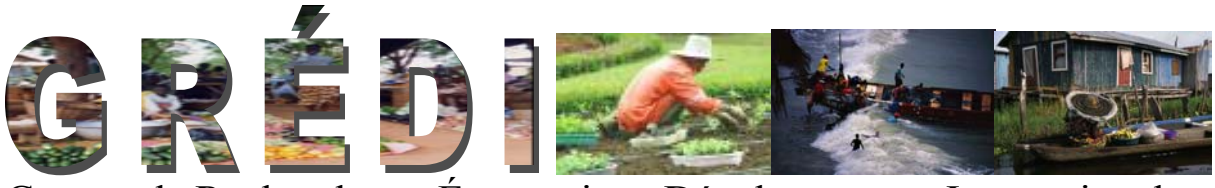

Groupe de Recherche en Économie et Développement International

Cahier de recherche / Working Paper

07-08

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

Federica Marzo

Fabrice Murtin

UNIVERSITÉ DE SHERBROOKE

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

Federica Marzo, Fabrice Murtin
CREST-INSEE

March 30, 2007

Abstract

In this paper we assess the causal impact of HIV/AIDS on monetary poverty using a panel data-set from South Africa and modeling the consequences of the illness on both earnings 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 to the labour market or to transfers networks. We solve both of them by proposing an original 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 very well-behaved. Splitting the sample into urban and rural population, we show that HIV/AIDS has a significant but moderate impact on poverty for urban population, because transfers partly compensate the fall of earnings entailed by the decrease in labour market participation. On the contrary, HIV/AIDS has an important impact on poverty for the rural population because it causes a fall of transfers. Surprisingly the effect on earnings is not significant . We argue that those results can be explained by the existence of an efficient public safety net in urban settings, while in contrast private transfers are subject to moral hazard and imperfect commitment that characterize risk-sharing in rural settings.

1 Introduction

In 2003 UNAIDS estimated that the number of people living with HIV/AIDS in the world was equal to 40 millions, 95% of them belonging to developing countries and more than 70% of them living in Sub-Saharan Africa. In those countries, the epidemics has had a dramatic impact on life expectancy, which has fallen to the level of the fifties according to UNDP (1997). Many negative externalities on growth and socio-economic development in general have been pointed at: the decrease of the labour force, the decrease in savings and investments, the reduction of school enrollment and teaching staff, the collapse of family and community structures (Marzo-Murtin 2006).

With respect to the labour market, the expected effects of the illness are progressive physical deterioration, resulting in decreasing productivity and reduced participation of affected workers, leading to a direct income loss and increased poverty. However,

affected households might adopt some coping strategies in order to face the economic consequences of the illness: for instance they might use their savings, sell assets such as livestock, or receive enhanced financial support from their relatives or their community. Those coping strategies might mitigate the increase of poverty due to HIV/AIDS.

The aim of this paper is precisely to analyse the relationship between the illness and poverty, gauging its consequences on both transitory and chronic - permanent - poverty. Such a decomposition is useful because as stated above, the combination of different coping strategies might result in different outcomes in the short-term and in the long-term. Coping strategies might not be sustainable on the long-term, or even have a negative effect on permanent income.

In practice, we decompose total income into earnings and non-labour income such as remittances or grants. Indeed, South Africa has built an important network of social grants, which represent a sizeable portion of total income especially for poor households. As the impact of HIV/AIDS might be positive on non-labour income, it is important to disentangle those two sources of income.

From an empirical perspective, the paper introduces an important methodological innovation. First, we assess the impact of HIV/AIDS on the levels of earnings and non-labour income taking into account selection effects. As emphasized by a large body of research[1], unobserved characteristics affecting participation to the labour market or to transfers networks might preclude the econometric analysis if they are correlated with some observable characteristics. Therefore, it is crucial to assess jointly the impact of the illness on both participation and income level.

Second, we account for the endogeneity of HIV/AIDS. Indeed if the illness increases poverty through various channels, it is likely that some unobserved factors affect both the likelihood of being affected by the desease and the standard of living, the sign of the correlation between those factors and the HIV/AIDS dummy variable remaining unclear. For instance, migrations may be a source of diffusion of the epidemics, but also a source of wealth because migrants generally send some remittances home. In that case the sign should 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.

The two econometric problems described above are major obstacles to the economic research on HIV/AIDS and to the derivation of credible public policies. In order to deal with those problems, we use a South-African household survey spanning over 6 years. The longitudinal dimension allows us to introduce some fixed-effects, hence to account for endogeneity of the HIV/AIDS variable. As we introduce fixed-effects both in the income level equation and the participation equation, and allow for non-null correlation between those two fixed-effects variables, we tackle the selection problem since unobserved variables affecting participation can be correlated with any observable variable via the correlation of the two fixed-effects variables.

Such an 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 simpler to simulate a distribution via Monte-Carlo Markov Chains methods

[1] see Heckman (1979)

(MCMC) rather than finding the mode of a distribution via maximization algorithms[2]. Beffy et al. (2004) and Murtin (2005) provide some studies based on such bayesian framework.

