Hiv/Aids And Poverty In South Africa: A Bayesian Estimation Of Selection Models With Correlated Fixed-Effects

Fabrice Murtin, Federica Marzo

HIV/AIDS AND POVERTY IN SOUTH AFRICA: A BAYESIAN ESTIMATION OF SELECTION MODELS WITH CORRELATED FIXED-EFFECTS

FABRICE MURTIN* AND FEDERICA MARZO†

Abstract

In this paper, we estimate the causal impact of human immunodeficiency virus/acquired immune deficiency syndrome (HIV/AIDS) on monetary poverty using a panel database from South Africa. We treat endogeneity and selection problems associated with HIV/AIDS by a selection model that includes correlated fixed effects both in the level and in the participation equations, which are estimated simultaneously via original Bayesian methods. We model the consequences of the illness on both labour income and income transfers, and disentangle between urban and rural households. While no significant impact of HIV/AIDS on labour income is found because of households' recomposition, we find a substantial fall in received transfers among rural population and a dramatic increase in chronic poverty.

JEL Classification: C11, C34, I14, J22
Keywords: HIV/AIDS, poverty, South Africa, Bayesian estimation, Monte Carlo Markov Chain, hybrid Gibbs sampling

1. INTRODUCTION

According to the latest Joint United Nations Programme on HIV/AIDS or UNAIDS estimations, at the end of 2007, as many as 32.9 million people in the world were living with HIV/AIDS, 95% of whom in developing countries and more than 67% in sub-Saharan Africa. Southern Africa remains the epicentre of the pandemic; with seven out of the 10 countries of the subregion having a prevalence rate above 15% (UNAIDS, 2009). In those countries, the epidemic has had a dramatic impact on demographic statistics, with life expectancy falling back to levels of the 1950s according to the United Nation Population Division (UNPD, 2008).

* Corresponding author: Economist, OECD Economics Department, 2 rue André Pascal, 75775 Paris Cedex 16, France. E-mail: Fabrice.Murtin@oecd.org
† Economist, The World Bank. E-mail: fmarzo@worldbank.org.
We thank Philippe Aghion, Jean-Claude Berthélémy, Frikkie Booysen, David Card, Thomas Dickinson, Marcel Fafchamps, Thierry Kamionka, Francis Kramarz, Murray Leibbrandt, Jeremy Magruder, Ted Miguel and Luigi Pistaferri for the fruitful discussions and helpful comments, as well as the seminar participants at CREST(INSEE), AES conference 2007 in Cape Town and UC Berkeley. Murtin acknowledges financial support from the Mellon Foundation when he was hosted by the Stanford Centre for the Study of Poverty and Inequality. The paper does not necessarily represent the views of the The paper does not necessarily represent the views of the Organisation for Economic Co-operation and Development or The World Bank and their member countries.

While conclusions on the impacts of the illness on gross domestic product (GDP) per capita remain mixed at the macroeconomic level,[1] a number of negative economic externalities have been established at the microeconomic level. In the context of the AIDS epidemic, a series of studies have highlighted the impact of adult deaths on consumption levels (Chapoto and Jayne, 2008), consumption growth (Beegle *et al.*, 2008), activities and income sources (Yamano and Jayne, 2004) and time reallocation (Beegle, 2005). Other well-established negative impacts of HIV/AIDS include a decrease in savings and investments, the reduction in school enrolment and teaching staff (Mutangadura *et al.*, 1999) and the weakening of family and community solidarity structures (Nyblade and Field-Nguer, 2001; Bond, 2002). In reaction to the illness, households are known to adopt various coping strategies to absorb income shocks, as described by Lundberg *et al.* (2000). Solidarity networks of relatives and neighbours emerge and may take the form of financial support, sharing of meals, fields or cattle, providing some labour or hosting new active persons in the household. As a result, individual and collective coping strategies serve to mitigate some of the short-term negative consequences of HIV/AIDS and generate a complex articulation between the incidence of the illness and various resource flows.

The aim of our paper is to assess the impact of HIV/AIDS on household income while accounting for its non-linear dynamics and the interaction between its sub-components. Using a panel of six waves spanning over three years as well as Bayesian methods, we estimate the impact of HIV/AIDS on labour income and non-labour income, which includes public grants and remittances. Such decomposition enables a finer understanding of the economic consequences of the illness, as each income component may be affected differently.

For each income source, we build a selection model that includes fixed effects, both in the income level equation and in the labour market participation equation. A valuable contribution of our methodology is to simultaneously solve the problems associated with time-constant unobserved heterogeneity and selection. First, the longitudinal dimension of our data allows us to take into account time-constant unobservable factors that may affect both household living standards and the likelihood of being affected by the virus. Second, our methodology allows us to address the selection problem because *correlated* fixed effects are introduced both in the income level equation and in the participation equation; thus, the non-null correlation between the two dimensions of fixed effects captures the correlation between unobserved variables that drives the usual selection bias.[2] Importantly, the estimation of correlated fixed effects makes the use of any instrument unnecessary. We believe that this original framework to be a useful alternative to instrumentation procedures, as long as panel data are available, and there is no obvious instrumental strategy.

However, we acknowledge important data limitations in our application case as the proportion of individuals experiencing a change in HIV status is reduced to 10% of the

[1] It is indeed unclear whether the illness would have a negative impact on human capital formation and GDP per capita (Corrigan *et al.*, 2005; McDonald and Roberts, 2006) or a positive net impact on GDP per capita channelled by reduced fertility (Young, 2005). As reported by Aghion *et al.* (2011), empirical evidence based on growth regressions unambiguously support the view that reduced life expectancy would have a negative impact on income growth.

[2] Another reason to model participation is the non-linearity of income dynamics induced by decisions to participate or not; the sharp decrease in income when individuals stop participating to the labour market is a key issue in the analysis of poverty implications of the illness.

original sample. Despite a good performance of the algorithm with such transition rates, identification of HIV/AIDS effects appears to be somewhat weak and subject to caution. Moreover, we do not address time-varying omitted factors linked to the illness, which may still create endogeneity bias.

The estimation of our model via simulated maximum likelihood might be difficult because of the presence of multiple sources of correlation (Hyslop, 1999) as well as censoring in the data. So we turn to Bayesian methods that enable the estimation of complex econometric models.[3] In this paper, we use a modified version of the Gibbs sampling algorithm introduced by Nobile (1998), called the hybrid Gibbs sampling; the idea being to combine the two building blocks of Bayesian econometrics, the Metropolis-Hastings and the Gibbs algorithms, in order to explicitly model the correlation between fixed effects and observable variables. This dramatically increases the speed of convergence of the classical Gibbs sampling algorithm in the context of endogenous explanatory variables.

The first surprising result of our estimation is that HIV/AIDS has no significant impact on household labour income, neither for urban nor for rural populations. This suggests that households manage to smooth labour income and participation at least in the short term, despite the negative shock of the illness. This finding underlines the importance of controlling for fixed effects; the observed negative correlation between the illness and household labour income is linked to other observed or unobserved determinants correlated with AIDS and not to the illness itself. However, at least two important caveats need to be lodged: first, the small sample size and the 10% HIV status transition rate are unlikely to yield strong identification of the HIV effect; then, the short duration of the survey is another important limitation as the impact of the illness becomes stronger over time. So it is important to flag that our conclusion pertains only to short-term effects of the disease and cannot be held representative of the situation at the national level. In other words, it could well be that a longer and larger survey identifies an impact of HIV/AIDS on household labour income.

Our second and main result shows a significant and asymmetric effect of the illness on income transfers among rural and urban settings. We find that the illness brings about an increase in participation in the transfer network for the urban population, attributable to public grants. On the other hand, in rural areas, AIDS seems to cause a sharp decline in the level of private transfers, increasing chronic poverty by as much as 50%. Plausibly, this effect could be linked to stigma and social discrimination associated with the illness as already noticed by other studies (*e.g.* Seeley, 2002).

The paper is organised as follows: the first section describes the existing literature on the economic consequences of AIDS. Then we introduce the econometric framework, test it on simulated data having similar statistical characteristics to our sample (*e.g.* mean, variance and within groups variance) and compare it with traditional estimators. In the third section, we describe the data, setting and results. The final section concludes.

2. HIV/AIDS AND POVERTY IN THE LITERATURE

The channels transmitting the economic shock of HIV are complex and multiple. A key contribution of this study is to disentangle the direct impact passing through labour

[3] It is generally simpler to simulate a complex distribution via Monte Carlo Markov Chains methods than finding its mode via maximisation algorithms.

income (representing productivity and labour participation effects) from possible indirect impacts channelled by transfers (public grants and private remittances) that may be the result of collective coping strategies. In this section, we review existing studies that address the economic consequences of the illness on both channels.

