HIV/AIDS and Poverty in South Africa: a Bayesian Estimation

Federica Marzo, Fabrice Murtin
CREST-INSEE

29th June 2007

Abstract

In this paper we assess the causal impact of HIV/AIDS on monetary poverty using a panel dataset from South Africa. We model the consequences of the illness on both labour income and transfers. Two major econometric problems are likely to bias the estimation: endogeneity of the HIV/AIDS dummy variable, and auto-selection of the individuals participating in the labour market or in transfers networks. We solve both of them by proposing an original econometric framework where we include correlated fixed-effects both in the level and the participation equations, which are estimated simultaneously with original Bayesian methods. The procedure is tested and well-behaved. Splitting the sample into urban and rural population, we show that HIV/AIDS has surprisingly no impact on household labour income for both groups. Thus the correlation between low labour income and incidence of the illness stems from other observed and unobserved characteristics correlated with incidence. For the urban population, we find that HIV/AIDS has a positive but small impact on the probability of receiving social and private transfers. On the contrary, HIV/AIDS increases poverty by a large extent among the rural population because it causes a sizeable fall in received transfers, a fact that might underline the social stigma associated to the illness.

1 Introduction

At the end of 2006, UNAIDS estimated that the number of people living with HIV/AIDS in the world was 39.5 millions, 95 percent of them belonging to developing countries and more than 60% of them living in Sub-Saharan Africa. Southern Africa remains the epicentre, hosting one third of total HIV-positive population and 34 percent of total deaths[1]. In those countries, the epidemic has been having a dramatic and tangible impact on demography, with life expectancy falling to the level of the fifties (according to UNDP, 1997)[2]. In recent years, some progress has been made on the understanding of the economic consequences

[1] UNAIDS 2006.

[2] See The United Nation Population Division ”The State of the World Population” report. Series for life expectancy and life expectancy without aids are reported.

of HIV/AIDS thanks, among other things, to an increased availability of data. Many negative economic externalities have been highlighted and, sometimes, quantified : the negative impact on the labour force, the decrease in savings and investments, the reduction of school enrollment and teaching staff, the collapse of family and community solidarity structures.

The impact over income generating activites is straightforward and is considered the first short term impact: in fact the symptomatic phase of the illness rapidly causes progressive, and often irreversible, physical deterioration, which results in decreasing productivity, reduced participation, and a diminution in affected household earning capacity. This dynamic, which entails a series of other economic consequences (*short term* impacts), could most likely turn into a shock on available income. However, households are known to adopt various coping strategies in order to absorb income shocks, especially when the surrounding environment is very risky. This is certainly the case for households afflicted with AIDS as well. Besides these forms of *households reactions*, several forms of *collective reactions* also exist, based on solidarity network of relatives and neighbours. These could consist in financial support, sharing meals, sharing fields or cattle, provide some labour or hosting new persons in the household (household recomposition), both new active members or orphans. Altogether, households and collective coping strategies are expected to mitigate the short terms consequences of HIV/AIDS. Despite these efforts, HIV/AIDS related morbidity and mortality have been associated with dramatic level of poverty in several Southern African countries, including South Africa (Booysen (2004), ONI et al. (2002) Jayne et al. (2005)) Botswana (Greenei et al. (2000)), and Malawi (Dorward et al. (2006)). Unfortunately, and due to the complex interactions underlying this dynamic, most of these studies document only *correlations* between HIV/AIDS and poverty, the assessment of a causal inference remaining problematic.

The aim of our paper is ti fill this gap and to analyse the complex relationship between AIDS and monetary poverty in South Africa, estimating its final causal impact on income, as well as on poverty level. In order to take into consideration possible short and long term impacts, we consider both transitory and chronic (permanent) poverty using a panel of six waves (three years). In fact, coping strategies might turn to be unsustainable on the long run, or even have a negative effect on permanent income. On the other hand, mutual insurance and support, although positive for the households receiving this support, might seriously affect the livelihood of non-affected households providing for it, especially in high prevalence settings. This is why in general we consider that for every households suffering from a shock, at least four other households are involved through negative externalities. In this sense, poverty might spread in entire communities and not only for directly affected households[3].

In practice, before computing Monte-Carlo simulation on overall monetary poverty, we estimate the impact of AIDS on income, decomposing it in labour income and non-labour

[3]This consideration she some doubts of the appropriateness of the classification of affected and non-affected households.

income (including public grants and remittances). Such a decomposition is useful to understand the economic consequences of the illness, since they might be different on each component. From an empirical perspective, the paper tackles two important methodological issues which represent major obstacles to the economic research on HIV/AIDS, together with data scarcity.

First, we take into account selection effects by relying on the longitudinal dimension of the data rather than on the use of any instruments, e.i. what affects the level of wage might determine participation in the labour market as well. As largely emphasized by the economic literature on the subject, unobservable characteristics affecting participation in the labour market or in financial transfers might hamper the econometric analysis if they are correlated with some observable characteristics. In this case, an econometric bias is likely to exist since HIV/AIDS plausibly affects participation in the labour market. A second important reason to focus on selection effects is that income realisations are not linear since they drop to zero when individuals do not participate. This non-linearity is central in the poverty analysis and this is why it is crucial to jointly assess the impact of the illness on both participation and income level. Our methodology allows us to solve the selection problem introducing fixed-effects both in the income level equation and the participation equation, and permitting a non-null correlation between those two fixed-effects (since unobserved heterogeneity affecting participation is allowed to influence the level equation too, via the correlation of the two fixed-effects variables).

Second, we take into account the endogeneity of HIV/AIDS. Indeed, the illness presumably increases poverty through various channels. At the same time, it is possible that some unobservable factors affect both the likelihood of being affected by the disease and the standards of living, even though the sign of the correlation between those factors and the HIV/AIDS dummy variable remains unclear. For instance, migrations may be a source of diffusion of the epidemic, but also a source of wealth through remittances or increased information (about market or about 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 a negative sign of the correlation. Once again, the longitudinal dimension of data allows us to introduce fixed-effects in the estimation, solving the problem of endogeneity of the HIV/AIDS variable.

Our econometric model requires sophisticated inference methods. Recently, there has been a surge of original Bayesian methods enabling the estimation of complicated econometric models. The use of such procedures stems mainly from the fact that it is more simple to simulate a distribution via Monte-Carlo Markov Chains methods (MCMC) rather than finding the mode of a distribution via maximization algorithms[4]. In this paper we use a modified version of the Gibbs sampling algorithm introduced by Nobile (1998), called the *hybrid Gibbs sampling*. The idea is to combine the two building blocks of Bayesian econometrics, the Metropolis-Hastings and the Gibbs algorithms. Following a procedure introduced and validated by Murtin (2006), we explicitly model the correlation between

[4]Beffy et al. (2004) and Murtin (2005) provide some studies based on such Bayesian framework.

fixed-effects and observable variables, dramatically fastening the convergence speed of the classical Gibbs sampling algorithm.

The first surprising result of our Bayesian estimations and Monte-Carlo simulations on poverty is that HIV/AIDS has no significant consequences on household labour income, neither for urban nor for rural population. This seem to indicate that households menage to smooth labour income and participation over time, despite the short term negative shock of the illness. This is most likely due to collective reactions positive effects, mainly household recomposition. In this sense we do not seem to be estimating the direct impact of the illness, but the final result of coping strategies. Therefore, the observed negative correlation between the illness and the household labour income is due to other observed or unobserved determinants, which are correlated with AIDS. However, the small sample size could explain the non significance of this effect.

