATIONAL
1996	3.1	6.5	8	8.1	25.1	15.5	15.8	17.5	19.9	14.17
1997	6.3	8.6	8.2	12.6	18.1	17.1	22.6	19.6	26.9	17
1998	5.2	9.9	11.5	15.9	21.3	22.5	22.8	30	32.5	22.8

Three-year trends by provinces show different patterns of growth of the HIV epidemic in each of the nine provinces. The data shows that whilst KZN had the highest rate, whilst the highest percentage point increase was observed in the Free State province. The high rates in KZN and Free State province essentially mean that approximately one in every three pregnant woman tested at the public sector clinics was HIV positive.

Examining age specific HIV prevalence we find that the highest rates are in the 20 –29 age group. This prevalence was significantly higher than all the other age cohorts. Trends of the HIV/AIDS epidemic in Africa have shown that during the exponential phase of the epidemic there is a rapid rise in the age cohort at highest risk of being infected, followed by saturation and a shift to younger age groups (Stoneburger et al 1996). The age-specific trends confirm that the HIV epidemic is at the exponential phase, whereby the highest rates are observed in the 20-29 age groups (Table 4). Hence, as the epidemic matures, AIDS related morbidity and mortality will be most visible in the 20-49 age group. In terms of labour market and poverty concerns, what is most significant is that HIV affects young adults, who are in their prime with respect to productivity and fertility. The loss of a large proportion of this age cohort could potentially increase the society's existing skills shortage. There would also be a larger pool of orphans due to AIDS, which would negatively impact on households of extended families to provide support, thus placing the state welfare system under further strain in terms providing child support grants.

Table 4: Estimated HIV Prevalence, by Age Category (ANC98)

Age Category	Estimate	Std. Err.	95% Confidence Interval	
Unspecified	0.11	0.031	0.0476	0.1680
12-19	0.19	0.007	0.1759	0.2043
20-29	0.24	0.005	0.2351	0.2537
30-49	0.16	0.006	0.1437	0.1666

Table 5: Estimated HIV Prevalence, by Education Category (ANC98)

Education Category	Estimate	Std. Err.	95% Confidence Interval	
Primary_	0.19	0.005	0.1839	0.2029
Incomplete	0.22	0.005	0.2067	0.2266
Matric_+	0.25	0.010	0.2263	0.2650

The HIV prevalence by education showed that women with primary school or less had significantly lower HIV prevalence than those with higher levels of education. The fact that higher rates were observed in women who attained some high school education or more highlights that the epidemic will impact on the relatively skilled and hence more employable groups, hence possibly eroding the skills base.

Table 6: Significance Tests on HIV Status, by Age and Education (ANC98)

HIV Status	Mean	Std. Err.	95% Confidence Interval	
Age				
Negative	25.64	0.060	25.5196	25.7563
Positive	24.89	0.099	24.6929	25.0792
***t*-test**	5.90			
Years of Education (not controlled for age)				
Negative	6.97	0.032	6.9099	7.0373
Positive	7.15	0.061	7.0313	7.2717
***t*-test**	-2.50			

The mean age differences by HIV status showed HIV positive women were significantly younger than the negative cohort (Table 6). However, as far as the mean number of years of education attained was concerned, the positive women had marginally attained higher levels of education although these were not significant.

Results from the Imputation Exercise

Having identified the working sample, the procedure as outlined above, was to segment the sample of women in the OHS according to pre-defined education and age categories. While the hypothetical example above referred to 4 education and age categories for each province, it was clear from the small sample size that we derived that such a 4*x*4*x*9 matrix would not hold up to any form of robust analysis. We eventually settled on a matrix of 3*x*2, with 3 age categories and 2 education categories, excluding a matching matrix with the province variable included[6]. The result was a distribution of the sample of pregnant women within the OHS as indicated in the table below.

[6] Initially, we had started out with a 3*x*3 matrix that had 3 instead of two education variables. However the resulting prevalence was so similar for the matric+ and tertiary categories, that it was decided to collapse these two categories into one. The prevalence figures for this are available from the authors on request.

Table 7: Number of Pregnant Women Attending Public Health Clinics

Category	Less than Matric	Matric +	Total
12- 19	260	17	277
Share of Total	*15.67*	*1.02*	*16.7*
20-29	591	238	829
Share of Total	*35.62*	*14.35*	*49.97*
30-49	464	89	553
Share of Total	*27.97*	*5.36*	*33.33*
Total	1315	344	1659
Share of Total	*79.26*	*20.74*	*100*

Source: OHS99

It is clear from the data above that the biggest proportion of the sample were those women between the ages of 20 and 29, who reported having less than a matric as their highest educational qualification. These women formed about 36% of the sample and numbered 591. The second-largest cell were women again with less than a matric, but between the ages of 30 and 49. Indeed, these two categories of women formed just under 65% of the entire sample of pregnant women. In addition, it is evident that the dominating characteristic in this sample is the level of educational attainment, where close to 80% of the women had not completed a matric.

Given that we had organised the sample of pregnant women according to the above categories, the next step was of course to determine the prevalence figures according to these categories, as derived from the Ante-natal Clinic data for 1998. The table below therefore presents these prevalence figures for all ante-natal clinic attendees. The prevalence figures here are drawn from an underlying sample of 15 089 women, of whom, 3170 tested positive for HIV/AIDS. The distribution of the full sample within the cells served as the denominator, while the distribution of those testing positive within the cells was used the numerator. In so doing, we arrived at the intra-cell prevalence figures according to the pre-defined education-age categories.

Table 8: HIV+ Prevalence % in ANC98, Education by Age Categories

Category	Less than Matric	Matric +	Total
12 - 19	18.05 *(0.013)*	25.1 *(0.112)*	19.04 *(0.011)*
20 - 29	23.94 *(0.009)*	19.71 *(0.029)*	24.13 *(0.008)*
30 - 49	15.59 *(0.007)*	17.11 *(0.029)*	15.35 *(0.007)*
Total	20.68 *(0.008)*	19.4 *(0.025)*	22.82 *(0.007)*

Source: ANC98

1: Note: Standard Errors are in parenthesis, and are corrected for according to frequency weights, the primary sampling unit and sampling stratification.

The data reveals a national prevalence figure of 22.82 % for this full sample. As with the 1998 HIV ANC survey, the highest prevalence was in the 20-29 age group, who also formed the largest sample of the women in the OHS dataset. Noticeably, the lowest prevalence is for women with less than a matric and between the ages of 30 and 49. These figures provide very early evidence that the pandemic has a strong youthful dimension and indeed does not seem to vary greatly across educational categories.

Given the above two tables, we now have our data sets aligned for the matching process. Specifically, the procedure was to randomly assign an HIV status to pregnant women in the OHS sample, according to the distributional parameters outlined in the ANC prevalence figures. Hence, for example, the prevalence for women in the age category 12-18 with less than a matric was reported in the ANC98 as 18.05%. Within the OHS data we then randomly assigned an HIV-positive status to 18.05% of all women in this particular cell. This process is repeated across all the categories, to arrive at an *imputed* HIV status for the pregnant women in the OHS sample. However, as alluded to earlier, the random process itself is subject to instability, in that different individuals are likely to be assigned an imputed HIV status with each random allocation undertaken. This is no doubt exacerbated by the small sample size in the OHS survey. In order to deal with this instability, we undertook 35 replications of the random assignment according to the ANC rates, and subsequently reported the mean of the variable under consideration. For example, the differing poverty estimates across imputed HIV-status reported on below, would represent the mean figures derived from the 35-round random allocation process. The table below therefore, in using this 35-round random replication process, presents the *ex-post* estimates of imputed HIV status within the OHS sample.

Table 9: Imputed HIV Status for Pregnant Clinic Attendees

Category	Less than Matric	Matric +	Total
12 - 19	18.08 *(0.023)*	23.53 *(0.103)*	18.41 *(0.023)*
20 - 29	23.86 *(0.018)*	19.75 *(0.026)*	22.68 *(0.015)*
30 - 49	15.52 *(0.017)*	16.85 *(0.040)*	15.73 *(0.015)*
Total	19.77 *(0.011)*	19.19 *(0.021)*	19.65 *(0.009)*

Source: OHS99 & ANC98

1: Note: Standard Errors are in parenthesis, and are corrected for according to the primary sampling unit and sampling stratification.

Although the imputed overall HIV prevalence was lower (19.65%) than the prevalence estimated from the ANC data (22.1%), age-specific figures showed consistency with the ANC dataset in that the highest prevalence was observed in the 20-29 age group. Given that we randomly distributed, through 35 replications, HIV-status across a sample size that is smaller than that contained in the ANC data, we expect some variation in the resulted imputed prevalence figures from the true estimates. Hence, we see for example that in the 12-19: Matric + cell, the imputed HIV+ prevalence was 23.53% as opposed to the antenatal survey estimate of 25.1%. The discrepancy in the estimates was not too large to render the imputed HIV-status figures unusable.