Whereas poverty is not necessarily a cause for spread of the epidemics,[4] several studies report that morbidity and mortality related to the illness are associated with dramatic rises in poverty in several Southern African countries, such as South Africa (Oni *et al.*, 2002; Booysen, 2004; Jayne *et al.*, 2005) or Malawi (Dorward and Mwale, 2005).

The channels through which the illness affects household consumption are three-fold. First, labour income falls as the production of the sick and often of their entourage decreases. Both hours worked and productivity decline sharply before a worker dies, and as a consequence, household labour income may drop by up to two thirds of mean income (Morris *et al.*, 2000; Loewenson *et al.*, 2007; Munthali, 2002).

Second, expenditures in social services, medical assistance and funerals increase. Steinberg *et al.* (2002) show that in South Africa affected households spend about one third of their income on health care, compared with a national average of 4%. The high cost of funerals has become a real threat for the economic security of households in South Africa as shown by Ayieko (1997) and can reach up to 40,000 ZAR.[5]

Third, the composition of the household often changes, with old persons coming in to help and new active people joining the household. As a consequence, the evolution of the dependency ratio is not clear (Mutangadura *et al.*, 1999; Rehle and Shisana, 2003; Epstein, 2004). Orphans are often hosted by extended families when one or both parents die. According to several Demographic and Health Surveys projects conducted in the early years of the decade on household composition, one fifth to one quarter of households in high prevalence African countries are fostering children (see Demographic and Health Surveys, 2000-2006).

While these coping strategies are centred around providing immediate relief, they may hold negative effects in the long term. For instance, the fall in labour income and the contemporary increase in health expenditure imply that households redistribute resources and time in favour of the persons living with HIV/AIDS, entailing a cut in consumption of basic goods and, possibly, malnutrition for other members of the household (Ainsworth and Dayton, 2003). Moreover, there is an evidence of dissaving or of asset sales, such as cattle and livestock, furniture and work instruments (Munthali, 2002). With production capacity weakened and savings exhausted, in the longer term consumption may decrease further. Human capital may be adversely affected because of children being withdrawn from school to look after the ill (Ainsworth *et al.*, 2002).[6]

Coping strategies also refer to external, collective actors such as the public sector, the surrounding community or extended family responding to the illness. Public transfers generally take the form of disability grants or orphans allowances, and have a heterogeneous

---

[4] Lachaud (2007) notably reports a positive correlation between HIV prevalence and household wealth; Sahn and Stifel (2003) find a positive correlation between education and AIDS incidence; while De Walque (2006) finds no significant association between prevalence and education.

[5] Over 5,700 US dollar (2007 exchange rate).

[6] However, Coombe (2002) suggests that the impact of the epidemic on school attendance is difficult to estimate accurately, as the reasons why children are withdrawn from school are usually unknown.

impact on poverty.[7] Private transfers composed of grants and remittances/gifts as a form of collective coping strategy may also constitute an important monetary support[8] and can make a difference with respect to the final impact of AIDS on households' livelihood. However, several studies report that traditional safety networks suffer severe stress when HIV/AIDS-related illnesses and mortality increase (Kawachi *et al.*, 1997, 1999; Kunitz, 2001; Seeley, 2002).

3. THE ECONOMETRIC FRAMEWORK

This section presents the model, the estimation algorithm and illustrates its benefits with respect to traditional panel models. The selection model displays multidimensional fixed effects, which first address the concerns stemming from unobserved heterogeneity. It also solves the selection problem because fixed effects are jointly introduced in the income level equation and in the participation equation with a non-null correlation that controls for the selection bias.

3.1 A Gaussian Model of Selection

We denote $e_{i,t}$ for the participation dummy and $y_{i,t}$ for income. Latent variables $e^*_{i,t}$ and $y^*_{i,t}$ correspond respectively to the unobserved propensity to participate and to individual productivity. The selection model is a system of two equations assuming Gaussian residuals[9]

$$\begin{aligned} & y_i^* | \beta^{(1)}, b_i^{(1)}, D^{(1)}, \sigma^2 \rightsquigarrow \left(X_i^{(1)} \beta^{(1)} + b_i^{(1)} \cdot i_T, \sigma^2 I_T \right) \\ & e_i^* | \beta^{(2)}, b_i^{(2)}, D^{(2)}, \sigma^2 \rightsquigarrow \left(X_i^{(2)} \beta^{(2)} + b_i^{(2)} \cdot i_T, I_T \right) \\ & \forall t, \quad e_{i,t} = I[e_{i,t}^* > 0], \quad y_{i,t} = e_{i,t} \cdot y_{i,t}^* \end{aligned} \tag{1}$$

where $D^{(j)}$ is the variance of fixed effects $b_i^{(j)}$, iT *is* a column vector of size T with all elements equal to 1, and I_T *is* the identity matrix. We use a Bayesian framework and focus on the distribution of parameters of interest $\left(\beta^{(j)}, b_i^{(j)}, D^{(j)}, \sigma^{2(j)} \right)_{j \in \{1,2\}}$. The former system can be written as a linear Gaussian panel model

$$Y_i^* | \beta, b_i, D, \sigma^2 = \mathcal{N} \left(X_i \beta + b_i \otimes i_T, \Sigma \right) \quad i \leq N, \quad t \leq T$$

$$Y_i^* = (y_{i,1}^*, \ldots, y_{i,T}^*, e_{i,1}^*, \ldots, e_{i,T}^*)$$

---

[7] Naidu (2004) shows that a large part of the income shock caused by the illness is absorbed by public grants in Soweto: The fall in income for affected household approaches 30%, but after public transfers are accounted for, it is reduced to 8%. However, Webb (1995) shows that in Southern Africa lower income communities are more likely to assume all of the cost related to the consequences of the illness (like hosting orphans), whereas high-income communities benefit from having the government as primary caregiver.

[8] See Case and Deaton (1998) on this issue.

[9] Conditional normality of the dependent variable corresponds to the log-normality of income variables in the case study presented later.

$$X_i = \begin{bmatrix} X_i^{(1)} & 0 \\ 0 & X_i^{(2)} \end{bmatrix}$$

$$\beta = [\beta^{(1)\prime} \beta^{(2)\prime}]'$$

$$b_i = [b_i^{(1)} b_i^{(2)}]'$$

$$D = \begin{bmatrix} D^{(1)} & D^{(1,2)} \\ D^{(1,2)} & D^{(2)} \end{bmatrix}$$

where $\otimes$ is the Kronecker product. There are two major issues arising in this context: the correlation structure of the model and the missing data because the dependant variable is partly observed or completely unobserved as in the case of the latent variable $e_{i,t}^*$.

3.2 The Correlation Structure

The first issue deals with endogeneity. In order to ease simulations, we assume that the conditional distributions of $y_{i,t}^*$ and $e_{i,t}^*$ are independent, in other words that the idiosyncratic residuals of each equation are non-correlated. This is reflected by non-diagonal terms of Σ set equal to zero. However, fixed effects can be correlated across the two equations, so that idiosyncratic shocks affecting wages and participation are non-correlated, but permanent shocks can be.[10]

Moreover, we would like to account for endogeneity of the observed variables, so that fixed effects have zero mean but not necessarily conditional zero mean (conditionally on observed variables). In short, $E[b_i \mid X_i] \neq 0$. Importantly, the correlation between fixed effects and endogenous variables shall be modelled if we want the Gibbs sampling algorithm to converge rapidly. The simplest solution is to assume that fixed effects are an index of the individual means of the endogenous variables plus a non-correlated component, as in Chamberlain (1984).[11] More precisely, one can decompose the vector of specific effects in the following way.

We note

$$\bar{X}_i = \begin{bmatrix} \bar{X}_i^{(1)} & 0 \\ 0 & \bar{X}_i^{(2)} \end{bmatrix}$$

$$\bar{X}_i^* = \bar{X}_i - \bar{X}$$

[10] Theoretically, it would be possible to allow for both sources of correlation, but the estimation would behave poorly, unless working with a large time dimension T.