Our second result show that there is a difference between rural and urban setting as for the impact of AIDS on transfers. The illness brings to an increase in participation in the transfers 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 transfers, increasing chronic poverty by a huge 50 percent. This effect might be interpreted as a direct impact f the illness on remittances, to a disruption of mutual insurance networks within the communities due to discrimination, or to a lower take-up of grants in general, as already showed by Booysen (2005). Transitory poverty remains constant for this subsample, suggesting that the increase in poverty is permanent.

In conclusion, the paper fits well in the literature. Although the analysis is not representative of the South African population, it offers a general framework to analyse the potential causality hidden behind some of the correlations largely described in the literature on AIDS, such as HIV/AIDS and labour market participation, participation in transfers system, the level of transfers received and so forth.While indicating that the effect of the illness on labour income is not significant at household level, this study stresses the importance of transfers, both public and private, from which the most vulnerable population seems to be excluded. The context in South Africa has hugely evolved during recent years, with ARV treatment being more and more available among the population. Scaling up efforts have positive consequences in terms of increased life expectancy for people living with AIDS and recovered earning capacity. This aspect must be considered when evaluating the conclusions of this study.

The paper is organised as follows: 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 characteristics to the actual ones, and compare it with traditional estimators. In a third section we describe the data, setting and results, and we assess, in section 4, the causal impact of HIV/AIDS on transient and chronic poverty by means of a Monte-Carlo simulation. Last section concludes.

2 HIV/AIDS and poverty in the literature

This paper represents an attempt to quantify the impact of HIV/AIDS on poverty. HIV/AIDS impoverishes people, this is a fact and several papers and empirical case studies show it (Barnett, Whiteside 2002).[5] Yet, the channels through which the illness affects households are numerous, and the final output of their interaction is not straightforward. For this reason, it is convenient to proceed to the definition and illustration of the different kind of impacts on household livelihood.

We call *short term impacts* the socio-economic consequences caused by the illness morbidity and mortality. Even if morbidity and mortality are spaced in time, we can reasonably consider that their consequences are short-term. Indeed, in absence of ARV treatment the duration between the onset of the symptomatic phase of AIDS and the death of the ill is about 12 to 18 months in African countries (Stillwaggon 2000). It is common knowledge that the most important direct economic consequences for the household is the decrease in productivity for the ill persons and, most often, for its entourage, consequently leading to a sharp decrease in household labour income, which can reach about two thirds of mean income (Morris, Burdge and Cheevers (2000)). From an aggregate perspective, this decrease in productivity is likely to have dramatic consequences on overall labour force, both in number and in structure, since most of the people living with HIV/AIDS belong to the 15-35 age band (Lisk 2002). Macro-economic modeling results on South Africa indicate that given the impact of HIV/AIDS, the rate of labour force growth will decline, resulting in a smaller labour force when compared to a non-aids scenario (BER 2001). Equally, prediction on the evolution of labour force fluctuates from a 18 percent decline (Quattek 2000) by 2015 to a 21 percent decline, always by 2015 (BER 2001). It is convenient to distinguish the impact on labour income passing through a fall in *productivity* (represented by wages) and the one passing through *participation* in the labour market (absenteeism or job abandon) With this respect, there is abundant evidence that both time worked and productivity decline sharply well before a worker dies or retires (Morros and Cheevers, 2000; Fox et Al.2004). More generally, several studies have shown that the death of an adult due to AIDS is more costly for the household than if the adult had died from another cause, that household consumption of basic necessities fell considerably and that the death of a female adult has a stronger impact on household recomposition (Ainsworth, Fransen and Over 1998). In fact, the second important short term impact is an increase in demand for and, consequently, the cost of social services, medical assistance and funerals. In particular, HIV/AIDS increases demand for health related goods, while labour income decreases. Steinberg et Al (2002) show that in South Africa affected households spend about one third of their income on health care, compared to a national average of 4 percent. Always in South Africa, the high costs of funerals has become a real threat for the economic security of the households (Ayeko 1997) and can reach the amount of 40.000

[5] But also Booysen 2004; Oni et al. 2002; Jayne et al 2005; Greenei et al. 2000; Yamano and Jayne 2002 Dorward et al. 2006. All these studies associate poverty and AIDS in Sub Saharan African countries.

ZAR (Steinberg et Al. 2002).[6]

Directly linked to the fall in household earning capacity there is a corresponding increase in the dependency ratio.[7] However, since infant and child mortality also increase and there might also be a recomposition effect (see below in the indirect impacts), with old persons coming to help but also new active people joining, this evolution in the dependency ratio is not clear cut (B.G.Epstein 2004; Rehele and Shisana 2003). Nevertheless, in absence of household recomposition in terms of new active comers, the number of inactive in the household should consistently increase, and this is the third important direct impact (Lisk 2002). In fact, there is strong evidence that orphans are hosted by extended families when one or both parents die. According to several DHS surveys conducted at the beginning of the years 2000 on household composition, one-fifth to one-quarter of households in high prevalence African countries are fostering children. This direct consequence supposedly has a negative impact on the livelihood of the hosting family and of the children in particular, in a context of diminished income and increased expenditure. In this sense, it is very disputable to limit the study of the illness consequences only on households directly suffering from death or morbidity episodes, since this would bring to the underestimation of the real impact of it on African social fabric.

Longer term impacts derive from households *reactions* to survive to the illness and its consequences. Importantly, these *coping strategies* aim at providing immediate relief but might have negative effects in the long-term. The fall in labour income and the contemporary increase in expenditure make households redistribute resources (and time) in favour of the persons living with HIV/AIDS, entailing a cut in basic consumption goods (such as clothing and electricity) and possibly entailing malnutrition for the other members of the household (Ainsworth and Dayton (2003)). Moreover, there is evidence of dissaving or of assets sales, such as cattle and livestock, furniture, work instruments.With production capacity weakened and exhausted savings, in the longer term consumption might decrease further, or stabilize on a low level and, if assets and saving are not replaced, future investments might be compromised. This concern every kind of investment, in physical capital, in social capital and in human capital. As for the latter, affected households tend to withdraw children from school (Ainsworth et al. 2002; Nampanya-Serpell, 2000), sending them to work or look after the ill.[8] The economic mechanism behind this attitude is straightforward: since the expected working time of the average worker declines, due to a reduction in life expectancy, the expected returns on investments in human capital also decline.

But coping strategies refer also to external actors' (such as the public sector or the surrounding community or extended family) *response* to the impact of the illness (*collective coping strategies*) and so do negative longer term externalities. Public transfers mainly come from public health services and take the form of destitution allowances, disability

[6]More than 5700 US dollar (2007 exchange rate).

[7]The Dependency ratio is the number of children and people of retirement age divided by the adult population available to support them (15-49 age band).

[8]However, Coombe (2002) suggests that the impact of the epidemic on school attendance is hard to estimate because the reasons why children are withdrawn from school are usually unknown

grants or orphans allowances. Naidu (2003) studies the evolution of income in Soweto (South Africa) and discovers that a large part of income shock caused by the illness is absorbed by public grants: the fall in income for affected household approaches 30 percent, but after public transfer it is limited to 8 percent..The informal sphere encompasses worker remittances or transfers, and belong to risk sharing or, more generally, traditional solidarity schemes. External support might assume different forms: extended family or even unrelated persons (neighbours or persons belonging to the same ethnic group) might join the household to help, or might give other kind of support, like paying for funerals or contributing to other expenses (Marzo Murtin 2057, Over and Mujinja 1993)). This phenomenon, although positive for households benefiting from it, since it can mitigate the negative impact of mortality on labour income (Mather et Al. 2004) might be negative in the longer term for supportive households (or for public finances) and, as a results, the longer term impact might be that entire communities turn out to be impoverished by the illness. In particular, several studies report that traditional safety networks are under severe stress when HIV/AIDS-related illnesses and mortality increase (Kawachi,Kennedy, Lochner and Porthrow-Stith, 1997; Kawachi, Kennedy and Glass 1999; Kunitz 2001 and Marzo, Murtin 2045).In conclusion non-labour income (grants and remittances/gift) as a form of collective coping strategy can constitute an important (monetary) support, especially for the poor, and can make the difference with respect to the final impact of AIDS on households livelihood.