The 35-round random distribution of imputed HIV status according to the two covariates (education and age) then allowed us to create the dummy variable for HIV-negative (0) and HIV-positive (1) pregnant women in the sample. The binary values were assigned according to the distribution stipulated in the matching process. This then represents the final stage of our matching process – namely that we now have within the OHS99 dataset, the sample of pregnant women between the ages of 12 and 49 who attended a public clinic in the last year and through the cell distributions of the ANC data, have been assigned an *imputed* HIV status. It is this assumed marker, an imputed HIV status, on these individuals that provides us with an entry point into the rest of the household survey data, which provides fairly detailed labour market, poverty and broader welfare information on individuals and households. It is to a descriptive analysis of this information that we now turn.

Socio-Economic Correlates of Imputed HIV-Status

One useful analytical tool in trying to determine some of the socio-economic factors associated with HIV prevalence, is to segment the descriptive statistics into the household and individual characteristics that are correlated with imputed HIV-status. Hence, we will attempt, in the sections that follow below, to interrogate the nature of the households that these women reside in, allowing for the fact that this is a relatively small sample. Secondly, clearly the individual characteristics of these women are important in a variety of different ways, as they may further illuminate on the nature of the pandemic.

Household Characteristics by Imputed HIV Status

Constantly mindful of the limitations inherent in a sample that it is relatively small, we nevertheless initiate the empirical overview in this section, by examining the spatial distribution of the epidemic. The distribution of households in this sample, indicated that about 56% (44%) of all women in the sample, lived in urban (rural) areas, indicating a distinct urban bias in the sample(Table 10).. According to imputed HIV status, there was a larger proportion of imputed HIV-positive women in urban areas, compared with imputed HIV-negative women – 56% compared with 54%. We report in the table below the prevalence figures for women residing in rural and urban households. The evidence suggests that the prevalence in urban areas is marginally higher

Table 10: Imputed HIV Status for Pregnant Clinic Attendees, By Location

Location	Estimate	Std. Err.	95% Conf. Interval	
Rural	19.64	0.0146	0.16763	0.22508
Urban	19.66	0.0131	0.17093	0.22231
Total	19.65	.0097585	.1773635	.2156443

If we assume however that the small sample size may drive the insignificance in the results, and take the higher urban prevalence as robust, these results would be consistent with other studies within Africa, where higher rates were observed in urban relative to rural settings (Alfowesco, Topouzis 1998). This is attributed to a range of factors such as population density rapid urbanization, unstable families and migrancy, all which are, reported to lead to higher HIV rates (Quinn, 1996).

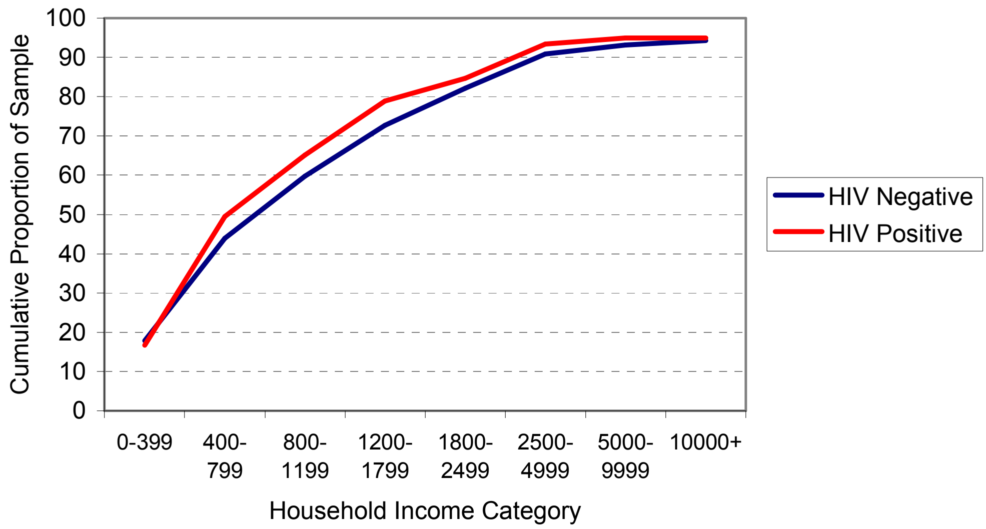
One key welfare aspect of household attachment amongst individuals is the level of household income infected individuals may have access to. Within the context of this paper, it was therefore important to examine the distribution of household income according to imputed HIV status. The income data within the OHS was only provided according to categories, and hence point estimates were not possible to derive[7]. The most optimal presentation of the data then, was to examine the distribution of household income across

[7] What this means of course is that measures of poverty and inequality, typically the headcount index, the poverty gap measure and Gini coefficient are not possible to derive with data that is in ranges rather than point estimates of income for each household. What would of course been very illuminating in this regard is to compare the poverty and inequality levels of households by imputed HIV status.

the different categories by imputed HIV status. We present the cumulative distribution of household monthly income across the two categories in the figure below.

Graphically, the two distributions suggest that there is first-order dominance (Figure 3). What this means is that for each income category provided the distribution of household income for HIV-positive women lies above that of the distribution for HIV-negative women[8]. It is only at the tail-ends of the distributions that there appears to be some convergence. This provisional evidence, seems to suggest that women with an imputed HIV-positive status emanate from poorer households.

Figure 3. Cumulative Density Function of Total Monthly Income by Imputed HIV Status (OHS99)



More specifically, the proportion of households with imputed HIV-negative women, earning a monthly income of R799 or less is about 44%. The corresponding figure for households with imputed HIV-positive women is approximately 49%. Further up the distribution, the percentage of households earning R1199 per month or less is 60% for imputed HIV-negative households and 69% for imputed HIV-positive households. Put differently, if we imposed an arbitrary household poverty line of say R1799 per month, then the headcount index for imputed HIV-negative households would be 73% and the measure for imputed HIV-positive households would be 79%. This crude measure of poverty, with a line imposed given the nature of the data, would at least provisionally suggest that there is a differential access to the level of income across imputed HIV-status in this sample of pregnant women. Apart from small sample size constraints, one of the critical caveats to this first result, is that as we do not have point estimates of monthly household income, it is not possible to accurately assess whether these two distributions are significantly different from each other.

[8] Theoretically, if we assume two distribution P(y) and N(y), then if for all values of y from 0 to y^{max} P(y) lies above N(y), then we can deduce that the distribution P(y) first-order dominates over the distribution N(y).

We therefore interrogate the data further, in an attempt to explain and validate these initial findings through testing for significant differences in the sub-section below.

Measures of Absolute and Relative Poverty Differences

Testing for significant differences in the form of the above cumulative distribution functions is not possible, as indicated above, without some sense of a continuous income variable. We unfortunately, as a result of the data, only have the income variable reported according to pre-specified categories. These categories, as Figure 3 above illustrates range from monthly household income of between 0 and R399 to R10000 or more. In an attempt at providing some semblance of a point measure, we have created point estimates within the categories. This was done by simply placing each household within a category range, at the mean of the range. Hence, for households in the 0-399 range, all were placed at a monthly income value of R199.5. We therefore derived a sample of household income for imputed HIV-positive and negative women, that effectively allows us to provide proxies of poverty measurement, and furthermore test for differences.

In order to measure poverty levels in this sample, according to monthly income, we utilized the general class of poverty measures first proposed by Foster, Greer and Thorbecke (1984), and know more widely as the 'FGT' measures of poverty. The FGT index of poverty measures can be represented in general form as:

$$P_\alpha = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{z-y_i}{z}\right)^{\alpha} \mid \left(y_i \le z\right) \qquad (1)$$

where n is the total sample size, z is the chosen poverty line, and y_i is the standard of living indicator of agent i. The parameter α measures how sensitive the index is to transfers between the poor units. Note that the index is conditional on the agent's income, y_i , being below the designated poverty line, z. The headcount index is generated when α=0, and in this case equation (1) is then simply the share of agents below the poverty line. The poverty gap measure (PG) is generated when α=1, and therefore for a given poverty line z is presented as:

$$P_1 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{z-y_i}{z}\right) \mid \left(y_i \le z\right) \qquad (2)$$

As is clear, the PG represents a direct measure of agents' incomes relative to the poverty line. A first advantage of the FGT index, is its additive decomposability, which allows for sub-group poverty measures to be summed to form a society-wide measure without any loss of generality. More directly though the PG allows a more nuanced assessment of relative poverty – something that the standard headcount index cannot provide. Utilizing this measure then, we derive in the table below, the headcount and poverty gap measures for imputed HIV-positive and HIV-negative households, as a direct complement to the cumulative distribution functions above. In the table below, we set a poverty line (z) of R1499.5, which is of course the midpoint of the 4th income category – 1200 to 1799 Rands per month.