[11] Note that even in the case of non-time varying regressors, the model remains identified because of the specification of prior distributions on each parameter. But identification is often weak in that case with vague priors, and convergence can be considerably slowed down. However, this is not a problem for us because all of our explanatory variables, including our main endogenous variable, a dummy for HIV/AIDS status, are time varying.

where $\bar{X}_i^{(j\in\{1,2\})}$ is the 1 by K vector of individual means of $X_i^{(j)}$ in equation (j) and $\bar{X}$ the 2 by $2K$ matrix composed of the overall mean of explanatory variables. We retain the following specification:

$$b_i = \bar{X}_i^* \cdot \lambda + \varepsilon_i \tag{2}$$

with $\lambda = [\lambda_1^{(1)} \ldots \lambda_K^{(1)} \lambda_K^{(2)} \ldots \lambda_K^{(2)}]'$, a 2K vector of scalars and $\varepsilon_i = [\varepsilon_i^{(1)} \varepsilon_i^{(2)})]'$ additional individual effects. Then,

$$\sum_i \bar{X}_i^{*\prime} b_i = \left(\sum_i \bar{X}_i^{*\prime} \bar{X}_i^*\right)\lambda + \sum_i \bar{X}_i^{*\prime} \varepsilon_i \tag{3}$$

With $E(\varepsilon_i \,|\, X_i^*) = 0$ and by decomposing the covariance of regressors and fixed effects, one obtains an estimate of λ:

$$\hat{\lambda} = N\left(\sum_i \bar{X}_i^{*\prime} \bar{X}_i^*\right)^{-1} \rho \circ \begin{pmatrix} sd\left(\bar{X}^{*\prime(1)}\right) \\ sd\left(\bar{X}^{*\prime(2)}\right) \end{pmatrix} \circ \begin{pmatrix} \sqrt{D^{(1)}} \\ \sqrt{D^{(2)}} \end{pmatrix} \text{ with } \varepsilon_i \perp \bar{X}_i^*$$

where ρ is the $2K$ vector of correlations between specific effects and individual means of regressors, $sd\left(\bar{X}^{*(j)}\right)_{j\in\{1,2\}}$ is the K vector containing the standard error across individuals of individual means of regressors in equation (j), $D^{(j)}$ is the variance of specific effects in equation (j), and ° is the element-by-element product. This expression is at the core of the Bayesian algorithm because one can derive fixed effects *bi* from the vector of correlations ρ and from fixed effects' variances $D^{(j)}$.

3.3 Missing Data

The second difficulty arises from the missing data, namely that $e_{i,t}^*$ is unobserved, as well as $y_{i,t}^*$ when $e_{i,t}^* < 0$. A powerful feature of the Bayesian approach is that the missing data can be treated just as other parameters of interest; they are simulated. Indeed, given the set of parameters Θ, the density of Y_i^* can be decomposed with Bayes' rule

$$\begin{aligned} f(Y_i^*\Theta) &= \prod_t f(y_{i,t}, e_{i,t}^* | \Theta, e_{i,t}^* > 0) \prod_t f(y_{i,}^*, e_{i,t}^* | \Theta, e_{i,t}^* \le 0) \\ &\propto \prod_t f(y_{i,t}, e_{i,t}^* | \Theta) f(e_{i,t}^* > 0 | \Theta, y_{i,t}, e_{i,t}^*) \prod_t f(y_{i,}^*, e_{i,t}^* | \Theta) f(e_{i,t}^* \le 0 | \Theta, y_{i,}^*, e_{i,t}^*) \\ &= \prod_t f(y_{i,t}, e_{i,t}^* | \Theta) 1_{e_{i,t}^* > 0} \prod_t f(y_{i,}^*, e_{i,t}^* | \Theta) 1_{e_{i,t}^* \le 0} \end{aligned}$$

Consequently, when $y_{i,t}$ is observed, the data augmentation step consists in drawing $e_{i,t}^*$ from its posterior distribution, namely a truncated normal distribution taking values upon the interval]0, +∞[. When $y_{i,t}$ is censored, ones draws the couple ($y_{i,t}^*$, $e_{i,t}^*$) from a bivariate normal variable truncated on the interval]−∞, 0[for the second component ($e_{i,t}^*$).

3.4 Estimation

Let us now describe the algorithm. The Bayesian algorithm aims to infer the conditional distribution $p(\Theta|Y)$ where $\Theta = (\beta, \rho, \varepsilon_i, D_\varepsilon, \sigma^2)$ are the parameters of interest, and Y is the

data. This distribution is proportional to the posterior distribution $p(Y|\Theta)\, p(\Theta)$ by Bayes rule. Some prior distributions $p(\Theta)$ are set on parameters and, for usual Gaussian panel models, priors and the sampling distribution (the likelihood) $p(Y|\Theta)$ are chosen from the same exponential family so that their product rearrange in closed form; the posterior distribution of each parameter has an explicit formulation. Importantly, the choice of a given prior does not imply any loss in generality as the variance of this prior distribution can be taken as large as desirable to avoid constraining the estimation.

Inference is achieved with a hybrid version of the Gibbs sampling algorithm. The Gibbs sampling algorithm is an iterative approach that draws from the conditional posterior distribution of each block of parameters[12] conditionally on former drawings of other blocks of parameters. This algorithm constitutes a Markov chain that converges towards the stationary distribution of parameters under fairly general conditions described by Tierney (1994). As the posterior distribution of the correlation ρ cannot be written in closed form, we simulate it using a Metropolis–Hasting step, which is the origin of the term "hybrid" Gibbs sampling. In a different context, this hybrid approach has first been introduced by Nobile (1998), and it is extensively described by Casella and Robert (2004). Priors and the detailed algorithm are fully described in annex.

3.5 Testing

We test this algorithm on a simulated dataset and show that the coefficients of all endogenous variables are perfectly estimated. For this test, 50,000 iterations of the hybrid Gibbs sampling were used. The model accounts for both specific effects and time effects, which are time dummies included into the set of regressors.

Formally, we simulate

$$\begin{aligned} y_{i,t}^{*} &= \mu^{(1)} + \delta_i^{(1)} + b_i^{(1)} + \beta^{(1)} X_{i,t} + \sigma u_{i,t} \\ e_i^{*} &= \mu^{(2)} + \delta_i^{(2)} + b_i^{(2)} + \beta^{(2)} X_{i,t} + v_{i,t} \\ \forall t, \quad e_{i,t} &= I[e_{i,t}^{*} > 0], \quad y_{i,t} = e_{i,t} \cdot y_{i,t}^{*} \\ b_i^{(j)} &= \nu_i + \varepsilon_i^{(j)}, \nu_i \,|\varepsilon_i^{(j)} \forall j \end{aligned} \tag{4}$$

The endogenous variable X is specified as a dummy variable that takes value one if $v_i > 0$ and 0 otherwise. Moreover, it is time varying as we allow some transitions from 0 to 1 for 10% of the population satisfying $X_{i,1} = 0$. Those transitions take place at a random date and are permanent. Hence, this endogenous variable replicates the statistical characteristics of the HIV/AIDS dummy variable observed from the data, it has an impact both on income level and on participation, and it is correlated to fixed effects $b_i^{(j)}$ via the time-constant component ν_i. As a result, the percentage of censored observations (those for which $e_i^{*} < 0$) is equal to 24% among the "non-affected" population, namely those for which $X_{i,t} = 0$, and 34% for the others. Again, these figures match the data related to labour income.

Table 1 presents the results for three different estimators: the Tobit random coefficients model, a fixed effects model applied to non-censored observations and the hybrid Gibbs sampling described earlier. As a result, it is clear that the Tobit model delivers poor

[12] In this context, the five blocks corresponding to β, ρ, ε_i, D_ε, σ^2.

Table 1. Test of the hybrid Gibbs sampling

	$\beta^{(1)}$	$\beta^{(2)}$	$\rho_{b_i^{(1)},x}$	$\rho_{b_i^{(2)},x}$	s^2	$D^{1,1}$	$D^{1,2}$	$D^{2,2}$
True values	−1	−1	0.54	0.52	1	1	0.5	1
Tobit estimates	−1.25	–	0	–	6.30	3.88	–	–
	(0.24)							
Fixed effects estimates[1]	−0.89	–	0.38	–	0.81	0.64	–	–
	(0.12)							
Hybrid Gibbs	−1.01	−1.039	0.56	0.56	0.97	0.99	0.43	1.18
	(0.10)	(0.15)	(0.05)	(0.05)	(0.04)	(0.13)	(0.15)	(0.26)

[1] On participating population only.

estimates of the income level equation, with a 25% downward bias on the coefficient of interest. This was expected because of the high number of censored data (with observed outcome zero) and because of the tight constraints set upon the functional form of the dependent variable in this model. Moreover, with a positive correlation between fixed effects and the endogenous variable, the magnitude of the effect is overestimated, entailing a downward bias because the coefficient is negative.

Fixed effects estimates produce a smaller bias because endogeneity is taken into account for the non-censored population; in that case, the coefficient is only overestimated by 11%. Even if that point estimate is not significantly different from the true value, this bias will contaminate poverty simulations. Again, the direction of the bias was expected because the correlation between the endogenous variable and unobserved determinants of participation was taken positive.

By contrast, the hybrid Gibbs sampling produces point estimates that are very close to their respective true value. Therefore, we obtain valid estimates of the impact of HIV/AIDS on income level and participation, which will be at the core of the poverty simulations in last section. Importantly, these estimates are robust to endogeneity bias and capture selection effects correctly.[13] Fig. 1 describes the successive drawings from the algorithm for the coefficients of interests and the correlations between fixed effects and endogenous variables.