Overall, the channels through which the economic shock is transmitted are the key points of our analysis. In this paper we will disentangle the impact passing through labour income (which represents productivity and labour participation) from the possible impact on transfers (public grants and private remittances), which might also be positive if it those transfers are interpreted as being is the result of collective coping strategies. Since the literature has shown that impacts and coping strategies depend on initial household endowments (Yamano and Jayne 2004)[9], both in wealth and in social capital, we distinguish between rural and urban households. We will also analyse the economic consequences on both the short and longer-term. As suggested above, indirect factors might indeed result in an alleviation of the income shock in the short-term, but not necessarily on the long-term.

3 The econometric framework

This section presents the model and illustrates its benefits with respect to other traditional approaches.

We note $e_{i,t}$ for the participation dummy and $y_{i,t}$ for income. The selection model is

[9]They show that impact of adult mortality on households is related to households' wealth: initial assets endowments account for a lot and negative impacts on (agricultural) production, assets and off-farm income is significant and negative only if the male head dies and the context is one of poverty.

a system of two equations assuming Gaussian residuals

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

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

$$\begin{aligned} Y_i^* | \beta, b_i, D, \sigma^2 &= \mathcal{N}\left(X_i\beta + b_i \otimes i_T, \Sigma\right) \qquad i \le N,\ t \le T \\ Y_i^* &= (y_{i,1}^*, .., y_{i,T}^*, e_{i,1}^*, .., e_{i,T}^*)' \\ X_i &= \begin{bmatrix} X_i^{(1)} & 0 \\ 0 & X_i^{(2)} \end{bmatrix} \\ \beta &= [\beta^{(1)\,\prime}\ \beta^{(2)\,\prime}]' \\ b_i &= [b_i^{(1)}\ b_i^{(2)}]' \\ D &= \begin{bmatrix} D^{(1)} & D^{(1,2)} \\ D^{(1,2)} & D^{(2)} \end{bmatrix} \\ \Sigma &= \begin{bmatrix} \sigma^2 I_T & 0 \\ 0 & I_T \end{bmatrix} \end{aligned}$$

where $\otimes$ is the Kronecker product. There are two major issues arising in this context: the correlation structure of the model, and missing data since the dependent variable is partly observed or completely unobserved as with the latent variable $e_{i,t}^*$. The first issue deals with endogeneity, the second with the selection problem. 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. 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. Moreover, we would like to account for endogeneity of the observed variables, so that fixed-effects have mean zero, but not necessarily conditional (on observed variables) zero mean. In short, E $[b_i|X_i] \neq 0$. As described by Murtin (2006), the correlation between fixed-effects and endogenous variables must be modeled if we want the Gibbs sampling algorithm to converge rapidly. The most simple 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). 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.[10] However, this is not a problem for us since the main endogenous variable, a dummy for HIV/AIDS status, is time-varying. More precisely, one can decompose the vector of specific effects in the following way.

Noting

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

where $\bar{X}_i^{(j\in\{1,2\})}$ is the 2 by $K.N$ matrix of the individual means of $X_i^{(j)}$ in equation (j), $\bar{X}$ the 2 by K matrix composed of the grand mean of the variables, one derives easily the following specification

$$\begin{aligned} B_i &= \bar{X}^*.1_N \otimes \lambda + \varepsilon \\ (1_N^{'} \otimes I_K).\bar{X}^{*\prime} B_i &= (1_N^{'} \otimes I_K).\bar{X}^{*\prime}\bar{X}^*.1_N \otimes \lambda + (1_N^{'} \otimes I_K).\bar{X}^{*\prime}\varepsilon \\ \sum_i \bar{X}_i^{*\prime}(b_i \otimes i_T) &= \sum_i \bar{X}_i^{*\prime}\bar{X}_i^*\lambda = \left(\sum_i \bar{X}_i^{*\prime}\bar{X}_i^*\right)\lambda + \sum_i \bar{X}_i^{*\prime}\varepsilon_i \quad \text{with } \varepsilon = [\varepsilon_1^{'} \quad \varepsilon_N^{'}]^{'} \\ \lambda &= N\left(\sum_i \bar{X}_i^{*\prime}\bar{X}_i^*\right)^{-1} \rho \circ \begin{pmatrix} sd(\bar{X}^{*(1)}) \\ sd(\bar{X}^{*(2)}) \end{pmatrix} \circ \begin{pmatrix} \sqrt{D^{(1)}} \\ \sqrt{D^{(2)}} \end{pmatrix} \text{ with } \varepsilon_i \perp \bar{X}_i^* \end{aligned}$$

where 1_N is a N vector column of 1, I_K KxK identity matrix, X the $2.T.N$ by $2.K.N$ matrix with diagonal block X_i and 0 elsewhere, $B_i = [b_1^{'}...b_N^{'}]'$, ρ the K vector of correlation between specific effects and individual means of regressors, $sd(\bar{X}^{*(j)})_{j\in\{1,2\}}$ the standard error of individual means of regressors in equation (j), $D^{(j)}$ the variance of specific effects in equation (j).

Then, the second difficulty to cope with is missing data, namely that $e_{i,t}^*$ is unobserved, as well as and $y_{i,t}^*$ when $e_{i,t}^* < 0$. A strength of the Bayesian approach is that missing data can be treated just as other parameters of interest: they are simulated. Indeed, it is straightforward that given the set of parameters Θ, the density of Y_i^* can be decomposed with Bayes rule

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

Hence when $y_{i,t}$ is observed the data augmentation step consists in drawing $e_{i,t}^*$ from its posterior distribution, namely a truncated normal taking values on the interval $]0, +\infty[$.

[10]In that case, identification might be weak if one is to specify vague and relatively uninformative priors, and convergence will be significantly slow down, though still achievable

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 $]-\infty, 0[$ for the second component $(e^*_{i,t})$.

Now, let us describe the algorithm. In a Bayesian setting the goal is to infer the conditional distribution $p(\Theta|Y)$, which is proportional to the posterior distribution $p(Y|\Theta)\,p(\Theta)$ by Bayes rule. Some prior distributions $p(\Theta)$ are set on the parameters and, for Gaussian panel models, priors and the sampling distribution[11] $p(Y|\Theta)$ 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. Again, in this context $\Theta = (\beta, \rho, \varepsilon_i, D_\varepsilon, \sigma^2)$. The choice of priors is far from being a limit to the estimation procedure, because prior information can be taken large enough not to be binding.

Inference is achieved with an hybrid version of the Gibbs sampling algorithm as in Murtin (2006). The Gibbs sampling algorithm is an iterative approach that draws from the conditional posterior distribution of each block of parameters[12] conditional on previous drawings of other blocks of parameters. This algorithm constitutes a Markov Chain that converges towards the stationary distribution of parameters under fairly general conditions.[13] As the posterior distribution of the correlation ρ cannot be written in closed-form, we simulate it using a Metropolis-Hasting step, which is at the origin of the term "hybrid" Gibbs sampling. Such an approach was introduced by Nobile (199x) and is extensively described by Casella-Roberts (2004). Priors and the detailed algorithm are fully described in annex 1.