Table 11: Headcount and Poverty Gap Measures for Imputed HIV-Status Households[9]

Poverty Measure	Estimate	Std. Error	Confidence Interval	
		HIV Negative		
Headcount Index	0.66	0.014	0.6290	0.6850
Poverty Gap	0.43	0.011	0.4045	0.4466
		HIV Positive		
Headcount Index	0.70	0.027	0.6499	0.7581
Poverty Gap	0.45	0.020	0.4082	0.4888
t-Statistics	**Headcount**	-1.52	**Poverty Gap**	-0.99

It is clear from the table that, at the poverty line of R1499.5 per month, 66% of households with imputed HIV-negative women were poor (Table 11). In addition, for these poor households, they were earning an income that placed them, on average, 43% below the stipulated poverty line. The latter is our measure of relative poverty, captured through this poverty gap estimate. In comparison, households with imputed HIV-positive women yielded a headcount estimate of 70%, and a poverty gap of 45%. Hence, the data provisionally suggests that imputed HIV-positive women may be residing in households that are poorer than those wherein imputed HIV-negative women live. Note that this statement is true for both the absolute measure of poverty (the headcount index) and the relative measure of poverty (the poverty gap). Given that we have the standard errors on each of these measures it was possible to calculate whether these differences in poverty were in fact significant. The results of the *t*-statistic reveal that for both the headcount and the poverty gap measure, the estimates for imputed HIV-positive and negative women are not statistically different from each other. However, it is possible that this is the result of a small sample size, rather than representing a robust acceptance of the null hypothesis that these poverty estimates are the same across both cohort of women. At a minimum though, the results do point to the fact that household vulnerability levels may be a critical marker in understanding the burden of the epidemic.

Household Size and Dependency Ratios

One of the key household characteristics that defines the socio-economic constraints faced by individuals, remains the size of the household. This measure provides an initial sense of the scale of income required within each household. We present in the table below household size, again of course as a mean of the 35 replications, for imputed HIV-positive and negative households, ranging from a 1 member household to an 8+ size. Firstly, it is evident that for this sample of pregnant women close to 90% were living in households with at least 3 members. Indeed the largest single category were households with 5 members in them.

---

[9] The *t*-statistic is calculated according to the formula $t = \frac{P_A - P_B}{s}$ *where* $s = \sqrt{\mathrm{var}(P_A) + \mathrm{var}(P_B)}$. Further details can be found in Ravallion & Datt (1996) and Kakwani (1993).

Table 12: Household Size Distribution by Imputed HIV Status

Household Size	Negative	Positive	Total
1	0.18	0.72	0.24
2	1.54	0.36	1.51
3	10.68	11.19	10.61
4	13.21	11.91	12.6
5	15.93	12.64	15.25
6	12.04	15.16	13.14
7	11.95	13	11.39
8+	34.47	35.02	35.25
Estimated Mean	6.93 (0.1094)	6.87 (0.1833)	6.91(0.0847)

1: Note: Standard Errors are in parenthesis, and are corrected for according to the primary sampling unit and sampling stratification.

The estimated means are particularly interesting though, in that they reflect a sample of pregnant women as a whole whose household size is in fact above the national mean- a figure we obtain from the aggregate sample of individuals in the OHS99. The national estimated mean household size is 5.84, while that for all African-headed households is 6.13 and for White-headed households the mean is 3.63. The values for both imputed positive and negative HIV women are above these national aggregates(Table 12). This suggests one component of vulnerability that may be linked more broadly to the status of all the women as being pregnant and users of public health services – the latter being a strong marker for an individual's position in the income distribution. In terms though, of the comparison between imputed HIV-positive and negative women in the above sample, the estimated mean household size for positive women, at 6.87, is marginally smaller than for negative women (6.93). However, we are again constrained by the small sample size in terms of deriving robust results. If it was found though that this result was significant, then clearly we would have evidence that counters other results on household size differences across HIV status (Bachman, 2003). In addition, a significant result of this sort would indeed, at least initially, be counter-intuitive in terms of the poverty estimates provided above. Interestingly though, if one examined the proportion of women located in households with 6 or more members by HIV status, we find that while these households formed 58.46% of all imputed negative households, in the case of imputed positive households, the figure was 63.18%.

One direct measure of dependency levels is of course the number of children residing within a household. We present the data below according to imputed HIV status, and defined any individual who is under the age of 12 as a child. The largest single category is amongst households with 2 children, and this is true for both HIV-status categories. Cumulatively it is evident that the samples are both fairly evenly divided according to households which have two or fewer children and those with 3 or more children.

Table 13: Distribution of Children in Household by Imputed HIV Status (OHS99)

Number of children in Household	Negative	Positive	Total
0	0.63	1.44	0.72
1	20.54	24.19	21.88
2	28.87	26.35	27.49
3	24.89	18.41	23.03
4	11.49	16.97	13.44
5+	13.56	12.63	13.43
Total	0.63	1.44	0.72
Estimated Statistics			
Mean	2.83	2.725	2.81
Std. Error	0.05407	0.0959	0.0424

The estimated means are instructive here, as they reveal a dependency ratio of about 2.8 for the sample of pregnant women as a whole(Table 12). In addition, note that the figures for imputed HIV-positive and negative women indicate that the mean number of children for the latter cohort is in fact higher (2.83) than for imputed HIV-positive women. Interestingly, the national mean figures were substantially lower than these reported for pregnant women attending public health clinics. For example the national mean number of children in a household is 1.75, rising to about 2 children per African household. Again, this reflects more generally on the broader sample drawn here, rather than any imputed HIV-status attribute. Ultimately though, the evidence here (again small sample size issues aside) suggests that there may be a marginally higher number of children in imputed HIV-negative households, compared with imputed HIV-positive households. Amongst other factors, this may reflect the impact of HIV on fertility, as evidence suggests that HIV infection lowers fertility.

Attachment to Wage Earners and the Quality of Support

The labour market remains probably the most important conduit for households to source income on a regular basis. More broadly, the notion here is that availability of wage income through formal or informal employment, determines the nature of household poverty and inequality in a society; and in all economies of course, the labour market remains the key access point for this income accumulation. Simply put, the labour market remains the filter through which the long-run distributional and poverty outcomes of a society are shaped. Within the context of this paper then, it is necessary to examine the degree to which imputed HIV-positive and negative women have access to employed individuals within the household. Specifically, we attempt to tabulate the number of employed individuals residing within each of these households – as a first step in gauging how effectively labour market activity within these respective households may be translating into lower levels of vulnerability.

Table 14 therefore provides the distribution of these households according to the number of employed individuals within each household[10]. It is clear firstly that in both categories, more

[10] Note that we are effectively including the pregnant women in the distribution of the number of employed within the households. Hence, the number of employed includes these women, whether they are working or not. We turn in the next section to individual characteristics to shed more light on their labour market status within the sample.

than half of the households have access to at least one employed individual. Note of course the this data does not tell us about the *quality* of this support in the form of monetary value of the wages earned by each of these employed individuals.

Table 14: Distribution of Employed Across Households by Imputed HIV Status

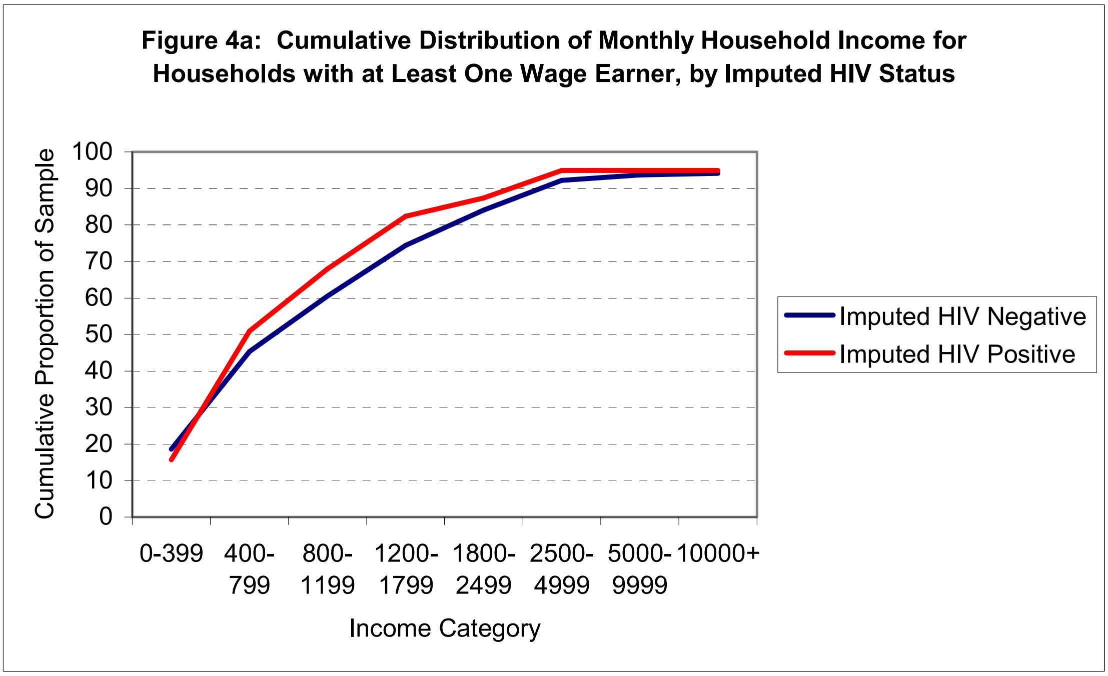
Number of Employed	Negative	Positive	Total
0	44.52	42.24	43.53
1	31.31	29.96	30.86
2	13.67	16.25	14.89
3+	10.5	11.55	10.72
Total Share	100	100	100
Estimated Statistics			
Mean	0.966	1.036	1.000
Std. Error	0.036	0.0717	0.0297

In the case of imputed HIV-negative women, approximately 55% of these household types had access to at least one wage-earner. In the case of imputed HIV-positive women, the figure was about 58%. At face value then, the imputed HIV-positive women appeared to have access to more income earners than imputed HIV-negative women. This fact is borne out by the estimated mean figures, where the imputed HIV-negative mean was marginally lower (0.966) than that for the imputed HIV-positive sample (1.036). Ostensibly then, this evidence may appear to contradict the initial cumulative distribution functions, which suggested that imputed HIV-positive women came from poorer households – given that they appeared to have access to a greater number of wage earners within a household.