4. DATA

4.1 The Survey

The impact of HIV/AIDS on poverty is studied using a panel composed of affected and non-affected households. A survey on households' quality of life and resources was conducted every six months in two districts belonging to the Free State province.[14] The first four rounds of interviews were completed in May/June and November/December of 2001 and in July/August and November/December of 2002. Rounds five and six of the

[13] It is worth underlining that this approach does not rely on any instrument in the participation equation. Instrumenting participation is often useful in cross-section regressions because identification of the correcting term, the Mills ratio, is weak though theoretically achieved. In our case, the algorithm performs well without any instrument, but it could be possible that with smaller time-dimension and low levels of within-variance, instrumentation becomes useful.

[14] Households were defined using the standard definition employed by Statistics South Africa in the October Household Survey, *i.e.* "a person or a group of persons who live together at least four nights a week" (Statistics South Africa, 2000:0317-E) and who share resources. Interviews were conducted with one key respondent only, namely the "person responsible for the daily organisation of the household, including household finances."

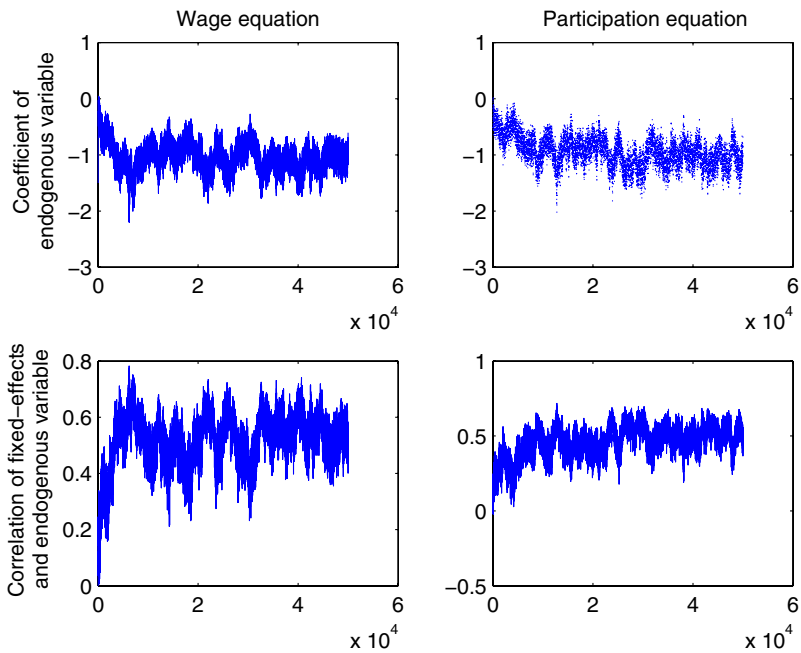

Figure 1. Convergence of the hybrid Gibbs sampling on simulated data

study were completed in July/August 2003 and May/June 2004, respectively. Thus, data span a period of over three years (see Booysen, 2004) for a detailed description of the survey and sampling procedure).

The balanced survey is composed of 332 households and 1,173 individuals with data available at each wave. There is a large heterogeneity between rural and urban households. Urban and rural populations are of equal size (around 166 households), as well as affected and non-affected households initially. Households are defined as affected if, at the time of the interview, someone belonging to it has declared being HIV positive. These households have been selected through non-governmental organizations and public services working in the field of HIV/AIDS. Informed consent prior to the utilisation of the data has been given by the concerned people or their relatives.[15]

Within the urban and rural populations, an equally sized comparison group of people identified as unaffected by HIV/AIDS at baseline was interviewed on a voluntary basis. They were meant to have similar characteristics to the affected households, thanks to the selection process.[16] Importantly, the classification of affected and non-affected has

[15] HIV-positive people who accepted to participate in the study have not necessarily informed their family about their serostatus, and the survey respondent is not necessarily the person recognised as being HIV positive. Moreover, we do not know at what time households have been affected, with the exception of those whose serostatus changed at one point during the three-year span of the survey.

[16] For each affected household successfully interviewed, the fieldworker chose randomly a neighbouring household living in close proximity to the affected household. In order to ensure that this household was at that time not directly affected by HIV/AIDS, the fieldworker asked to the respondent some key questions, namely whether someone in the household has being treated for

been revised wave after wave; households who experienced AIDS-related illness or death over subsequent waves were reclassified as "newly affected." This group is made of 33 households, about 10% of the original sample. Admittedly, the share of households is relatively low, which may seem problematic as identification of HIV/AIDS effects is based on the size of the latter group. However, it is important to stress that former simulations demonstrated the capacity of the Bayesian algorithm to assess HIV/AIDS effects under the same conditions (*i.e.* a small sample and a low transition rate).

The second important feature of the survey is the choice of the settings. According to Statistics South Africa (2000), the Welkom magisterial district, situated in the Goldfields, is the third richest in the Free State province. It can be defined as an urban setting.[17] By contrast, the rural magisterial district of Witsieshoek, which is within the boundaries of the former homeland of Qwaqwa, is the poorest in the Free State province and is ranked among the poorest in the country, with very little infrastructure and few social services.

It is important to note that the findings from this study cannot be generalised to households across South Africa because of the small sample size, a feature shared by most other HIV/AIDS impact studies and the context-specific location of the survey. The results reported in these pages are more representative of the Free State province, which has the second highest prevalence of HIV/AIDS and is also the province with the second highest prevalence growth rate (Cohen, 2002).

4.2 Descriptive Statistics

The original sample is composed of 167 urban and 166 rural households. As shown by Table 2, there are important differences between urban and rural populations. On

Table 2. Descriptive statistics (waves 1 to 6)

	Urban			**Rural**		
	Total	**Affected**	**Non-affected**	**Total**	**Affected**	**Non-affected**
N	167	87	80	165	88	77
Age of head	50.5	51.6	49.4	48.6	48.4	49.0
Education of head	7.6	7.2	7.9	6.7	6.3	7.2
Dependency ratio	0.70	0.78	0.61	0.71	0.68	0.74
Active people	3.3	3.5	3.1	2.6	2.6	2.5
Average labour income[1]	709	593	817	535	363	702
Participation rate (labour income)	63.7	58.8	69.0	49.7	45.7	54.3
Average grants[1]	250	251	249	193	197	189
Participation rate (Grants)	47.2	56.7	36.8	53.2	54.1	52.2
Average transfers[1,2]	278	266	294	238.8	223.6	257
Participation rate (Transfers)	62.6	70.0	54.4	75.2	75.8	74.6
Average total income[3]	625	535	724	445.2	335.5	571.6
Participation rate (Total)	97.2	96.4	98.1	97.7	97.6	97.8
Chronic poverty	22.7	24.0	21.3	33.3	40.2	25.4
Transitory poverty	8.0	8.8	7.1	9.1	10.1	7.8

[1] Computed on individuals who participate.
[2] Social grants plus private transfers.
[3] Computed on the whole population.

Tuberculosis, pneumonia and other diseases linked to AIDS over the past six months. Only those displaying negative answers were retained in the control group.

[17] The distinction between rural and urban setting is based on economic activities and on the mode of governance (traditional vs. modern), rather than on the differences in dwellings equipments and infrastructure endowments.

average, the urban households are more educated than the rural ones, and their income[18] is 40% higher. Income transfers, which include monetary transfers and in-kind transfers, constitute a smaller proportion of household income in urban areas, about 45% vs. 55% in rural areas. Moreover, higher mean income translates into a lower chronic poverty rate:[19] 23% vs. 33% for a 250 ZAR threshold (about 1.2 dollars a day). Unemployment figures over the whole period confirm the vulnerability of the rural population when compared with the urban: The unemployment rate reaches 37.3% over the six waves for the urban population and 50.3% for the rural one (this is consistent with national unemployment figures).

HIV-affected and non-affected households differ substantially in terms of labour income, which is defined as earnings coming from any form of formal or informal working activity, including waged activities and subsistence agriculture.

For the urban population, affected households' total income is 20% lower, labour income about 27% lower relative to non-affected households, and their unemployment rate is 10% higher. They rely proportionally more on non-labour income (defined as all income components not belonging to working activities), namely all public grants as well as private remittances and in-kind gifts. These results strengthen the evidence brought by Barnett and Joint United Nations Programme on HIV/AIDS (2002) that state and community support can make a difference for urban households affected by HIV.[20]

For the rural population, the unemployment rate is 10% higher among the affected households, and labour income is 50% lower. Overall, total income is 40% lower among the affected rural households. As an illustration, Fig. 2 reports the cumulative density functions of labour income; of labour income and private transfers; of labour income, private and public transfers, for both urban and rural population. According to these graphs, public social grants are the most important component of non-labour income for the urban population, and public and private transfers are roughly equal in share for rural households.

---

[18] For total income and its components, we have used real figures in adult equivalent terms: $\frac{y_i}{n^{0.6}}$, where y is income, n is the total number of household members and 0.6 is the adult equivalent coefficient.