We test this algorithm on a simulated dataset and show that the coefficients of all endogenous variables are perfectly estimated. For this test 100 000 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

$$\begin{aligned} y^*_{i,t} &= \mu^{(1)} + \delta^{(1)}_t + b^{(1)}_i + \beta^{(1)} X_{i,t} + \sigma u_{i,t} \\ e^*_i &= \mu^{(2)} + \delta^{(2)}_t + b^{(2)}_i + \beta^{(2)} X_{i,t} + v_{i,t} \\ \forall t, \quad e_{i,t} &= I\,[e^*_{i,t} > 0], \quad y_{i,t} = e_{i,t} \cdot y^*_{i,t} \\ b^{(j)}_i &= \nu_i + \epsilon^{(j)}_i, \ \nu_i | \epsilon^{(j)}_i \forall j \end{aligned} \tag{2}$$

The endogenous variable X is specified as a dummy variable that takes value one if $\nu_i > 0$ and 0 otherwise. Moreover, we specify a time-varying component by allowing 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 all the characteristics of the HIV/AIDS dummy variable in the data, it has an impact both on income level and participation, and is correlated to fixed-effects $b^{(j)}_i$ via the time-constant component ν_i. As a result, the percentage of censored observations is equal to 24% among the "non-affected" population, namely those for which $X_{i,t} = 0$, and 34% for

[11] Namely, the likelihood.

[12] In this context the 5 blocks corresponding to $\beta, \rho, \varepsilon_i, D_\varepsilon, \sigma^2$.

[13] See Tierney (1994)

the others. Again, these figures match the data on earnings. Table 1 presents the results for three different estimators: a Tobit random-coefficients model, a fixed-effects model applied to non-censored observations, and the hybrid Gibbs sampling described above. As a result, it is clear that the Tobit model delivers poor estimates of the income level equation, with a 25% downward bias. This was expected: with a positive correlation between fixed-effects and the endogenous variable, estimates overestimate the magnitude of the effect, hence entail a downward bias since 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 our poverty simulations. Again, the direction of the bias was expected since the correlation between the endogenous variable and unobserved determinants of participation was taken positive: as a result the observed "affected" population was not representative of the total "affected" population, with only those with a good draw from the fixed-effects distribution being included into the sample of estimation. This underestimates the effect of the HIV/AIDS variable on income level. In 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 correctly selection effects.[14] Figure 1 describes the successive drawings from the algorithm for the coefficients of interests and the correlations between fixed effects and endogenous variables.

4 Data

The impact of HIV/AIDS on poverty is studied with the help of a panel 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.[15]

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

The balanced survey is composed of 331 households and 1 173 individuals with data

[14]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.

[15]Households were defined using the standard definition employed by Statistics South Africa in the October Household Survey (OHS), i.e. "a person or a group of persons who live together at least four nights a week" (Statistics South Africa, 1995: 0317-E) and who share resources. Interviews were conducted with one key respondent only, namely the "person responsible for the daily organization of the household, including household finances".

available at each wave. The survey has two main characteristics, which represent also major problems: the selection process of affected and non-affected households, and the large heterogeneity between rural and urban households.

Households are defined as *affected* if, at the time of the interview, someone belonging to it has declared being HIV-positive. Households belonging to this group have been selected through NGO's and public services working in the field of HIV/AIDS. Informed consent prior to the utilisation of the data has been given by concerned people or by their relatives. An important aspect needs to be underlined: 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. A comparison group of equal size not directly affected by HIV/AIDS at baseline was interviewed on a voluntary basis. They were meant to have similar characteristics to affected households thanks to the selection process.[16] Nevertheless, one cannot be completely sure of having selected a non-affected household, since people might be unaware of their serostatus. Importantly, the classification of affected and non-affected has been revised wave after wave: households who experienced illness or death over subsequent waves were reclassified as "newly affected". This group is made of 33 households, about 10% of the original sample.[17]

The second important feature of the survey is the choice of the settings. According to Statistics SA (2000), the Welkom magisterial district, situated in the Goldfields, is the third richest in the Free State province, with a headcount poverty ratio of 0.34 and average monthly household expenditure of 2364 ZAR. It can be defined as an urban setting.[18] In 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 poor infrastructure and social services. The headcount poverty ratio in this district is 0.69, while average monthly household expenditure amounts to 807 ZAR. Thus, the particular selection of study sites allows us to compare the household impact of HIV/AIDS on poverty among communities that differ substantially in terms of standard of living and access to basic services. In both setting prevalence is very high and the Freestate has the second highest prevalence of HIV/AIDS and is also the province with the second highest prevalence growth (Cohen, 2000).

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

[17] Households originally classified as non-affected but who declared in the second wave having experienced HIV/AIDS related deaths or illness in the recent past were reclassified as "affected".

[18] The distinction between rural and urban setting is made on economic activities difference and on a governance bases (traditional vs modern), and not on the differences in dwellings equipments and infrastructure endowments.

To conclude this section, it is important to note that the findings from this study cannot be generalized to households across South Africa because of the small sample size, a feature shared by most other HIV/AIDS impact studies, and because of the sampling procedure. Yet, the results reported in these pages, albeit context -specific, do present a telling picture of the socio-economic impact of the HIV/AIDS epidemic.

5 Estimation

5.1 Descriptive statistics

The sample is composed of 167 urban and 166 rural households. Among these, the proportion of affected households is by construction equal in size, and about 55%. We define *affected* households those who declared to the interviewer being so, or having suffered from AIDS-related illnesses or deaths (according to symptomatology). In contrast, *non affected* households are those having declared the contrary.[19] First, as shown in Table 1, there are important differences between urban and rural populations. On average urban households are more educated than rural ones and their income is 40% higher Grants and remittances constitute a smaller proportion of household income in urban areas, about 45% versus 55% in rural areas. Moreover, higher mean income translates into a lower chronic poverty rate[20]: 23% versus 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 to the urban: 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).

Second, HIV/AIDS splits clearly each group of population into two separate subgroups in terms of labour income. For the urban population, affected households earn about 27% less than non-affected households and their unemployment rate is 10% higher. On the other hand, they rely more on non-labor income since 70% participate in the transfers network (public or private) versus 54% for the non-affected. Most of this increase can be attributed 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. On average, affected households earn 20% less than non-affected in terms of total income. For the rural population, the unemployment rate is 10% higher among affected households and labour income 49% lower. Both subgroups are comparable with respect to transfers, most likely because rural setting, and former Qwaqwa in particular, have very poor access to public social services, regardless from households characteristics. Overall, total income is 40% lower among affected rural households. As an illustration, figure 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. From these graphs we see that social grants are the most important component of transfers

[19]For more details, see Booysen 2004 for the description of the survey.

[20]For the definition of headcount index for chronic and transient poverty, see below.

income for the urban population and that public and private transfers are equal in share for rural households.For total income and its components, we have used real figures in adult equivalent terms (meaning that y real figures have been divided by the total number of persons in the household power 0.6).[21]

In conclusion, HIV/AIDS seemingly entails a fall in labour income partly compensated by transfers among urban households, where public transfers seem to be play the role of collective coping strategy. On the contrary a large drop of labour income with unchanged transfers characterises rural population. Since these are only correlations, next section is going to show that some of them reflect a causality, with some others being spuriously driven by endogeneity.

5.2 Procedure and results

As a starting point it is important to tackle the issue of simultaneity: potentially, 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 a regressor, and on the other way around non-labour income may enter the labour income equation. We explain here why we have rejected the second specification.