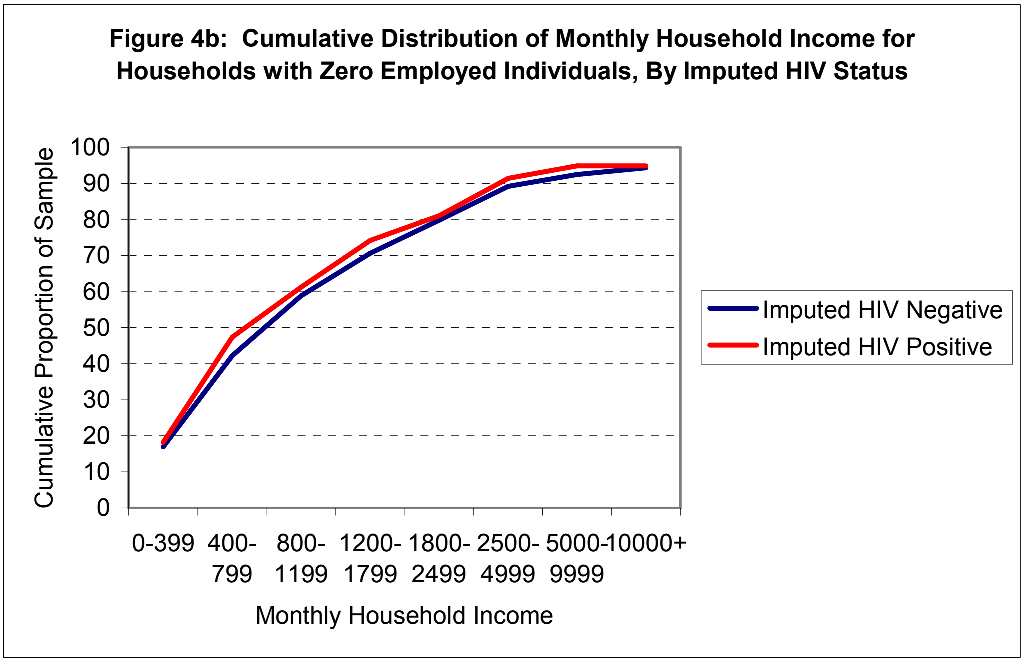
There would appear to be one important caveat to these results: While it may be possible that more women from the HIV-positive cohort were located in households with more (or the same number of) earners, ultimately the quality of the support offered could be lower. Put differently, while the number of employed may be larger, the aggregate value of income earned by these individuals could still potentially be smaller than in the case of imputed HIV-negative households as shown in the initial analysis that examined cumulative income distributions and poverty measures for the full sample (Figure 3 and Table 11). Again, we accept completely that the small sample size problem may render this analysis highly questionable. However, we undertook it at a minimum, as an indication of the linkages that can and should be made when trying to understand household attachment patterns amongst HIV-afflicted individuals.

We produced below then cumulative distribution functions of imputed HIV-status households, with the first figure examining the distribution for those household with 1 or more earners (Figure 4a), and the second figure illustrating the comparative distributions for households with zero earners (Figure 4b). The results are most instructive here, and reflect on the importance of linking 'quantity of support' in terms of number of earners to 'quality of earners' in the form of the value of income entering the household. It is clear, at least graphically, from Figure XXXa that apart from the lowest income band, there is clear evidence that the proportion of imputed HIV-positive households earning below any given income band is greater than the imputed HIV-negative cohort. For example the figures show that while 60% of HIV-negative households earned below this figureR1199 per month, 68%

of imputed positive households earned below this figure. This immediately begins to suggest that while the quantity of support available to imputed HIV-positive households may be greater (or the same as) imputed HIV-negative households, the quality of support in the form of monetary value is different.



We examined in turn, in the Figure 4b below, the cumulative distribution functions for those households who reported having zero wage income earners. As a second category of access to income support, the initial tabulation in Figure 4a, suggested that the imputed HIV-positive cohort had a lower proportion of these household types.Figure 4b shows though, the quality of the support offered differed markedly.



Hence, we found that for every income category, the cumulative function for imputed HIV-positive households was above that of imputed HIV-negative households. Put differently, across all household income bands, first order dominance holds – with the HIV-positive distribution first-order dominating the HIV-negative distribution. For example, within the income range 1200-1799 per month, the proportion of imputed HIV-negative households below this range is 70.61%, compared with 74.14% for imputed HIV-positive households. In addition, within the range R400-799, the proportion below this is 42.24% for imputed HIV-negative households and 47.40% for imputed HIV-positive households.

We attempted then, to provide point estimates of vulnerability through the estimation of the headcount and poverty gap measures for the two samples. As with the above estimation, we retained the poverty line of R1499.5 per month. The results in Table 15 present the poverty measures for the imputed HIV-positive and negative women who lived in households with at least one wage earner. The results for those with zero wage earners are provided in the appendix below. Firstly, it is clear that both the headcount and poverty gap measures were higher for imputed HIV-positive households. In terms of the latter, while imputed HIV-positive households were on average 46% below the poverty line, the figure for imputed HIV-negative households was 44%. Put differently, imputed HIV-positive households were relatively more poor than imputed HIV-negative households, where both had at least one wage earner residing.

Table 15: Headcount and Poverty Gap Measures for Imputed HIV-Status Households With At Least One Wage Earner [11]

Poverty Measure	Estimate	Std. Error	Confidence Interval	
HIV Negative				
Headcount Index	0.67	0.019	0.6281	0.7030
Poverty Gap	0.44	0.014	0.4070	0.4637
HIV Positive				
Headcount Index	0.73	0.035	0.6618	0.8007
Poverty Gap	0.46	0.026	0.4064	0.5097
t-Statistics	**Headcount**	-1.64**	**Poverty Gap**	-0.76

Note: Poverty Line z, is set at 1499.5 per month per household.
**: Significant at the 10% level.

We also undertook, as with the above table (Table 15), the *t*-test for a significant difference in the two poverty measures. The headcount measure is significant at the 10% level, while the difference in the poverty gap estimates was insignificant. The table above, was repeated for households with zero wage earners, at the same poverty line (See Appendix Table 24). Here again the headcount and poverty gap were higher for imputed HIV-positive households. However, for both measures, the *t*-statistic was insignificant.

Finally, we also undertook sensitivity tests, but raising the poverty line to R2149.5 per month, the midpoint of the range, 1800 to 2499 (Appendix Table 26). These results indicate the same broad estimated trends, with the imputed HIV-positive poverty results higher than those for the imputed HIV-negative households. Notably, the result is significant in the case of the differences in the headcount across 1+ wage earner households.

These results are an important extension to our initial finding - namely that imputed HIV-positive women appeared to be residing in poorer homes than imputed HIV-negative women. These result are suggestive of two key features of the vulnerability profile of the pandemic. Firstly, imputed HIV-positive and negative cohort were located in similar proportions, within wage-earning households. Secondly though, the access to wage earners within the household, amongst imputed HIV-positive women location was not optimal, given that they resided in households with lower levels of income. The households with either no wage earners or at least one, were poorer in the case of HIV-positive women, than HIV-negative women. Put differently, the labour market participants that positive women were accessing within the households are likely to be more vulnerable than those accessed by imputed HIV-negative women. Finally, it is important to add, once again, that we view these results as tentative, given the low levels of significance in poverty differences. However, we do believe that the lack of significance in the *t*-tests is driven by the small sample size, with

[11] The *t*-statistic is calculated according to the formula $t = \frac{P_A - P_B}{s}$ *where* $s = \sqrt{\mathrm{var}(P_A) + \mathrm{var}(P_B)}$. Further details can be found in Ravallion & Datt (1996) and Kakwani (1993).

the broad direction and nature of differences likely to remain intact, should a larger sample size be utilized.