[19] Chronic poverty is defined as in Jalan and Ravallion (2000). Let y_i represent total simulated income, $\overline{y}_i$ represent its average, z represent the poverty threshold, and P^α be the Foster-Greer-Thorbecke class of poverty measures. Total poverty of household i is then the expectation over time $P_i^\alpha = E\left[\left(\frac{z - y_{i,t}}{z}\right)^\alpha 1_{y_{y,t}<z}\right]$. Then, we define chronic poverty for a household i as $C_i^\alpha = \mathrm{E}\left[\left(\frac{z - \overline{y}_i}{z}\right)^\alpha 1_{y_i<z}\right]$. In practice, we have tested two cases with $\alpha \in 0, 2$ in order to capture not only the prevalence but also the intensity of poverty. As results are always identical, we focus on the headcount measure and report only the statistics obtained with $\alpha = 0$.

[20] In quantitative terms, 70% of urban affected households participate in transfer networks (public or private) vs. 54% for the non-affected. Most of the difference in participation rates is attributable to social grants (public aid such as disability or destitution grants, old pension and child fostering grants), as shown by the participation rate in social grants schemes.

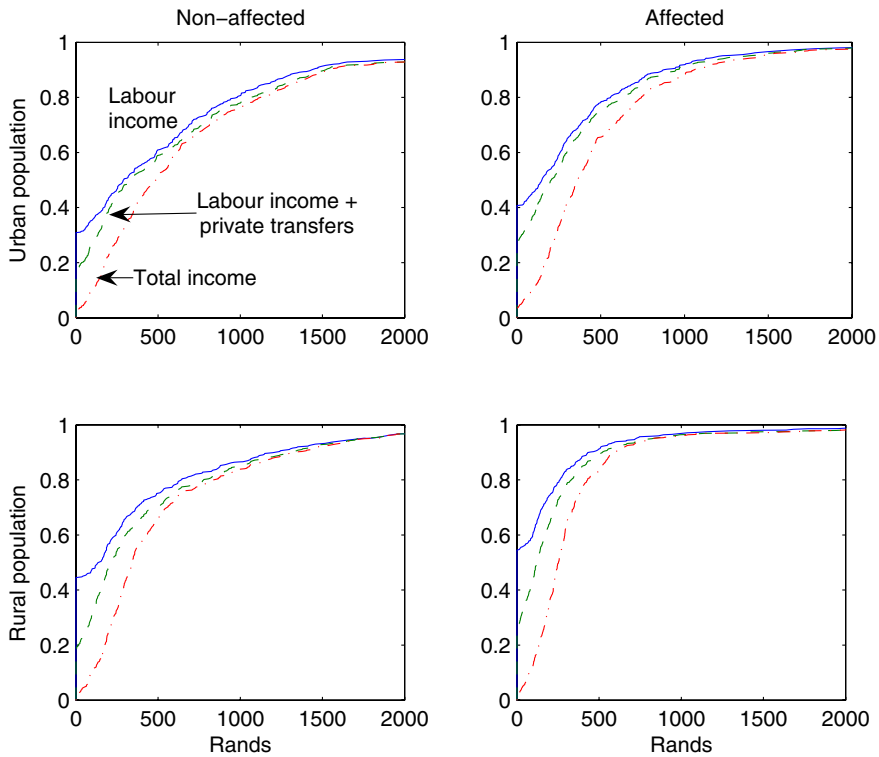

Figure 2. Cumulative distribution functions for affected and non-affected households

Summing up, descriptive statistics highlight first a negative correlation between HIV/AIDS and labour income among both urban and rural households; second, urban households seem to have better access to public support measures, a finding that may suggest a better access to facilities and information, whereas rural households do not seem to benefit from higher income transfers. At this point, these statistics are only cross-sectional correlations; and in the next section, we investigate whether they also reflect causality.

5. ESTIMATION

5.1 A Sequential Estimation

As a starting point, let us address the issue of simultaneity: labour income and non-labour income are simultaneously determined if households are rational and have expectations on both sources of income, so that labour income may enter the non-labour income equation as an explanatory variable and vice-versa. We assume, however, that non-labour income does not affect labour income.

First, in the empirical literature, several authors have made a similar assumption: studying the determinants of remittances and labour income, Maytra and Ray (2003) exclude remittances from the labour income equation. Second, the effect of grants on labour income is not clear cut. Regarding South African data, Jensen (2004) does not find any significant impact of old-age pensions on the household labour supply or

composition. Booysen (2005) finds some ambiguous results, with old-age pension and disability grants being associated with lower employment, while child-fostering support is associated with increased labour supply. Hence, it is difficult to know what would be the total effect when those various grants are pooled together.

Last, the earlier assumption reflects an intuitive economic mechanism: HIV/AIDS represents a negative income shock in first place, whose consequences are tackled through various coping strategies, including extended family or community remittances and public grants. Along this line, most of the non-labour income is intended to complement labour income. Its increase is a consequence of the illness that triggers health-related financial aid or compensation for the decrease in labour income. One could argue that accurate expectations of transfers might have a negative impact on earnings via a substitution effect. However, this view would neglect the fact that most of our sample population live in poverty and would hardly diminish their consumption level.

In practice, we first estimate the selection system for labour income. We derive a latent labour income y_i for each non-working individual as described earlier, we use this variable as a regressor in the non-labour income selection system and run the estimation. In the latter case, we specify two endogenous variables, HIV/AIDS and the latent labour income variable. In particular, fixed effects control for the endogeneity of labour income inside the non-labour income equation. The set of regressors is the same across both the level and participation equation; as previously mentioned, the correlation between fixed effects captures any potential selection bias, and this correlation is identified without any exclusion restriction as soon as a longitudinal dimension is available in the data.

5.2 Results

Results from the Bayesian estimation are reported in the first part of Table 3. For both urban and rural populations, we find that the illness does not have any impact on the level of labour income or on the participation. This is surprising, especially for the rural population, because the descriptive statistics seemed to indicate that affected households had lower levels of income and higher levels of unemployment. This finding reveals that the negative correlation between HIV/AIDS and labour income was spurious and carried by some observed or pre-existing unobserved differences correlated with the incidence of the illness. In particular, among the rural population, the negative correlation between HIV/AIDS and labour income is carried by unobserved heterogeneity.[21] Because the impact of the illness on labour income level is not significant, it means that households adapt somehow to the new situation, increasing labour participation of non-ill persons. We discuss and further investigate this issue in the next subsection.

In a second step, we estimate the non-labour income equations, as illustrated by the second part of Table 3. Among the urban population, the dummy HIV/AIDS is positive and significant in the participation equation of non-labour income, meaning that affected

[21] Ex-ante, the sign of the correlation between unobserved factors and the HIV/AIDS dummy variable is ambiguous. For instance, the propensity to migrate may increase the risk of contamination but might also be a source of wealth through increased information about market or public services. In that case, the sign might be positive. On the contrary, communities might be unequally able to adopt both health technology, such as condoms and productive technologies, driving the negative sign of the correlation.

Table 3. Bayesian estimation of HIV/AIDS impact

	Urban		Rural	
	y^*	e^*	y^*	e^*
Labour income				
HIV/AIDS	−0.113	−0.618	0.247	−0.063
	(0.234)	(0.475)	(0.247)	(0.356)
Dependency ratio	−0.146**	−0.014	−0.147*	−0.449**
	(0.066)	(0.133)	(0.078)	(0.140)
Education of head	0.050	0.093	−0.035	−0.027
	(0.036)	(0.077)	(0.033)	(0.067)
Squared education	0.002	−0.003	0.008**	0.006
	(0.003)	(0.006)	(0.002)	(0.005)
Age of head	0.037**	0.070**	0.014	0.059**
	(0.017)	(0.031)	(0.014)	(0.027)
Squared age	−0.000	−0.001**	−0.000	−0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Female head	−0.346**	−0.641**	−0.359**	−0.403
	(0.090)	(0.190)	(0.101)	(0.199)
ρb_i, *HIV/AIDS*	−0.09	0.09	−0.41**	−0.12
	(0.16)	(0.16)	(0.15)	(0.11)
Non-labour income				
HIV/AIDS	−0.119	0.919*	−0.450**	0.064
	(0.286)	(0.520)	(0.169)	(0.418)
Labour income	−0.104*	−0.068	0.054	−0.150
	(0.057)	(0.100)	(0.043)	(0.108)
Dependency ratio	0.108	0.558**	0.034	0.586**
	(0.069)	(0.151)	(0.050)	(0.179)
Education of head	−0.060	0.005	−0.019	0.003
	(0.041)	(0.079)	(0.027)	(0.077)
Squared education	0.005*	0.000	0.004**	0.001
	(0.003)	(0.006)	(0.002)	(0.006)
Age of head	0.041**	−0.151**	0.018	−0.041
	(0.019)	(0.040)	(0.013)	(0.038)
Squared age	−0.000	0.002**	0.000	0.001*
	(0.000)	(0.000)	(0.000)	(0.000)
Female head	0.243*	0.313*	−0.048	−0.048
	(0.108)	(0.182)	(0.090)	(0.213)
ρb_i, *HIV/AIDS*	−0.02	−0.17	0.33**	−0.01
	(0.18)	(0.21)	(0.11)	(0.14)

Note: Regressions include year dummies.
*, ** indicate significance at a 5/10% confidence level.

households receive more often income transfers. This is consistent with the fact that urban settings have an easier access to social services. Moreover, we observe a substitution effect between labour income and transfers in the sense that the former has a negative impact on the latter, which likely reflects eligibly criteria for social grants.