First, in the empirical literature many 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. Regarding South-African data, Jensen (2004) does not find any significant impact of old-age pensions on the household labour supply or composition. On the other hand, 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 force. Hence, it is difficult to infer any causal interpretation from these correlations. Although difficult to justify empirically, our assumption is consistent with the economic mechanism we have in mind to explain the dynamic of income: HIV/AIDS is a short term income shock, whose consequences are tackled through various coping strategies, including extended family or community remittances and public grants.. In other words, the increase of non-labour income is a consequence either of the decrease in earnings caused by the illness (let us call it indirect impact) or of the illness itself (let us call it direct impact[22]). Of course, one could argue that, if correctly anticipated, the transfer windfall might have a negative impact on earnings via a substitution effect. According to us, this view neglects the fact that most of our sample population live in poverty, and would hardly diminish their commitment to work, even in presence of external monetary support: in other words, their labour supply is plausibly constrained.

[21] $\frac{y_i}{n^{0.6}}$, were y is real income (or one of its component), n is the total number of household member and 0.6 is the adult equivalent coefficient.

[22] In this context the terms *direct* and *indirect* do not refer to the broader definition of AIDS impacts seen above.

To sum up, we are interested in the impacts of HIV/AIDS on both labour income and non-labour income, as well as in the indirect impact on transfers passing through a decrease in labour income. We neglect the indirect negative impact of HIV/AIDS on labour income passing through a decrease/increase in transfers. By the way, we expect the impact on transfers to be positive if non labour income is the result of collective coping strategy (public or private), and negative if the illness somehow affects preexisting transfers (through discrimination or directly affecting people sending remittances).

A Bayesian procedure is used to estimate the impact of HIV/AIDS on labour income[23], while specifying a unique endogenous variable, the HIV/AIDS dummy variable[24](indicating whether the household belongs to the group of *affected* or to the control group). Results 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 nor on participation. This is surprising, especially for the rural population, since descriptive statistics seemed to indicate that affected households had lower levels of income and higher levels of unemployment. This result is explained by other observed and unobserved characteristics correlated with HIV/AIDS and having a negative impact on labour income: for the urban population, the dependency ratio, the age and gender of household head; for the rural population, the dependency ratio, age, education and gender of household head. The negative correlation between fixed-effects and HIV/AIDS in the level equation for the rural population explains a large part of the negative correlation between labour income and the illness, as well as pre-existing differences in household composition and unobservable variables. Since the impact of the illness on labour income level is not significant, it means that households adapt to the new situation, increasing labour participation of non-ill persons, or fostering new active members in the households (recomposition coping strategy). This story be no longer true had HIV/AIDS a significant impact on household composition, an hypothesis that is discussed in next sub-section. Figures 3 and 4 illustrate the convergence of the Bayesian algorithm.

On a second step, we estimate the non-labour income[25] equations.[26] We specify two endogenous variables, HIV/AIDS and the latent labour income variable estimated before. As a result, fixed-effects are a linear combination of those two variables' within-averages and another independent component. The set of regressors is the same across both the level and participation equation: in a Bayesian framework, there is no need of excluding any variable since identification stems from the priors.[27] Results are illustrated in the second

[23]Labour income is defined as earnings coming from any form of formal or informal working activity, including waged activities and subsistence agriculture.

[24]in this paper one is only interested in the causal impact of this particular variable on poverty.

[25]Non labour income is defined as all the income components not belonging to the sphere of working activities, namely all public grants as well as private remittances and in-kind gifts.

[26]Figure 5 and 6 depict the Bayesian algorithm.

[27]Those prior distributions ensure that the target distribution, the posterior distribution, has an unique global mode, contrary to the likelihood that is constant over a manifold of parameters. Priors need not to be well specified since they are vague and do not contribute by more than 1% to final estimates. However, the distributional form of priors could be more important. A robustness analysis is left aside for future work in order to gauge the consequences of mis-specification in the distributional forms.

part of Table 3. Although not significant on the level of non-labour income, the dummy HIV/AIDS has a positive impact on participation for the urban population, meaning that households are targeted according to their serostatus. 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 since the former has a negative impact on the latter. On the contrary, the impact on the level of rural population non labour income is negative and significant, which was not visible from the descriptive statistics. Again, this is due to the positive correlation between fixed-effects and HIV/AIDS. Since affected households were more prone to receive transfers for other reasons than the health conditions of its members (responding to informal solidarity schemes), the causal impact of the illness was hidden. This result is important: it entails that affected rural households are deprived from vital monetary resources *because of* their illness. In this sense, and unlike public transfers, private transfers do not seem to be a coping strategy, triggered by the illness, since they do not increase to respond to it. Figures 5 and 6 depict the Bayesian algorithm.

A comparison with Table 4 (of the fixed effects estimation) enlightens 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 results in Table 2 where the selection bias of fixed-effects estimate was found to be small. Again, this does not lower the interest of the Bayesian procedure, since participation equations need to be estimated in order to capture the non-linear dynamics of income and the impact on poverty.

Summing up, HIV/AIDS does not have any impact on labour income for both groups. The null impact on earnings my be interpreted as a high degree of substitutability of labour participation between the different members of the household or thanks to recomposition effects (new active members from the extended family coming in to help).On the contrary, HIV/AIDS affects urban households by increasing their participation in the transfers network. In particular, they receive more grants, which represent more than 70 percent of transfers, as shown by descriptive statistics. On the other hand, HIV/AIDS decreases the amount of transfers perceived by rural households, most probably private transfers. How to interpret the latter effect? It could be due to the stigma associated to HIV/AIDS, to the disruption of informal mechanisms of risk-sharing and private transfers caused by the increase of risk associated to the illness[28], or to a lower take-up rate of social grants by the victims of HIV/AIDS. Regressing respectively the log of social grants and the log of private transfers on the former set of regressors for rural individuals having positive transfers, we found in both cases a negative but insignificant coefficient for the HIV/AIDS variable plausibly due to the small sample size.

[28]Because non-affected households have a low gain expectations in an insurance system set up with affected households that might disappear in a close future, they rationally reduce their transfers. In another study, Marzo (2007) shows that HIV/AIDS diminishes the number of private transfers within communities with the help of a representative two-waves Zambian panel. See Jayne et al. (2006) for a global description of the data, and a study by Fafchamps (2006) on the determinants of risk sharing.

5.3 Discussion

In this subsection we come back on the derivation of former results as well as on their interpretation. It has been emphasized above that some explanatory variables correlated with HIV/AIDS were driving the correlation between income and the illness. Such observed variables are for instance the dependency ratio and the dummy for female-headed households. Taking those variables as exogenous ones, the findings above describe the impact of HIV/AIDS on labour income, as well as the potential indirect impact on transfers passing through labour income - at last found to be irrelevant for both populations since labour income levels are never affected by the illness. The problem is that we would underestimate the total impact of HIV/AIDS, had the illness an impact on households composition. If husbands are more likely to be affected than wives[29], then households may likely be headed by a female in case the male spouse passed away. Similarly, affected households may ask some relatives to join the household in order to assist the ill. Those relatives are likely to be inactive people joining to help (grand parents or children), they might increase the dependency ratio and contribute to affect available household equivalent labour income. At the same time, they might cause an increase in non labour income through social grants. In other words, the dependency ratio may capture some of the illness consequences.

In order to test these ideas, 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 ten groups the change in sero status entailed a significant change in the dependency ratio. Since morbidity and mortality rates are important among affected households, it is somehow surprising that the dependency ratio does not vary accordingly. This might be explained by a mutual support mechanism by which active people from extended family or from the community join the household to help. Some descriptive statistics give credit to this explanation: considering the newly-affected population, we find that the dependency ratio is in fact *higher* in wave 6 than in wave 1, with averages being respectively equal to 0.79 and 0.63, although medians are both equal to 0.5. Besides, affected and non-affected households receive new members in almost equal proportions (respectively 26.1% and 22.4%), but in affected households 18.1% of new members contribute to household income versus 10.8% for non-affected households. This proportion raises to 25.8% for newly-affected households. This finding could be an important explanation for the fact that the illness does not impact labour income: in reality what we estimate is not the direct impact, but the result of coping strategies, including household recomposition.