Attachment to the Unemployed and Non-Wage Earners

As an extension to the above, it would be useful to determine the nature of intra-household dynamics in the form of the access to non-wage earners, such as pensioners, that this sample displays. In addition, it is important to gauge the burden placed on these households by the unemployed. We turn to the latter in the table below, which examines the distribution of the unemployed across the two household types. It is evident firstly, that almost all households in the sample report the presence of an unemployed individual[12]. In the case of the imputed HIV-positive sample, there are no households with zero unemployed individuals in them.

In the case of imputed HIV-positive and imputed HIV-negative households, 63% of both cohorts report having between one and four unemployed individuals within the households. In addition, as the figure in the appendix indicates, this distribution peaks in households with 3 unemployed individuals, for both groups. Hence, for both categories of pregnant women, a disproportionate share report living in households with 3 unemployed individuals in them.

Table 16: Distribution of Unemployed Across Households by Imputed HIV Status

Number of Unemployed	Negative	Positive	Total
0	0.09	0.00	0.06
1	3.44	3.96	4.24
2	15.75	16.97	16.21
3	23.53	24.55	22.78
4	19.91	17.33	19.17
5+	37.28	37.19	37.54
Total	100	100	100
Estimated Statistics			
Mean	4.327	4.137	4.248
Std. Error	0.0746	0.1181	0.057

The estimated statistics reveal that the mean for both samples is just over 4 individuals per household. In the case of imputed HIV-positive households, it is in fact slightly lower at 4.137 as opposed to 4.327 for imputed HIV-negative households. Again though, it is important to determine whether the levels of income across these household types may be different. In other words, while the distribution of the unemployed across the two cohorts may be similar, it is possible that imputed HIV-positive households do still emanate from poorer homes. The headcount and poverty gap estimates for imputed HIV-positive and negative households according to those with at least 1 unemployed person in them, bear this out. Hence, we find that the headcount index for imputed HIV-positive women in this case is at least 4 percentage points higher than that for imputed HIV-negative women (70% as opposed to 66%) and a 2 percentage point difference in the poverty gap measure (45% versus 43%). This reiterates our two earlier results, namely that imputed HIV-positive women are residing in poorer households - a result that we would of course expect given

---

12 As the section below will elucidate on, we use the expanded definition of unemployment here to differentiate between the unemployed, employed and economically inactive.

that we are essentially replicating the sample of women in the aggregate figures presented above.

It has been shown that income transfers entering households remain one of the more effective instruments for poverty alleviation. South Africa is no exception, and the old age pension in particular has been shown to be a highly optimal transfer in terms of reaching poor households. We try, in the table below to examine whether access to the old age pension, and another important income transfer – remittances – differs across these two household categories[13]. Firstly, it is clear that in the case of both old age pensions and remittances, the majority of households in either HIV category report having no access to these transfers. In the case of old age pensions, about 80% of all households in the sample have no pensioner living with them. For remittances the figure is higher, with about 85% of all households not being recipients of remittances. Hence, by the two major income transfers, households appear to have minimal access to these individuals. For this reason, we tabulated the access that households have to any form of income transfer, which would include the two noted already, as well as those such as disability grants, the child support grant, unemployment insurance, care dependency grant and so on. In this case, clearly the level of support provided increases, as between 69 (imputed HIV-negative) and 70% (imputed HIV-positive) of households report accessing at least one form of these transfers.

[13] We have excluded other income transfers such as the chid support grant, disability grant and unemployment insurance, given that a very small proportion of the sample in either of the two categories reported having access to these transfers. For example, in the case of the disability grant, between 95 and 97% of households in each of the two categories reported having no access to this grant.

Table 17: Presence of Non-Wage Earner(s) in Household by Imputed HIV-Status Households

Pensioner(s) Present	Negative	Positive	Total
Pensioner Present?			
No	80.9	78.7	80.23
Yes	19.1	21.3	19.77
Total	100	100	100
Estimated Statistics			
Mean	0.2253	0.2599	0.2344
Std. Error	0.0151	0.0322	0.0125
Remitter Present?			
Remitter(s)	**Negative**	**Positive**	**Total**
No	84.44	86.28	84.69
Yes	15.56	13.72	15.31
	100	100	100
Estimated Statistics			
Mean	0.1855	0.1877	0.1886
Std. Error	0.0146	0.0399	0.0134
Any Non-Wage Earner Present?			
No. of non-wage income earners	**Negative**	**Positive**	**Total**
No	30.68	29.96	29.97
Yes	69.32	70.04	70.03
Total	100	100	100
Estimated Statistics			
Mean	1.3819	1.472	1.42
Std. Error	0.04231	0.0898	0.0352

The estimated means reveal that across both transfer types and the category covering all grants, imputed HIV-positive households appear to have access to a marginally larger number of these recipients for example, the mean number of old age pensioners in imputed HIV-positive households is 0.26, compared with 0.23 for imputed HIV-negative households. Clearly though, these differences to all intents and purposes, suggest that imputed positive and negative households are accessing the same quantum of income transfer recipients.

Predictably, the next question to be answered is that despite being able to access the same number of recipients, does this fact enable imputed HIV-positive households to be less poor, or at least not poorer, than imputed HIV-negative households? The results are provided in the Table 18, where we tabulate the poverty measures for those households with at least one non-wage income earner resident, by imputed HIV status. Table 18 makes it plain that the same access to the number of income transfers is not sufficient to make imputed HIV-positive less, or as poor as imputed HIV-negative households. Hence, by both the headcount and poverty gap measures, imputed HIV-positive households are poorer.

Table 18: Headcount and Poverty Gap Measures for Imputed HIV-Status Households With At Least One Non-Wage Income Earner

Poverty Measure	Estimate	Std. Error	Confidence Interval	
		HIV Negative		
Headcount Index	0.67	0.017	0.6363	0.7031
Poverty Gap	0.43	0.013	0.4050	0.4549
		HIV Positive		
Headcount Index	0.74	0.031	0.6802	0.8044
Poverty Gap	0.47	0.024	0.4209	0.5141
t-Statistics	**Headcount**	-2.03**	**Poverty Gap**	-1.40

Note: Poverty Line z, is set at 1499.5 per month per household.
**: Significant at the 10% level.

The absolute measure is statistically significant, indicating a robust measure of poverty differences between the two samples of pregnant women, while the relative measure of poverty is insignificant. Ultimately though, this suggests that income transfers in and of themselves are, at least in this small sample, not able to lift a greater proportion of imputed HIV-positive households out of poverty, relative to imputed HIV-negative households.

The above data raises an important question concerning intra-household dynamics that may be resulting in these skewed poverty outcomes. Despite having ostensibly the same access levels to both wage and non-wage earners and yielding similar household sizes, together with dependency ratios, imputed HIV-positive households seem to be consistently poorer than imputed HIV-negative households. We turn now to the individual characteristics of the sample, in an attempt to deal in greater detail with this issue.

Individual Characteristics by Imputed HIV Status

Having dealt with the characteristics of the households wherein these two cohorts of pregnant women reside, it is important to try and assess how their individual characteristics may differ across a range of covariates. In the sections below then, we deal with the differing labour market status of these individuals, wages earned, their marital status, relationship to the head of the household and so on. The broad aim of this section then would be to try and understand what specific individual characteristics by imputed HIV status may possibly be shaping the differential household outcomes observed above.

Labour Market Characteristics by Imputed HIV Status

It is possible that the differential household poverty levels amongst imputed HIV-positive and negative women is in fact driven by the nature of their labour market status. The table below attempts to detail these labour market characteristics, for the two cohorts. We provide below a snapshot of labour market status according to both the strict (or official) definition of unemployment and the expanded definition. We report in each case the share of women within each labour market category, namely; not economically active (out of the labour force), employed and unemployed. As a derivation of these, we also report the unemployment rates for the two groups.

The table firstly makes it clear that, according to either definitions, the share of imputed HIV-positive women out of the labour force is marginally lower, with 64 and 41% not economically active according to the two definitions.