Among the rural population, the impact of HIV/AIDS on the level of non-labour income is negative and significant, which was not observable from descriptive statistics. Again, this is due to the positive correlation between fixed effects and HIV/AIDS that cancels out with the true negative impact of the illness. This result is important; it suggests that affected rural households are deprived of vital monetary resources because of their illness. What explains this finding is the fall in private rather than public transfers. Indeed, the percentage of newly affected households receiving public transfers increases after the infection from 52% up to 66%. There are two interpretations to the fall in private transfers. One is that the latter effect is a consequence of return migrations, because the people who used to send remittances are the most likely to be affected by HIV/AIDS.[22]

[22] Some basic statistics show that the amount of remittances received by the group of newly affected people decreases slightly from 120 ZAR to 104 ZAR after the illness has been declared.

Table 4. Fixed effects estimates of HIV/AIDS impact on labour income[1] – Participating households

	Urban	Rural	Urban Ill[2]	Rural Ill[2]
Labour income				
HIV/AIDS	–0.009	0.367	0.146	0.393
	(0.312)	(0.230)	(0.356)	(0.252)
Dependency ratio	–0.170**	–0.027	–0.235**	–0.021
	(0.081)	(0.104)	(0.098)	(0.122)
Education of head	0.071	0.019	0.034	0.048
	(0.044)	(0.040)	(0.051)	(0.053)
Squared education	–0.002	0.000	0.003	–0.002
	(0.003)	(0.003)	(0.004)	(0.004)
Age of head	0.016	0.005	0.022	–0.162**
	(0.022)	(0.022)	(0.024)	(0.073)
Squared age	–0.000	0.000	–0.000	0.002**
	(0.000)	(0.000)	(0.000)	(0.001)
Female head	–0.258**	–0.153	–0.189	0.068
	(0.126)	(0.156)	(0.151)	(0.221)
ρb_i, *HIV/AIDS*	–0.15	–0.54	–0.16	–0.42

[1] Regressions include year dummies.
[2] Households who have been affected by illness/death at least once.
** indicate significance at a 5/10% confidence level.

However, the percentage of newly affected households receiving remittances increases slightly after the infection, from 27% to 34%. Alternatively, the fall in private transfers may be the consequence of social stigma associated with the illness as already underlined by other authors (Kawachi *et al.*, 1997, 1999; Kunitz, 2001; Seeley, 2002).

Table 4 displays a simple fixed-effects model estimation run on labour income and highlights the role played by the selection bias. It turns out that Bayesian and fixed-effects estimates of the role of HIV/AIDS on income levels are relatively close and both non-significant for urban and rural populations. Therefore, the selection bias is negligible. This was already suggested by the results of simulations presented in Table 2, where the selection bias had little influence on fixed-effects estimates. Again, this does not lower the interest of the Bayesian procedure, because participation equations need to be estimated in order to capture the non-linear impact of the illness on poverty.

Summing up, HIV/AIDS does not seem to have a significant impact on labour income in any group. The null impact on earnings may be interpreted as a high degree of substitutability of labour participation between the different members of the household, as shown by Chapoto and Jayne (2008) in the case of Zambia or Beegle (2005) in Tanzania. Also, the relatively short time span of the data is perhaps a significant limitation. On the other hand, HIV/AIDS affects urban households by increasing their participation in the transfer networks. In particular, they are more likely to receive public grants, which represent more than 70% of transfers as shown by descriptive statistics. Moreover, HIV/AIDS decreases the amount of private transfers received by rural households. Providing a strong explanation for this phenomenon is impossible with this dataset. However, we tend to view social discrimination or economic segregation derived from the breaking of traditional solidarity schemes as plausible candidates for an explanation.

5.3 Discussion

This section first provides the impact of HIV/AIDS on poverty using the former estimates, then discusses robustness issues.

(i) Poverty Impact We illustrate the magnitude of former estimates by calculating the implied impact of HIV/AIDS on poverty.[23] As expected, HIV/AIDS does not have any impact on labour income; it decreases chronic poverty slightly in the urban setting because of the increased probability of receiving (public) transfers. By contrast, HIV/AIDS diminishes the level of transfers (rather than mean participation) by 36% for the rural population. This has huge consequences in terms of chronic poverty by increasing the headcount level from 41.5% to about 60%. This stresses the importance of negative externalities of HIV/AIDS and its impact on poverty levels. Considering the squared poverty gap provides the same conclusion: The increase in poverty because of HIV/AIDS is large and permanent among rural households.

(ii) Omitted Variables To start with, controlling for fixed effects will not remove the influence of time-varying unobserved components. On this aspect, there is admittedly little to be done empirically, as this limitation is faced by any panel fixed effects estimation.

Besides, some explanatory variables correlated with HIV/AIDS, such as the dependency ratio and the dummy for female-headed households, possibly drive the correlation between income and the illness. In the former analysis, we treated these variables as exogenous ones, ignoring the possible impact that HIV/AIDS may have upon them. We now examine the effect of the illness on these two variables.

There are reasons to believe that the dependency ratio is an endogenous variable, but it is unclear whether it should increase or decrease in the wake of HIV incidence. Affected households may ask some relatives to join the household in order to assist the ill, or reallocate tasks among the members of the household as depicted by Mutangadura *et al*. (1999). If the relatives joining the household only do so to take care of the ill as elderly might do for instance, they may increase the dependency ratio. If active adults join the household to substitute for the ill in labour tasks, then the dependency ratio may stay constant or even decrease.

In practice, we regressed the dependency ratio on the HIV/AIDS variable, quadratics in education and age and time dummies in a fixed-effects model for different groups: the urban population, the urban population with positive labour income, with null labour income, with positive transfers, with null transfers and similarly for the rural population. For none of these 10 groups, the change in serologic status entailed a significant change in the dependency ratio. This result takes root in a mutual support mechanism by which active people from extended family or from the community join the household to help. Descriptive statistics support this explanation: Considering the newly affected

[23] A Monte Carlo experiment is run to generate the distribution of poverty indicators conditionally on suffering from HIV/AIDS or not. We draw from the empirical distribution of idiosyncratic residuals, whose cumulative distribution functions are computed with a Kaplan–Meyer procedure. This procedure enables us to reconstruct 1,000 counterfactual values for level and participation of labour income and non-labour income, from which we derive total income and poverty measures. We decompose the economic consequences of HIV/AIDS into four distinct channels – labour income levels, labour market participation and their counterparts for non-labour income. In each case, we compare the outcomes of two counterfactual populations, one with a positive serostatus (HIV/AIDS variable equal to 1) and a reference group with a negative serostatus. We use point estimates for all coefficients except for HIV/AIDS when the latter variable is not significant; in this case, its coefficient is set equal to zero.

population, we find that the dependency ratios before and after the infection are quite unchanged as their median values are both equal to 0.5, and percentiles 75 are both equal to 1. Whereas affected and non-affected households receive new members in almost equal proportions (respectively equal to 26.1% and 22.4%), in non-affected households 10.8% of new members contribute to household income vs. 18.1% for affected households. This proportion rises to 25.8% for newly affected households.

These findings offer an explanation for the fact that the illness does not impact labour income and that the dependency ratio remains somewhat constant: Both labour income and the dependency ratio are rapidly smoothed through household recomposition or extended labour supply after the incidence of the illness. As mentioned earlier, this result is consistent with Chapoto and Jayne (2008) or Beegle (2005), who find small and insignificant changes in labour supply of households experiencing a prime-age adult death. Like us, the latter study argues that household recomposition is one explanation of the lack of increase in hours worked by surviving household members.

On a second step, the gender of the household head was analysed as a dependant variable, and it was found again that the HIV/AIDS dummy was not significant, except for the group with strictly positive labour income in urban settings. Running a fixed effects model for labour income on this group while excluding the female-headed household dummy, we found a lower negative coefficient for HIV/AIDS but still not significant.

As a conclusion, modelling the potential impact of HIV/AIDS on the dependency ratio and the gender of household would not deliver significant effects and modify the main conclusions of the paper.