Conducting the same regressions with the gender of the household head variable, we found again that the HIV/AIDS variable was not significant, excepting for the group with strictly positive labour income in urban settings. Running a fixed-effects model for labour

[29] And widespread evidence show that they certainly are the first to be affected.

income of this group while excluding the female headed household dummy, we found a smaller coefficient for HIV/AIDS, but still not significant. Overall, it is unlikely that modeling the potential impact of HIV/AIDS on the dependency ratio and the gender of household we would capture strong effects and modify the main conclusions of the paper.

However, it is true that 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 negative effect of HIV/AIDS on urban labour income participation is almost significant. This somewhat mitigates the conclusion that the illness has no impact on labour income. A small sample size is indeed a limit of the analysis. On the other hand, data availability remains one major obstacle to conduct this kind of analysis. In our case, it is important to maximize the time span in order to obtain the maximum of credible transitions from one status to another. Some might argue that estimates based on a 10% transition rate are not credible enough given the small sample size. Nevertheless, our algorithm has given exact estimates when tested under the same conditions.

Lastly, the affected population is heterogeneous. As shown by Booysen (2004), the small group of affected households not having experienced morbidity or mortality episodes has very low poverty incidence rates. These are wealthy households, who could have a large impact on our estimates. To test this idea, we report in the second part of Table 4 the fixed-effects regressions for labour income estimated on the urban and rural population who experienced at least one case of illness or death over the period, regardless from their serostatus. This population represents 74.4% of total population. Conclusions are similar to previous ones since the HIV/AIDS variable is not significant. We refrain from running regressions on a smaller population due to small sample size and increased risk of selection bias.

6 The Impact of HIV/AIDS on Chronic and Transient Poverty

In this section we address the impact of HIV/AIDS on poverty. A Monte-Carlo experiment is run to generate the distribution of poverty indicators conditional on suffering from HIV/AIDS or not. We draw in the empirical distribution of idiosyncratic residuals, whose cumulative distribution functions are computed with a Kaplan-Meyer procedure. That enables us to reconstruct 1000 counterfactual levels of labour and non-labour income, from which we derive total income and poverty measures.

We define here our measures of chronic and transitory poverty. Let y_i represent total simulated income, $\hat{y}_i$ its average, z the poverty threshold, and P^α be the Foster-Greer-Thorbecke (1984) class of poverty measures. Total poverty of household i is then the expectation over time $P_i^\alpha = E\left[\left(\frac{z-y_{i,t}}{z}\right)^\alpha 1_{y_{i,t}<z}\right]$. Following the approach initiated by Jalan and Ravallion (2000), we define chronic and transitory poverty for an household i as

the expectation over time respectively of

$$\begin{aligned} C_i^\alpha &= \mathrm{E}\left[\left(\frac{z-\hat{y}_i}{z}\right)^\alpha 1_{\hat{y}_i<z}\right] \\ T_i^\alpha &= P_i^\alpha - C_i^\alpha \end{aligned}$$

In practice we consider two cases with $\alpha \in 0, 2$ in order to capture not only prevalence but also the intensity of poverty - the latter index (the square poverty gap index) taking into account inequality.

This procedure allows us to decompose the economic consequences of HIV/AIDS transmitted through the distinct channels - labour income levels, labour market participation, and their counterparts for non-labour income. In each case one compares the outcomes of two counterfactual populations, one having a HIV/AIDS variable equal to 1 and a reference group where it is equal to 0. We use the point estimates for all coefficients except for HIV/AIDS when the latter variable is not significant: in this case the coefficient is taken equal to zero.

Table 5 and last figure sum up the main results: HIV/AIDS does not have any impact on labour income; it decreases slightly chronic poverty in the urban setting because of the increased probability of receiving (public) transfers. In 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.

Importantly, Table 5 shows that most of the poverty increases are permanent: transient poverty varies marginally for the urban population, and it even decreases for the rural population. Considering the squared poverty gap gives the same kind of conclusion: the increase of poverty due to HIV/AIDS is permanent.

Last, it is worth mentioning that those results would be somehow modified if one had accounted for the distribution of resources within the household. As quoted before, Ainsworth and Dayton (2003) report a redistribution in favour of the ill. Hence a sizeable proportion of members from households above the poverty threshold may be in fact below this threshold. As this effect works also the other way around, it is hard to assess the influence of income repartition within the household. We leave this question opened for further research.

7 Conclusion

In this paper we have analysed the causal impact of HIV/AIDS on poverty. Using a Bayesian framework, we have introduced an econometric framework that accounts for both self-selection and endogeneity effects. A causal analysis of the economic consequences of HIV/AIDS becomes feasible. First, we find that the illness does not have any final significant impact on household labour income, which shows that households manege to keep the overall labour supply constant despite the illness. In this sense, we do not seem to be estimating the short impact of AIDS, but rather the result of coping strategies: the null impact on earnings is interpreted as a high degree of substitutability of labour participation between the different members of the household or thanks to recomposition effects (new active members from the extended family coming in to help). Further work could focus on individuals as unit of analysis, taking for granted that the illness, in absence of ARV treatments, causes progressive physical deterioration and eventually death. This would eliminate the recomposition effect. The small sample size might as well partly explain this result. Second, we find much heterogeneity between urban and rural populations regarding the impact of HIV/AIDS on non-labour income. The illness increases the probability of receiving transfers for the urban population. In particular, we deduce that they increasingly receive public grants due to their serostatus, since access to basic services is easier in urban setting and since public grants represent more than 70 percent of transfers for this group. Nevertheless, this effect is small in terms of chronic poverty reduction. On the other hand, HIV/AIDS decreases the amount of transfers perceived by rural households. Since participation in public transfers remains the same, this reduction in level likely concerns private remittances. This could be caused by stigma largely associated with AIDS, and bringing to the disruption of informal safety nets. Since affected households are condemned and since AIDS is a reflection of immoral behaviour (or of supernatural punishment), the gain expected from including them in a transfers scheme in terms of reciprocity (and public recognition and esteem) are diminished. In consequences, transfers toward affected households are reduced. The illness causes a huge 50% increase of chronic poverty among rural population while transitory poverty decreases, a sign that the illness involves a permanent fall in household equivalent total income. In a context of decreasing private support in terms of transfers, public support should be strengthen, together with overall access to social basic services.

References

Beffy, M., Buchinsky M., Fougere D., Kamionka T. and F. Kramarz (2006). Th Returns to Seniorty in France (and Why they are Lower than in the United States), CEPR wp 5486.

Coate, S. and M. Ravallion (1993). Reciprocity without commitment : Characterization and performance of informal insurance arrangements. Journal of Development Economics. Vol. 40(1), pages 1-24.

Fafchamps, M. and F. Gubert (2007). Risk Sharing and Network Formation. American Economic Review Papers and Proceedings, forthcoming.

Heckman, J.J. (1979). Sample Selection Bias As a Specification Error. *Econometrica*, 47(1), 153-162.

Hulme, D. and A. Shepherd (2003). Conceptualizing Chronic Poverty. *World Development*, vol. 31(3), pp.403-424.

Hyslop, D. (2003). State Dependance, Serial Correlation and Heterogeneity in Intertempora Participation Behavior: Monte Carlo Evidence and Empirical Results for Married Women. *Econometrica*, 67(6), p. 1255-1294.

Jensen, RT. (2004). Do private transfers 'displace' the benefits of public transfers? Evidence from South Africa. Journal of Public Economics vol. 88 (1-2), pp. 89-112.