Table 19: Labour Market Status by Imputed HIV Status (OHS99)

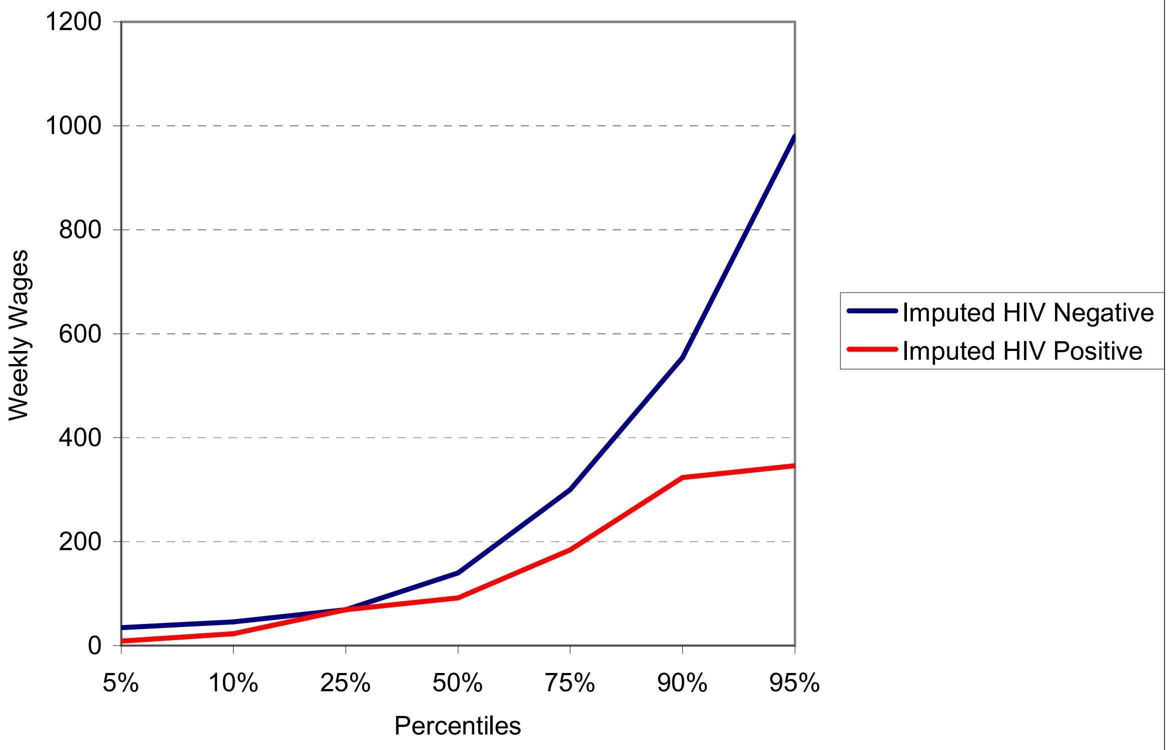
Labour Market Status	Negative	Positive	Total
Official Definition			
Not Economically Active	68.3	64.49	67.65
Employed	18.12	15.58	16.96
Unemployed	13.59	19.93	15.39
Total	100	100	100
Estimated Statistics			
Unemployment rate			
Expanded Definition			
Not Economically Active	43.3	40.94	43.39
Employed	18.12	15.58	16.96
Unemployed	38.59	43.48	39.65
Total	100	100	100
Estimated Statistics			
Unemployment rate	40.49(0.0066)	42.40(0.115)	41.19(0.0052)

1: Note: Standard Errors are in parenthesis, and are corrected for according to the primary sampling unit and sampling stratification.

In terms of active participants though, it is evident that the share of women who are employed is greater for imputed HIV-negative women (18%) compared with imputed HIV-positive women (16%). Hence, this is an early suggestion that imputed HIV-positive women are less likely to be in employment (assuming they are in the labour force) than imputed HIV-negative women. This differential share in employment is of course a reflection of the contrast in access to wage income, a vital component of household income. These employment shares are then masked in the differing unemployment rates across the two groups, which we provide in estimated form with their standard errors in the table above. It is evident that the imputed HIV-positive strict unemployment rate is 36%, while for the negative group it is 6 percentage points lower. In terms of the expanded definition, the positive rate is 42%, compared with a rate of 40% for imputed HIV-negative women. Provisionally, what this evidence suggests is that, by both definitions of unemployment, imputed HIV-positive women are more likely than imputed-negative women to be unemployed, and therefore face a higher likelihood of being zero earners. In this manner the labour market enters as one of the key, if not the key, factor in understanding the differential household poverty levels noted in the previous section. The fact that imputed-positive women are more likely to be unemployed, serves as one of the primary mechanisms for ensuring that the households they reside in are poorer than those households with imputed HIV-negative women.

Having established that there is a lower probability of employment amongst the imputed-negative cohort, it is also relevant to determine what the characteristics are for those women across both groups who do have employment. We attempt then, in the data that proceeds below, to examine employment by a series of covariates for these two samples. The broader aim again, is to assess whether these figures reinforce the initial result of differing levels of vulnerability for the two groups. Figure 5 below examines weekly wages by imputed HIV status, for those women who are employed[14]. We have presented the data according to percentiles in the respective wage distributions, ranging from the 5th to the 95th percentile for each of the distributions. The figures are particularly revealing, in that they show distinct patterns in the distributions.

Figure 5. Percentile Distribution of Weekly Wages, by Imputed HIV Status



In the bottom-quarter of the wage distribution (5th to 25th percentile), it is evident that weekly wages are fairly closely correlated relative to the distribution – from the 50th to the 95th percentiles. Throughout the distribution however, the weekly wages of imputed HIV-positive women is below that of imputed HIV-negative women. For example, the 10th percentile wage is R32.10 per week for positive women while it is R46.20 for negative women. Even at the median (the 50th percentile), the weekly wage of imputed positive

[14] The OHS99 does have a question on actual wages earned, but respondents can provide it in any frequency format (weekly, monthly or annual), which is specified in the questionnaire. Hence, it was necessary to convert all monthly and annual wages to their weekly equivalent. This was done using a factor of 0.231 for monthly wages and a factor of 0.0192 for annual wages.

women is lower at R92.40 compared with R140 for imputed negative women. Despite this differential in the bottom half of the distribution, it is clear though that it is in the upper portions of the distribution, that the differentials begin to widen even further. As the data in the appendix on the differing ratios across the percentiles indicates, the ratio of HIV-negative to HIV-positive wages declines steadily to the 50th percentile, and then increases thereafter. For example, while at the 50th percentile the weekly wage of imputed negative women is about 150% of the same percentile wage for imputed positive women, at the 90th (95th) percentile, this ratio falls to 171 (282%). There would appear then to be two key deductions from the wage data above. Firstly, that imputed-positive women who are employed are earning lower weekly wages than those who are imputed-negative. Secondly, this differential tends to decline steadily until the median of the distribution, after which the differentials begin to widen again.

Mindful of our small sample size, the above would provide at least tentative evidence for a second labour market factor accounting for the contrasting household poverty levels. Hence, not only are imputed HIV-positive women more likely to be unemployed, but the evidence here points to the fact that should they be employed, they are likely to earn a wage that is invariably lower than there counterparts who are imputed-negative.

In trying to understand what factors may be shaping these relatively lower earnings amongst imputed HIV-positive women, we tabulate below the distribution of employment amongst the two cohorts by sector, occupation and level of formalisation. Beginning with the latter, the data below suggests that the majority of women in this sample are employed in the formal sector.

Table 20: Formal-Informal Sector Distribution of the Employed by Imputed HIV Status (OHS99)

Sector	Negative	Positive	Total
Formal	75.8	58.33	72.37
Informal	24.2	41.67	27.63
Total	100	100	100

However, the relative distributions indicate that imputed positive women are more likely to be in the informal sector than imputed negative women. This would partly explain the lower wages earned by imputed positive women in the figure above. Indeed, more detailed evidence shows the mean wage of imputed positive women in the informal sector is about 60% of the wage earned by imputed negative women in the formal sector. In addition, the mean wage within the formal sector is higher for imputed negative women, when compared with the imputed positive sample. Examining the relationship of sectoral distribution by HIV status through a Chi square analysis for categorical data, we find a significant relationship between being employed in the informal sector and having a positive HIV status (p=0.03). The Relative Risk of women being HIV positive if they were employed in the informal sector versus the formal sector was 1.89 (CI 95%: 1.05-3.38).

In trying to expand on the role of the labour market in engendering these household poverty outcomes, we further examined in the two tables below, the distribution of employment by sector and occupation for these two cohorts. Table 21 shows the share of employment by main sector. It is evident that the dominant sectors for imputed negative women are Wholesale & Retail Trade and Private Households. The second of these captures, in the main, those women employed as domestic workers. It is clear also that Agriculture and Community Services, constitute over a quarter of the sample of employed imputed positive women. The latter sector is predominantly represented by the public sector.

Table 21: Sectoral Distribution of the Employed by Imputed HIV Status (OHS99)

Sector	Negative	Positive	Total
Agriculture	15.5	13.95	14.23
Manufacturing	12.5	11.63	12.1
Construction	0.00	2.33	0.36
Wholesale & Retail Trade	34.5	20.93	32.74
Transport, Storage, Communication	0.00	2.33	0.36
Finance	4.50	6.98	4.63
Community, Social & Personal Services	16.00	13.95	15.66
Private Households	16.00	25.58	18.86
Not Specified	1.00	2.33	1.07
Total	100	100	100

The dominance of both the above sectors is not surprising, given that these sectors are large employers in the national economy, and tend also to employ larger shares of women than other sectors such as Mining and Financial & Business Services. These figures alone do not do much by way of additionally explaining the differential levels of household vulnerability, although they certainly suggest employment distributions that are distinct across the two cohorts. Marked differentials exist with respect to women working in private households by HIV status, namely a smaller proportion of HIV negative women are employed at the private household level than HIV positive (16% versus 26%)[15].