(iii) Sample Size The small sample size (about 170 households in each group) could entail large standard errors in estimates, maybe reducing the significance of some effects, in particular the HIV/AIDS impact on household labour income. Moreover, it could be argued that estimates based on a 10% transition rate (the share of newly affected households) are not credible enough given the small sample size. It is true that within a fixed effects panel data model, identification relies on the share of within variance, or said differently, on the degree of time variation contained in explanatory variables. With such a low proportion of individuals shifting HIV status across survey's waves, it is clear that our assessment of HIV/AIDS effects cannot be viewed as being strongly identified or very robust. However, in defence of our results, the Bayesian algorithm has given exact estimates when tested under the same conditions (*i.e.* a small simulated data sample with a dummy interest variable displaying a 10% transition rate over time). So the performance of the Bayesian framework remains good when applied to the same type of data sample. However, it is fair to acknowledge that the negative effect of HIV/AIDS on urban labour income participation is almost significant, and that a larger or longer survey could well deliver significant effects. This is an important caveat to the conclusion that the illness does not have any significant impact on labour income among urban households.

(iv) Measurement Errors In addition, measurement errors on the HIV/AIDS variable could be a cause of bias in the estimation. They could potentially explain why the HIV/AIDS dummy is not significant in labour income regressions. Indeed, measurement errors tend generally to bias the estimates towards zero, providing that measurement

errors are white noises independent from the observations. We rule out this possibility because we think that measurement errors more likely concern affected households who declare to be non-affected. In that case, measurement errors cannot be viewed as independent white noises anymore, and in this context, point estimates are consistent.[24]

6. CONCLUSION

In this paper, we have analysed the relationship between HIV/AIDS, income sources and poverty. Using original Bayesian methods, we have introduced an econometric framework that accounts for both self-selection and time-constant unobserved heterogeneity. We find that the illness does not have any significant impact on household labour income, suggesting that households manage to keep the overall labour supply constant despite the illness, at least on the short term. Although supported by other studies, this result might be influenced by the small size of our data sample or by its limited time coverage. Moreover, there is only a 10% transition rate in HIV status in the sample, which weakens the identification of HIV/AIDS effects within our fixed effects panel data model.

We also find substantial heterogeneity across urban and rural populations regarding the impact of HIV/AIDS on non-labour income. Among the urban population, the illness increases the probability of receiving (public) transfers. On the other hand, HIV/AIDS decreases the amount of (private) transfers received by rural households. Although the latter effect is difficult to interpret, we suspect that economic segregation and discrimination/stigma play a crucial role in this regard.

As a result, HIV/AIDS causes a massive 50% increase in chronic poverty among rural population, suggesting that the illness involves a permanent fall in household equivalent total income. In a context of decreasing private income transfers, public support should be strengthened together with the overall access to social basic services. These results also reinforce the relevance of fighting the diffusion of HIV in order to fight poverty. Prevention would be efficiently supported by free HIV testing, as argued by Thornton (2008), while the scaling up of antiretroviral programmes could reduce morbidity and mortality, increase labour income through increased participation, hours worked and children's school participation, as suggested by Graff Zivin *et al.* (2009) and Thirumurthy *et al.* (2008).

REFERENCES

AGHION, P., HOWITT, P. and MURTIN, F. (2011). The relationship between health and growth: When Lucas meets Nelson-Phelps. *Review of Economics and Institutions*, 2(1): 1-24.

AINSWORTH, M. and DAYTON, J. (2003). The impact of the AIDS epidemic on the health of older persons in Northwestern Tanzania. *World Development*, 31(1): 131-148.

———, BEEGLE, K. and KODA, G. (2002). The impact of adult mortality and parental deaths on primary school enrollment in Northwestern Tanzania. *Journal of Development Studies*, 41(3): 412-439.

AYIEKO, M. (1997). From single parents to child-headed households: The case of children orphaned by AIDS in Kisumu and Siaya districts. *Medicine and International Health*, 8: 507-511.

BARNETT, T. and JOINT UNITED NATIONS PROGRAMME ON HIV/AIDS. (2002). *The Private Sector Responds to the Epidemic: Debswana-a Global Benchmark*. Geneva: UNAIDS.

[24] The observed variable can be written as $x_i = x_i^* 1_{\varepsilon_i < 1-\lambda}$ where x_i^* is the true variable, ε_i is an independent uniform on [0, 1] and λ is the fraction of affected households who declare to be not affected. By independence of x_i^* and ε_I, it is clear that point estimates are unchanged.

BEEGLE, K. (2005). Labor effects of adult mortality in Tanzanian households. *Economic Development and Cultural Change*, 53(3): 655-683.
———, DE WEERDT, J. and DERCON, S. (2008). Adult mortality and consumption growth in the age of HIV/AIDS. *Economic Development and Cultural Change*, 56: 299-326.
BOND, V. (2002). The dimensions and wider context of HIV/AIDS stigma and resulting discrimination in southern Africa. In A. Sisask, (ed), *One Step Further: Responses to HIV/AIDS*. Gothenburg, Sweden: SIDA/UNRISD, 29-53.
BOOYSEN, F. (2004). Income and poverty dynamics in HIV/AIDS-affected households in the Free State province of South Africa. *South African Journal of Economics*, 72(3): 522-545.
——— (2005). The role of social grants in mitigating the socio-economic impact of HIV/AIDS in two Free State communities. *South African Journal of Economics*, 73: 545-563.
CASE, A. and DEATON, A. (1998). Large cash transfers to the elderly in South Africa. *Economic Journal*, 108(450): 1330-1361.
CASELLA, G. and ROBERT, C. P. (2004). *Monte Carlo Statistical Methods*. New York: Springer-Verlag.
CHAMBERLAIN, G. (1984). Panel data. In Z. Griliches and M. Intriligator (eds), *Handbook in Econometrics, Vol. 2*. Elsevier: North Holland, 1247-1318.
CHAPOTO, A. and JAYNE, T. (2008). Impact of AIDS-related mortality on rural farm households welfare in Zambia. *Economic Development and Cultural Change*, 56(2): 327-374.
COHEN J. (2002). Is AIDS in Africa a distinct disease? *Science*, 288(5474): 2153-2155.
COOMBE, C. (2002). Mitigating the impact of HIV/AIDS on education supply, demand and quality. In G. A. Cornia (ed), *AIDS, Public Policy and Child Well-Being*. Florence: UNICEF-IRC.
CORRIGAN, P., GLOMM, G. and MENDEZ, F. (2005). AIDS crisis and growth. *Journal of Development Economics*, 77(1): 107-124.
DEMOGRAPHIC AND HEALTH SURVEYS (2000-2006). Available at: http://www.measuredhs.com/accesssurveys/ [Accessed 29 November 2012].
DE WALQUE, D. (2006). *Who Gets AIDS and How? The Determinants of HIV Infection and Sexual Behaviors in Burkina Faso, Cameroon, Ghana, Kenya and Tanzania*. World Bank Policy Research Working Paper No. 3844. World Bank, Washington, DC.
DORWARD, A. and MWALE, I. (2005). Labor market and wage impacts of HIV/AIDS in rural Malawi. *Review of Agricultural Economics*, 28(3): 429-439.
EPSTEIN, B. (2004). The demographic impact of HIV/AIDS. In M. Haacker (ed), *The Macroeconomics of HIV/AIDS*. Washington, DC: International Monetary Fund, 1-40.
GRAFF ZIVIN, J., THIRUMURTHY, H. and GOLDSTEIN, M. (2009). AIDS treatment and intrahousehold resource allocation: Children's nutrition and schooling in Kenya. *Journal of Public Economics*, 93(7-8): 1008-1015.
HYSLOP, D. (1999). State dependence, derial correlation and heterogeneity in intertemporal labor force participation of married women. *Econometrica*, 67(6): 1255-1294.
JALAN, J. and RAVALLION, M. (2000). Is transient poverty different? Evidence for rural China. *Journal of Development Studies*, 36(6): 82-99.
JAYNE, T. S., VILLARREAL, M., PINGALI, P. and HEMRICH, G. (2005). HIV/AIDS and the agricultural sector in Eastern and Southern Africa: Anticipating the consequences. ESA Working Paper No. 04-06, Agricultural and Development Economics Division, Food and Agriculture Organization, Rome.
JENSEN, R. T. (2004). Do private transfers 'displace' the benefits of public transfers? Evidence from South Africa. *Journal of Public Economics*, 88(1-2): 89-112.
KAWACHI, I., KENNEDY, B. P. and GLASS R. (1999). Social capital and self-rated health: A contextual analysis. *American Journal of Public Health*, 89: 1187-1193.
———, ———, LOCHNER, K. and PROTHROW-SMITH, D. (1997). Social capital, income inequality and mortality. *American Journal of Public Health*, 87(9): 1491-1498.
KUNITZ, S. J. (2001). Accounts of social capital: the mixed health effects of personal communities and voluntary groups. In D. Leon and G. Walt (eds), *Poverty, Inequality and Health: An International Perspective*. Oxford: OUP.
LACHAUD, J.-P. (2007). HIV prevalence and poverty in Africa: Micro- and macro- econometric evidences applied to Burkina Faso. *Journal of Health Economics*, 26(3): 483-504.
LOEWENSON, R. (2007). Exploring equity and inclusion in the responses to AIDS, *AIDS Care*, 19(Suppl. 1): 2-11.
LUNDBERG, M., OVER, M. and MUJINJA, P. (2000). *Sources of Financial Assistance for Households Suffering an Adult Death in Kagera, Tanzania*. World Bank Policy Research Working Paper No. 2508. World Bank, Washington, DC.
MAYTRA, P. and RAY, R. (2003). The effect of transfers on household expenditure patterns and poverty in South Africa. *Journal of Development Economics*, 71(1): 23-49.
MCDONALD, S. and ROBERTS, J. (2006). AIDS and economic growth: A human capital approach. *Journal of Development Economics*, 80(1): 228-250.
MORRIS, C. N., BURDGE, D. R. and CHEEVERS, E. (2000). Economic impact of HIV infection in a cohort of male sugar mill workers in South Africa. *The South African Journal of Economics*, 68(5): 933-946.
MUNTHALI, A. (2002). *Adaptive Strategies and Coping Mechanisms of Families and Communities Affected by HIV/AIDS in Malawi*. Paper prepared for the UNRISD project on HIV and Development. Geneva: UNRISD.
MUTANGADURA, G., MUKURAZITA, D. and JACKSON, H. (1999). *Review of Household and Community Responses to the HIV/AIDS Epidemic in the Rural Areas of Sub-Saharan Africa*. Geneva: UNAIDS.