Ligon E., Thomas J. and T. Worrall (2000). Mutual Insurance, Individual Savings and Limited Commitment. Review of Economic Dynamics, vol. 3(2), pages 216-246.

Marzo, F. and F. Murtin (2006). The Impact of HIV/AIDS on Risk-sharing Networks in Zambia, mimeo.

Maytra, P. and R. Ray (2003). The effect of transfers on household expenditure patterns and poverty in South Africa. Journal of Development Economics, Vol. 71(1), p.23-49.

Murtin, F. (2006). Inequality of Opportunity in the United States: a Bayesian Estimation of Correlated Random Models with Time-Constant Regressors, mimeo.

Nobile, A. (1998). A Hybrid Markov Chain for the Bayesian Analysis of the Multinomial Probit Model. Statistics and Computing, 8, 229 - 242.

Ravallion, M. (1996).Issues in Measuring and Modelling Poverty. The Economic Journal, Vol. 106, No. 438, pp. 1328-1343

Townsend, R. (1994). Risk and Insurance in Village India. Econometrica, 62(3), pp. 539-91.

B Tables

Table 1 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

Table 2 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 (0.24)	-	0	-	6.30	3.88	-	-
Fixed-effects estimates[1]	-0.89 (0.12)	-	0.38	-	0.81	0.64	-	-
Hybrid Gibbs	-1.01 (0.10)	-1.039 (0.15)	0.56 (0.05)	0.56 (0.05)	0.97 (0.04)	0.99 (0.13)	0.43 (0.15)	1.18 (0.26)

[1] on participating population only.

Table 3 - Bayesian Estimation of HIV/AIDS Impact

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

Regressions include year dummies

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)
$\rho_{b_i,HIV/AIDS}$	-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.

Table 5 - HIV/AIDS and Poverty Microsimulations

	Urban		Rural	
Labour Income	**Average Income**	**Participation Rate**	**Average Income**	**Participation Rate**
Reference Group	525	73	168	47
Affected Households	525	73	168	47
Transfers				
Reference Group	246	59	187	86
Affected Households	284	74	119	86
Total Income				
Reference Group	772	96	355	95.4
Affected Households	810	98	287	95.4
Headcount Poverty Rates (in %)	**Chronic**	**Transitory**	**Chronic**	**Transitory**
Reference Group	17.5 (1.7)	9.6 (1.6)	41.5 (2.9)	11.7 (2.4)
Affected Households	15.6 (1.7)	9.3 (1.5)	59.8 (2.8)	4.6 (2.3)
Squared Poverty Gap	**Chronic**	**Transitory**	**Chronic**	**Transitory**
Reference Group	0.033 (0.005)	0.069 (0.005)	0.240 (0.013)	0.184 (0.011)
Affected Households	0.024 (0.004)	0.054 (0.004)	0.338 (0.013)	0.163 (0.010)

C Figures

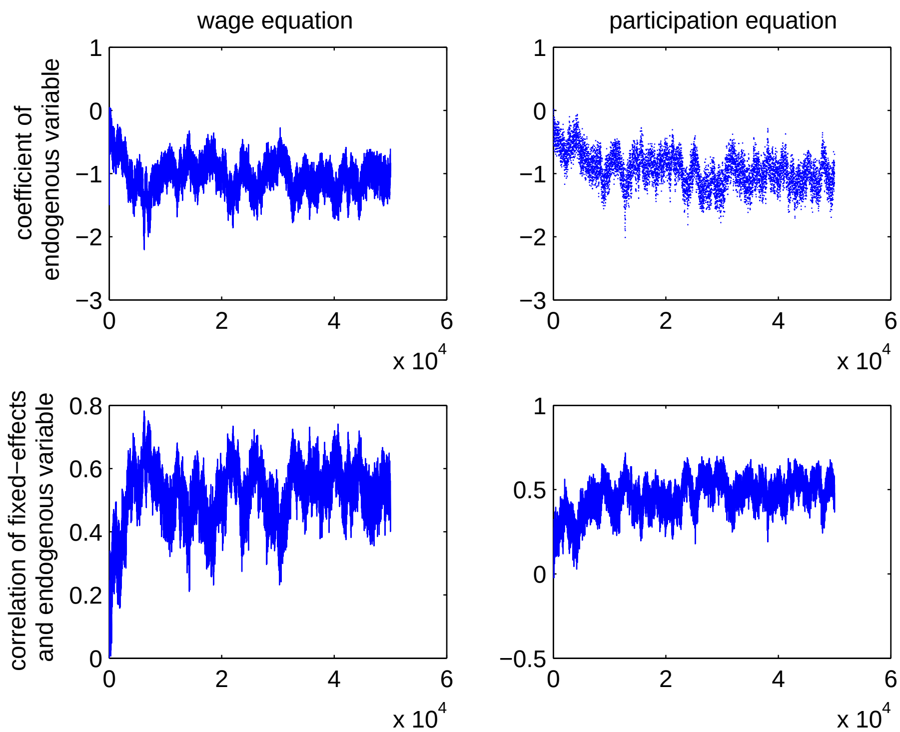

Figure 1: Convergence of the Hybrid Gibbs Sampling on Simulated Data

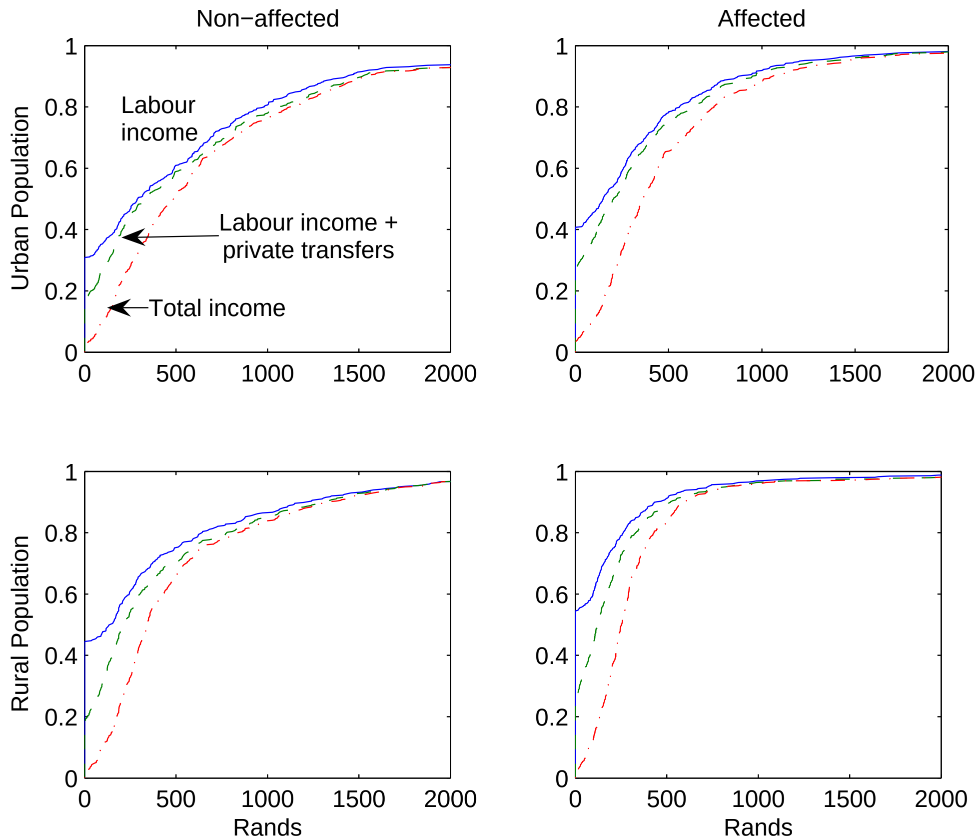

Figure 2: Cumulative Distribution Functions for affected and non-affected households - urban and rural populations