Ultimately though, the above has yielded a number of important corollaries that match well with the empirical description of significant differences in household vulnerability between these two groups of pregnant women. Firstly, it is evident that one of the key factors driving this differential household poverty status, is the fact that imputed positive women are less likely than imputed negative women to find employment in the labour market. The estimated unemployment rates, confirm that this difference in employment probabilities across the two groups. Secondly, for those women that are employed, there is initial (although insignificant) evidence that their earnings are lower if they have an imputed HIV-positive status. This evidence is then corroborated by the sectoral distributions, which reveal that employed HIV-positive women are more likely to be employed as household domestic workers and less likely to be working in the higher-paying formal sector than their imputed-positive counterparts. Collectively this data is strongly suggestive of the notion that the differential individual labour market characteristics of imputed HIV-positive women is negatively influencing their household poverty status, relative to the imputed HIV-negative cohort.

[15] Note that we did attempt to run a table on the contrasting employment distributions by occupation across the two cohorts. However, close to 20% of the sample of imputed negative women and a quarter of imputed positive women reported an unspecified occupation, rendering this data unusable.

Marital Status and Relationship to Head of Household

In an attempt at trying to understand intra-household dynamics in more detail, we consider below two particular individual characteristics, that may be of relevance when trying to understand the possible differences between these two cohorts. We therefore consider firstly the marital status of individuals, using the categories provided within the OHS99 questionnaire. The rankings for both samples are the same, with the majority of women in both samples reporting not having married before. The higher proportion of imputed positive women (60% as against 52%) who have never married may be partly due to the fact that they are slightly younger on average than imputed negative women.

Table 22: Marital Status by Imputed HIV Status (OHS99)

Marital Status	Negative	Positive	Total
Married – Civil	16.30	13.82	15.88
Married – Traditional	16.03	12.36	15.45
Living together with partner	11.38	11.64	11.21
Widow	1.82	1.82	1.58
Divorced/Separated	2.28	0.36	1.94
Never married	52.19	60.00	53.94
Total	100.00	100.00	100.00

The estimated figures show that the mean age for imputed positive women is about 26 years, compared with a mean of 27 for imputed negative women, although of course it is clear that this differential could not serve as the sole explanation for the marital status discrepancy. Of course more broadly, the data reflects on the fact that over half of the sample of all pregnant women are relatively young and report not having been married before. As a consequence, the data also suggests that a higher proportion of imputed negative women are married, either according to civil or traditional rights.

In terms of the intra-household dynamics, Table 23 attempts to locate the relationship of the particular pregnant women to the head of the household. Interestingly, in at least 15% of cases for both cohorts, these women reported being the head of household themselves. Notably, a larger share of imputed negative women report themselves as the head of the household.

Table 23: Relationship to Head of Household, by Imputed HIV Status (OHS99)

Relationship to Head	Negative	Positive	Total
Head	16.29	13.00	15.19
Wife/Partner	28.42	25.99	27.67
Daughter/Step Child/Adopted Child	32.58	38.99	34.6
Sister	3.35	4.69	3.86
Grandchild	6.52	6.14	6.21
Other relative	12.04	9.39	11.57
Non-related persons	0.81	1.81	0.90

In the majority of cases, women reported being either a biological, adopted or step-daughter to the head of household, with a larger proportion of imputed positive women in this category relative to the imputed negative cohort. As larger proportion of imputed negative women reported being the wife or partner to the head. While the age differences noted above, certainly feed into this differential status, there is clearly some provisional evidence in this albeit restrictive, sample, that imputed positive women are more likely than imputed negative women to be living with their parents within the household.

Conclusion

This study through an exploratory, but promising methodology, aimed to provide a tentative analysis of the relationship between HIV, poverty and labour markets. It is clear from the paper that the relationship between poverty, labour markets and HIV is not homogenous but multi-dimensional in character. The analysis examined these inter-relationships at both the household and individual level. The key findings from the analysis suggest that imputed HIV positive women come from poorer households than imputed negative women.

The relationship between imputed HIV status and the various descriptors of poverty and labour market status is palpable and apparent at the household level. The household level analysis revealed that that there is a differential access to the level of income across imputed HIV-status in this sample of pregnant women. More detailed analysis gauging poverty through the use of indices such as the headcount ratio and the poverty gap showed that the imputed HIV-positive cohort were significantly different from the uninfected group in that the former resided in poorer households. This is a critical result as it provides initial evidence suggesting that imputed HIV-positive households are worse off than the imputed HIV-negative households in terms of both absolute and relative poverty levels.

Apart from absolute and relative poverty, various other interrelated factors are reported in the literature to be associated with HIV such as income inequality, deteriorating health and education services, migration due to economic hardship and civil strife, all of which disrupt the household and family life (MAP1998). However, it is critical to note that whilst the evidence in this study suggests an association between poverty and HIV status, given that this was a cross-sectional analysis, it does not in any way provide any definitive evidence on the causal link between HIV/AIDS and poverty.

The household characteristics data revealed that the size of the household by imputed HIV status was marginally smaller for the positive women compared with the negative women. Noting, however that all women in the study were pregnant women attending public sector clinics this sub-group generally come from larger households than the national norm of 5.85. However, the findings of this study are contrary to a South African study examining households in rural and urban areas in the Free State Province, which revealed significant differences in household size by HIV affected and unaffected households (Bachmann and Booysen, 2003).

Apart from household size another household indicator is an individual's attachment to wage earners, whereby the degree of support can be gauged by their access to wage earners. This analysis revealed that imputed HIV-positive women appear to have access to more income earners than imputed HIV-negative women even though they reside in poorer households. However, when we further examine the quality of the support of the wage earners, we find that this level of support to imputed positive women is lower.Thus, the results indicate that households with at least one wage earner, are poorer in the case of HIV-positive women, than HIV-negative women.

In the case of access to non-wage earners, such as pensioners we find that the quantity of support provided to imputed HIV-positive women is greater than for imputed negative women. However, in trying to measure the quality of support, the poverty measures for those households with at least one non-wage income earner resident; by imputed HIV status shows that for the headcount and poverty gap indices, imputed HIV-positive households are poorer. These findings provisionally suggest that income transfers are not sufficient to enough to lift a greater proportion of imputed HIV-positive households out of poverty, relative to imputed HIV-negative households. Despite having ostensibly the same access to both wage earners, and increased access to non-wage earners (in households with similar dependency ratios), imputed HIV-positive households seem to be consistently poorer than imputed HIV-negative households.

Individual characteristics by imputed HIV status were examined in order to further understand why HIV positive women come from poorer homes. The key finding that emerged from this imputation process include the composite observation that women who are most affected by the HIV epidemic are essentially young (aged 20-29), living in urban settings and unemployed. These findings are consistent with a national household survey that examined socio-behavioral factors in relation to HIV infection in South Africa (Shisana & Simbayi, 2002). Higher HIV rates in urban settings have been attributed to the combination of factors such as rapid urbanization, migration, unemployment and unstable families or communities (Quinn 1996).

The differing labour market status of these individuals by HIV status, which manifeststangentially in marital status and relationship to the head of the household, appears to be a key marker for explaining the contrasting household vulnerability levels between these two groups of pregnant women. The study revealed that by both definitions of unemployment, imputed HIV-positive women are more likely than imputed-negative women to be unemployed, and therefore face a higher likelihood of being zero earners. In addition, the imputed-positive women who are employed are also earning lower weekly wages than those who are imputed-negative with the wage differentials tending to be very wide at the tail-ends of the distributions.

The wage differentials are supported by further analysis that examined the sector of employment by HIV status. Imputed positive women are more likely to be employed in the informal sector than imputed negative women. In addition, the mean wage within both the formal and informal sector is higher for imputed negative women, when compared with the imputed positive sample. The high proportion of women in the Retail and household domestic worker sectors is consistent with national trends as these sectors are either larger employers or disproportionate employers of women.

Given that the employment profile suggests that the imputed positive women are more likely than imputed negative women to be unemployed, and if employed, they are likely to be employed at the bottom-end of the occupational ladder, with contrasting lower wages, this profile validates the empirical description of significant differences in household vulnerability between these two groups of pregnant women. The key appears to be employment status of the positive women, which then has a deleterious impact on their household poverty status, relative to the imputed HIV-negative cohort.

On examination of social-safety nets that support these women and strategies the positive women undertake to survive we find that the majority of them resided in households as being either a biological, adopted or stepdaughter to the head of household. This highlights that pensioners may be cross-us subsidizing the entire family and absorbing the impact of poverty in the household due to unemployment of the HIV positive women. In contrast a larger proportion of imputed negative women reported being the wife or partner to the head.