NAIDU V. (2004). Economic Impact of HIV/AIDS on Urban Households – A Pilot Study in Soweto. International Conference AIDS. 2004 Jul 11-16; 15.
NOBILE, A. (1998). A hybrid Markov chain for the Bayesian analysis of the multinomial probit model. *Statistics and Computing*, 8: 229-242.
NYBLADE, L. and FIELD-NGUER, M. L. (2001). *Women, Communities, and the Prevention of Mother-to-Child Transmission of HIV: Issues and Findings From Community Research in Botswana and Zambia*. Washington: International Centre for Research on Women.
ONI, S. A., OBI, C. L., OKORIE, A., THABEDE, D. and JORDAN, A. (2002). The economic impact of HIV/AIDS on rural households, small holder agricultural production and orphans in the Limpopo provinceof South Africa. *The South African Journal of Economics*, 70(7): 1173-1192.
REHLE, T. M. and SHISANA, O. (2003). Epidemiological and demographic HIV/AIDS projections: South Africa. *African Journal of AIDS Research*, 2(1): 1-8.
SAHN, D. E. and STIFEL, D. C. (2003). Progress toward the millennium development goals in Africa. *World Development*, 31(1): 23-52.
SEELEY, J. (2002). *Thinking with the Livelihoods Framework in the Context of the HIV/AIDS Epidemic*. Research Paper, Livelihoods Connect, Institute of Development Studies, University of Sussex.
STATISTICS SOUTH AFRICA (2000). *Income and Expenditure Survey*. Pretoria.
STEINBERG, M., JOHNSON, S., SCHIERHOUT, G., NDEGWA, D., HALL, K. and RUSSELL, B. (2002). *Hitting Home: How Households Cope with the Impact of the HIV/AIDS Epidemic*. Henry J. Kaiser Family Foundation Report.
THIRUMURTHY, H., GRAFF ZIVIN, J. and GOLDSTEIN, M. (2008). The economic impact of AIDS treatment: Labor supply in Western Kenya. *Journal of Human Resources*, 43(3): 511-552.
THORNTON, R. L. (2008). The demand for, and impact of, learning HIV status. *American Economic Review*, 98(5): 1829-1863.
TIERNEY, L. (1994). Markov chains for exploring posterior distributions. *The Annals of Statistics*, 22(4): 1701-1728.
UNAIDS (2009). Report on the global AIDS epidemic 2009.
UNPD (2008). United Nation Population Division, The State of the World Population.
WEBB, D. (1995). Community responses to AIDS orphans: Whose responsibility? *AIDS Analysis Africa*, 5: 6-7.
YAMANO, T. and JAYNE, T. S. (2004). Measuring the impacts of working-age adult mortality on small-scale farm households in Kenya. *World Development*, 32(1): 91-119.
YOUNG, A. (2005). The gift of the dying: The tragedy of AIDS and the welfare of future African generations. *Quarterly Journal of Economics*, 120(2): 423-66.

APPENDIX

The Algorithm

We use the following priors for Θ:

$$\begin{aligned}
\beta &\rightsquigarrow \mathcal{N}_K(\beta^0, B_0) \\
\rho &\rightsquigarrow \mathcal{U}([-1;1]) \\
\varepsilon_i | D_\varepsilon &\rightsquigarrow \mathcal{N}_2(0, D_\varepsilon) \\
\sigma^{-2} &\rightsquigarrow \mathcal{G}\left(\frac{\nu_0}{2}; \frac{\delta_0}{2}\right) \\
D_\varepsilon^{-1} &\rightsquigarrow \mathcal{W}_2(\rho_0; R_0)
\end{aligned} \tag{5}$$

The algorithm is the following:

Algorithm: Hybrid Gibbs Sampling for Selection Model with Correlated Specific Effects

1. At iterate $(j+1)$, sample

$$\beta^{(j+1)} \rightsquigarrow N_K\left(B^{(j)}\left(B_0^{-1}\beta^0 + (\sigma^{-2})^{(j)} \sum_{i=1}^{N} X_i'(Y_i^* - b_i^{(j)} \otimes i_T),\right.\right.$$

$$\left.\left. B^{(j)} = \left(B_0^{-1} + (\sigma^{-2})^{(j)} \sum_{i=1}^{N} X_i' X_i\right)^{-1}\right)\right.$$

2. M-H step:

a Draw a candidate value for $\rho^{(j+1)}$:

$$\rho^{(c)} = \rho^{(j)} + \tau u, \quad u \rightsquigarrow \mathcal{U}([-1;1])$$

b Evaluate the acceptance ratio α:

$$\alpha = \min\left(1, \frac{\pi(\rho^{(c)})}{\pi(\rho^{(j)})}\right)$$

where π is the posterior distribution of ρ. With uniform distributions set on ρ, the former ratio confounds with the likelihood ratio.

c Draw a random number $r \rightsquigarrow \mathcal{U}(0;1])$ and return

d Define

$$\lambda^{(j+1)} = N\left(\sum_i \bar{X}_i^{*\prime} \bar{X}_i^*\right)^{-1} \rho^{(j+1)} \circ \begin{pmatrix} sd\left(\bar{X}^{*(1)}\right) \\ sd\left(\bar{X}^{*(2)}\right) \end{pmatrix} \circ \begin{pmatrix} \sqrt{D^{(1)}, (j+1)} \\ \sqrt{D^{(2),(j+1)}} \end{pmatrix}$$

3. Sample

$$\varepsilon_i^{(j+1)} \rightsquigarrow \mathcal{N}\left(D_{\varepsilon,i}^{(j)} (\Sigma^{-1})^{(j)} \otimes i_T' \left(Y_i - X_i \beta^{(j+1)} - \bar{X}_i^* \lambda^{(j+1)}, D\left((D_\varepsilon^{-1})^{(j)} + T(\Sigma^{-1})^{(j)}\right)^{-1}\right)\right)$$

4. Sample

$$(D_\varepsilon^{-1})^{(j+1)} \rightsquigarrow \mathcal{W}_2\left(\rho_0 + N; \left(R_0^{-1} + \sum_{i=1}^N \varepsilon_i^{(j+1)} \varepsilon_i^{(j+1)\prime}\right)^{-1}\right)$$

and define the specific effects and their variance

$$b_i^{(j+1)} = \bar{X}_i^* \lambda^{(j+1)} + \varepsilon_i^{(j+1)}$$

$$D^{(j+1)} = \mathrm{Var}(\bar{X}_i^* \lambda^{(j+1)} + D_\varepsilon{}^{(j+1)})$$

5. Sample

$$(\sigma^{-1})^{(j+1)} \rightsquigarrow \mathcal{G}\left(\frac{\nu_0 + NT}{2}; \frac{1}{2}\left(\delta_0 + \sum_{i=1}^N U_i^{(j+1)\prime} U_i^{(j+1)}\right)\right)$$

where $U_i^{(j+1)} = Y_i - X_i \beta^{(j+1)} - b_i^{(j+1)} \cdot i_T$

6. Go to 1.