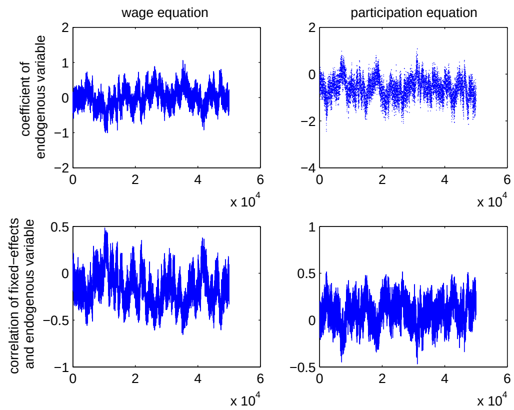

Figure 3: Bayesian Estimation of labour income - Urban Population

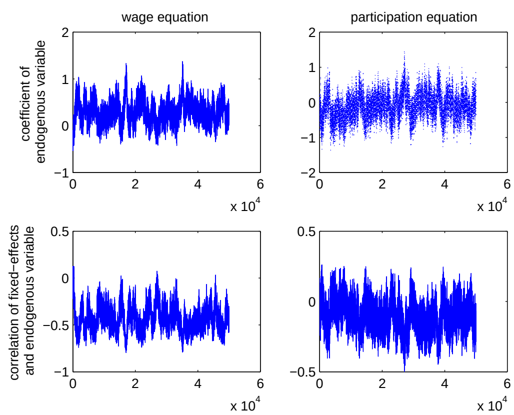

Figure 4: Bayesian Estimation of labour income - Rural Population

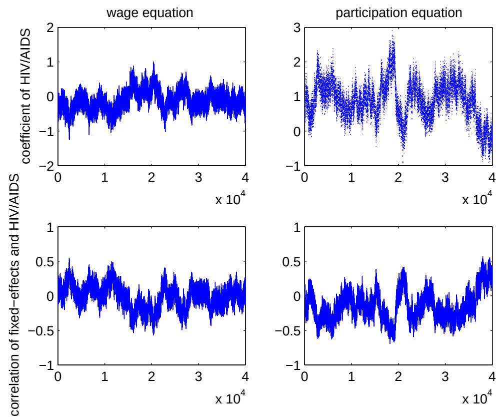

Figure 5: Bayesian Estimation of Non-Labour Income, Level Equation - Urban Population

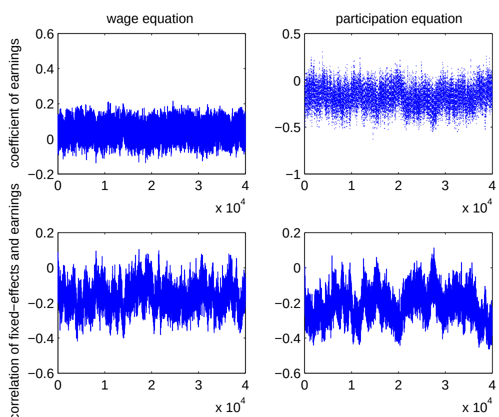

Figure 6: Bayesian Estimation of NonLabour Income, Participation Equation - Urban Population

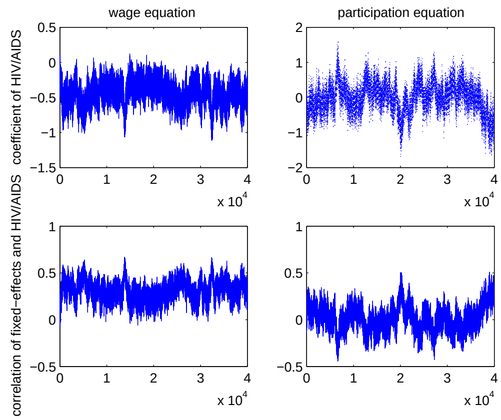

Figure 7: Bayesian Estimation of Non-Labour Income, Level Equation - Rural Population

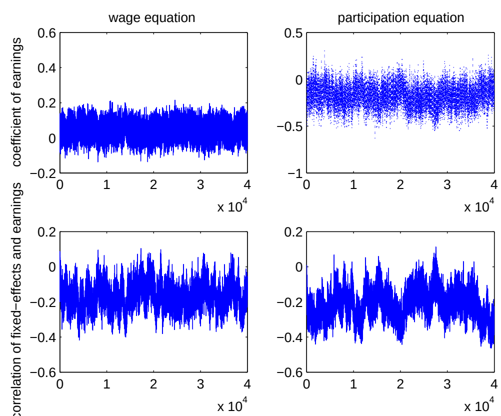

Figure 8: Bayesian Estimation of NonLabour Income, Participation Equation - Rural Population

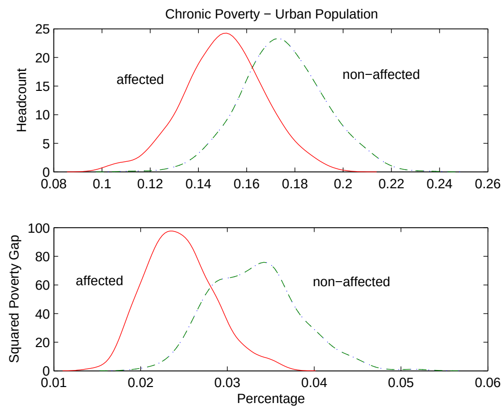

Figure 9: Microsimulation of Poverty Rates - Urban Population

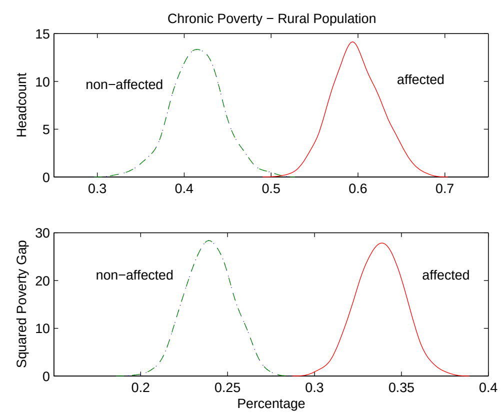

Figure 10: Microsimulation of Poverty Rates - Rural Population

D The algorithm

We use the following priors for Θ :

$$\begin{array}{rcl}
\beta & \rightsquigarrow & \mathcal{N}_K\left(\beta^0, B_0\right) \\
\rho & \rightsquigarrow & \mathcal{U}([-1;1]) \\
\varepsilon_i \mid D_\varepsilon & \rightsquigarrow & \mathcal{N}_2\left(0, D_\varepsilon\right) \\
\sigma^{-2} & \rightsquigarrow & \mathcal{G}\left(\frac{\nu_0}{2};\frac{\delta_0}{2}\right) \\
D_\varepsilon^{-1} & \rightsquigarrow & \mathcal{W}_2\left(\rho_0; R_0\right)
\end{array} \tag{3}$$

The algorithm is the following :

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

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

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

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 :

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

where π is the posterior distribution of ρ

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

$$\rho^{(j+1)}=\begin{cases}\rho^{(c)} & \text{if } r \leq \alpha \\ \rho^{(j)} & \text{otherwise}\end{cases}$$

d Define

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

3. Sample

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

4. Sample

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

and define the specific effects and their variance

$$\begin{aligned} b_i^{(j+1)} &= \bar{X}_i^* \lambda^{(j+1)} + \varepsilon_i^{(j+1)} \\ D^{(j+1)} &= \text{Var}\,(\bar{X}_i^* \lambda^{(j+1)}) + D_{\varepsilon}{}^{(j+1)} \end{aligned}$$

5. Sample

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

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

6. Go to 1