In South Africa, there is a dearth of national level information on a HIV, poverty and its impact of labour markets. However, there is substantial work on modeling and projecting the epidemic using the HIV antenatal survey data and on a more limited scale, some localized analysis of the relationship between poverty and HIV at household level. This study presents an innovative method of using available data namely the ANC and OHS data to simulate the impact of HIV on household and individuals with respect to poverty and labour markets. This analysis provides an initial window into how the HIV epidemic affects young pregnant women attending public sector clinics. Examining the impact of the epidemic in this group is significant as they represent young sexually active adults, who form the foundation of the current and future labour force. This analysis consistently showed through individual and household level analysis that women who are imputed positive are poorer than the imputed HIV negative group. Whilst absolute magnitude and severity of poverty and HIV cannot be measured nor can the causality be tested with the current available data, it provides one with substantial evidence that point to the relationship between these factors. It also provides sufficient evidence that both at policy level and programme implementation level, the HIV epidemic cannot be managed without systematic and rigorous efforts that address poverty at individual and household levels.

References

Awofeso, N (1995) AIDS and tuberculosis/leprosy in Nigeria: the urbanization factor. *Acta Leprol.* 9(3): 149-51.

Bachmann M and Booysen F (2003). Health Economic impact ofHIV/AIDS on South African households: A cohort study. BMC Public Health. 3(1):14

Department of Health (1998). National HIV Sero–Prevalence Survey of Women Attending Public Health Antenatal Clinics in South Africa.

Department of Health (2000). National HIV Sero–Prevalence Survey of Women Attending Public Health Antenatal Clinics in South Africa.

Foster J.E. Greer J & Thorbecke,E. 1984. A Class of Decomposable Poverty Measures, *Econometrica.*Vol 52.

Kakwani N (1993), Statistical Inference in the Measurement of Poverty, *The Review of Economics & Statistics*, 75(4).

MAP (1998). Status and Trends of HIV /AIDS epidemicsin the world. UNAIDSoll A, O niel G, Peckham CS *et al* (2000). The public health applications of unlinked anonymous sero-prevalence monitoring of HIV in the United Kingdom. International *Journ Epi.* 29: 1-10.

Quinn TC. Global burden of the HIV pandemic (1996). *Lancet.* Jul 13; 348(9020): 99-106.

Ravallion,M & G.Datt (1996) Is Targeting through a Work Requirement Efficient? Some Evidence for Rural India. In: D. van der Walle & K.Nead (eds) *Public Spending and the Poor: Theory and Evidence.* The Johns Hopkins University Press. Baltimore, Maryland.

Shaikh & Abdullah (2002). The Provincial and District Survey Report; Western Cape.

Shisana O, Simbayi L (2002). Nelson Mandela /HSRC study of HIV/AIDS .*South African National HIV prevalence, Behavioural Risks and Mass Media.* Human Sciences Research Council. ISBN 0-7969-2007-9; 58-59.

World Health Organisation. HIV and AIDS (2000). *WHO report on Global Surveillance of Epidemic – Prone Infectious Diseases.* Pg 18?

Word Health Organization (1989). *Unlinked anonymous screening for public health surveillance of HIV infections. Proposed International guidelines. Global Programme on AIDS* , GPA/SFI/89.3:1-.

Topouzis D (1998) *The Implications of HIV/AIDS for Rural Development Policy and Programming: Focus on Sub-Saharan Africa.* Sustainable Development Department, Faohiv And Development Programme, UNDP

Appendix

Figure 6. Cumulative Distribution of Household Size by Imputed HIV Status

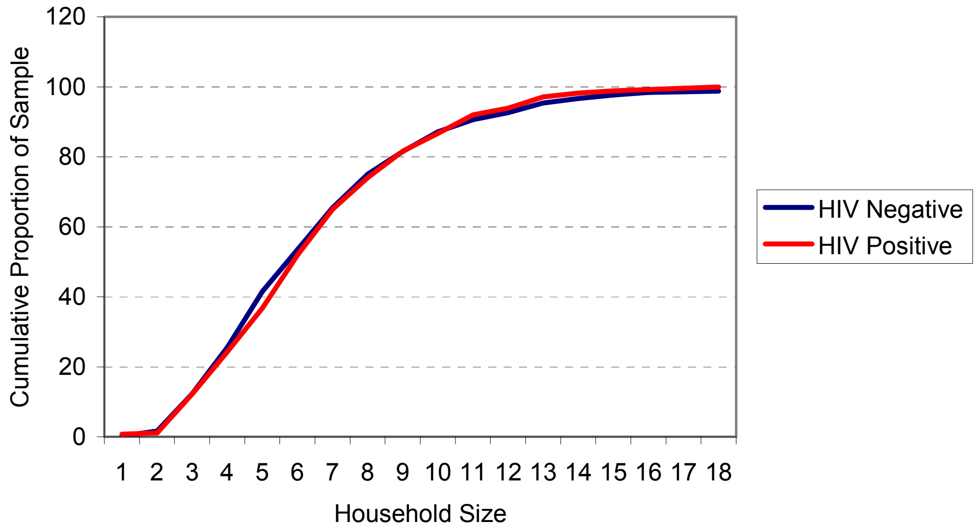


Figure 7. Cumulative Distribution of Number of Children by Imputed HIV Status

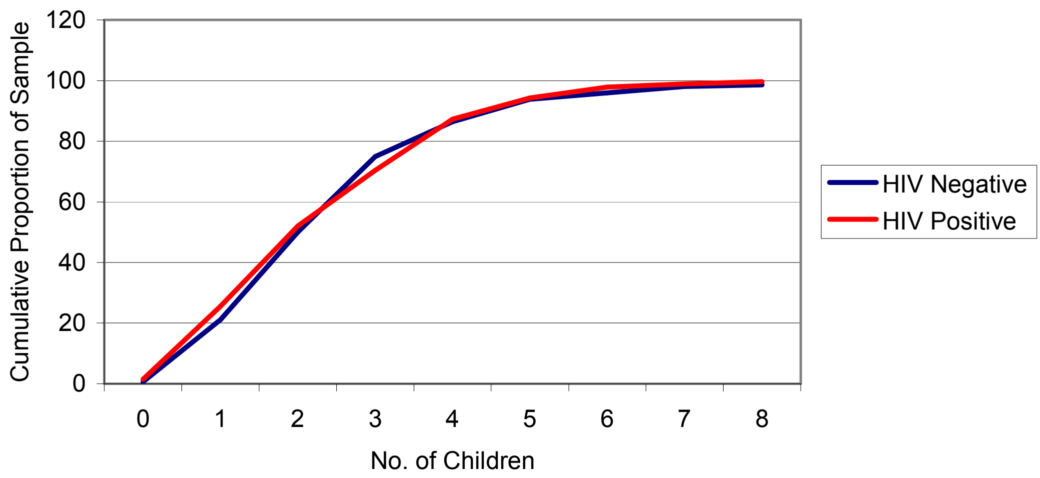


Table 24: Headcount and Poverty Gap Measures for Imputed HIV-Status Households With Zero Wage Earners [16]

Poverty Measure	Estimate	Std. Error	Confidence Interval	
HIV Negative				
Headcount Index	0.65	0.022	0.6039	0.6887
Poverty Gap	0.41	0.016	0.3819	0.4449
HIV Positive				
Headcount Index	0.67	0.044	0.5800	0.7534
Poverty Gap	0.44	0.033	0.3704	0.5005
t-Statistics	**Headcount**	-0.42	**Poverty Gap**	-0.60

Note: Poverty Line z, is set at 1499.5 per month per household.

[16] The *t*-statistic is calculated according to the formula $t = \frac{P_A - P_B}{s}$ *where* $s = \sqrt{\mathrm{var}(P_A) + \mathrm{var}(P_B)}$. Further details can be found in Ravallion & Datt (1996) and Kakwani (1993).

Table 25: Headcount and Poverty Gap Measures for Imputed HIV-Status Households With At Least One Wage Earner, *z= 2149.5*

Poverty Measure	Estimate	Std. Error	Confidence Interval	
		HIV Negative		
Headcount Index	0.80	0.016	0.7727	0.8357
Poverty Gap	0.55	0.014	0.5203	0.5735
		HIV Positive		
Headcount Index	0.88	0.026	0.8232	0.9268
Poverty Gap	0.58	0.023	0.5379	0.6304
t-Statistics	**Headcount**	-2.30*	**Poverty Gap**	-1.38

Note: Poverty Line z, is set at 2149.5 per month per household.

*: Significant the 5% level

Table 26: Headcount and Poverty Gap Measures for Imputed HIV-Status Households With Zero Wage Earners, *z= 2149.5*

Poverty Measure	Estimate	Std. Error	Confidence Interval	
		HIV Negative		
Headcount Index	0.76	0.019	0.7266	0.8019
Poverty Gap	0.52	0.016	0.4889	0.5501
		HIV Positive		
Headcount Index	0.79	0.037	0.7206	0.8691
Poverty Gap	0.54	0.031	0.4822	0.6061
t-Statistics	**Headcount**	-0.73	**Poverty Gap**	-0.71

Note: Poverty Line z, is set at 2149.5 per month per household.

Figure 8: Distribution of Number of Unemployed by Imputed HIV Status Households

Figure 9. Ratio of Weekly Wages: Imputed Negative to Imputed Positive Women