In this paper we use a modified version of the Gibbs sampling algorithm introduced by Nobile (1998), called the hybrid Gibbs sampling. The idea behind such class of algorithms 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 model explicitely the correlation between fixed-effects and observed variables, which dramatically fastens the convergence speed of the classical Gibbs sampling algorithm.

As a result, we find that HIV/AIDS has completely different consequences on urban and rural populations. With a negative impact on labour market participation compensated by an increase of participation in transfers networks, the impact of HIV/AIDS on poverty is moderate for the urban population. We provide evidence that the existence of public grants explain this result. Although the illness surprisingly does not have any causal impact on earnings for the rural population, it does cause a fall of transfers levels, increasing the chronic poverty rates by a huge 50%. Interestingly, endogeneity drives entirely the spurious correlation between earnings or transfers and HIV/AIDS. This effect on transfers is interpreted as a disruption of risk-sharing networks within the communities, which are subject to moral hazard as described by Coate-Ravallion (1998), Ligon-Worral (2000), Fafchamps (2002) and others.

The paper is organized as follows: fist section encompasses a brief description of the literature on poverty and HIV/AIDS. Then we introduce the econometric framework, test it on simulated data having similar characteristics to actual one, and compare it with traditional estimators. In a third section we describe the data and the results, and assess in section 4 the causal impact of HIV/AIDS on transient and chronic poverty. Last section concludes.

2 HIV/AIDS and poverty in the literature

This paper represents one of the first attempts to analyse the impact of HIV/AIDS on poverty. HIV/AIDS impoverishes people, this is a fact (Barnett, Whiteside (2002)). Yet the channels through which the illness affects households are numerous, and it is convenient to disentangle direct impacts from indirect ones.

Direct impacts are the consequences in terms of morbidity and mortality. Even if morbidity and mortality are spaced in time, we can reasonnably consider that their consequences are short-term. Indeed, 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). These direct economic consequences for the household are a decrease in productivity of the ill and most of the time of its entourage, consequently leading to a sharp decrease in household income. As a matter of fact, this fall of income can reach about two thirds of mean income (Morris, Burdge and Cheevers (2000)). From

[2]Simulated maximum likelihood is indeed a possible alternative. Hyslop (2003) achieves the inference of a dynamical logit model with fixed-effects. But the dependance to initial conditions and maximization problems occuring with competing sources of correlation make this estimation very delicate.

an aggregate perspective, this decrease in productivity is likely to have dramatic consequences on the labour force since most of the people living with HIV/AIDS belong to the 15-35 age band.
Indirect impacts derive from household reactions to the illness, namely the coping strategies of the household. Importantly, they aim at providing immediate relief but often have negative effects on the long-term. For instance, households redistribute resources in favour of the persons living with HIV/AIDS, possibly entailing malnutrition for the other members of the household (Ainsworth and Dayton (2003)). Moreover, they often use past savings or sell assets such as cattle and livestock, furniture, work instruments. With production capacity fragilized and exhausted savings, consumption will start decreasing, causing an increase in malnutrition. Another long-term consequence worth mentionning concerns the education sector, where both supply and demand are affected by the epidemics: on one hand the teaching staff has been seriously struck by the illness, recording one of the highest prevalence rate in many countries; on the other hand, affected households tend to withdraw children from school, sending them to work or look after the ill[3].
Overall, the channels through which the economic shock is vehiculed are the key point of the analysis. In this paper we will disentangle the direct impact passed by earnings from indirect impacts going through transfers and remittances. We will also analyse the economic consequences on both the short and long-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. For that purpose, it is necessary to tackle two key econometric problems, endogeneity and selection bias, relying on an original framework that is described below.

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 a system of two equations assuming gaussian residuals

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

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

[3]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

variables. The former system can be written as a linear gaussian panel model

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

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 dependant 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, one assumes that the conditional distributions of $y_{i,t}^*$ and $e_{i,t}^*$ are independant, 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 0. 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. Then one would like to account for endogeneity of the observed variables, so that fixed-effects have mean zero, but not necessarily conditional zero mean, conditionally on observed variables. Briefly, E $[b_i|X_i]$ 0. As described in Murtin (2006), the correlation between fixed-effects and endogenous variables must be modeled if one is to expect 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[4]. This is however not our case 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

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

(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 $e_{i,t}^*$, 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 $]0, +\infty[$. 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 $(e_{i,t}^*)$.

Let us describe now 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[5] $p(Y|\Theta)$ 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 as vague as one wishes.

[5] namely, the likelihood

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 [6] conditionnally 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[7]. As the posterior distribution of the correlation ρ cannot be written in closed-form, one simulates 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 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

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

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 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%. This is still sizeable. Even if that point estimate is not significatively different from the true value, this bias will contaminate the poverty simulations. Again, the direction of the bias was expected since the correlation between the endogenous variable and unobserved determinants of participation

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

[7] see Tierney (1994)

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. A large number of iterations, namely 50 000, is needed to obtain convergence, which is validated graphically. Therefore, one obtains 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[8].

4 Estimation

4.1 The Survey

The impact of HIV/AIDS on poverty is studied with the help of a panel of affected and non-affected households. A survey on the quality of life and the ressources of households was conducted approximately every six or seven months in two districts belonging to the Freestate province[9].

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 (waves) five and six of the study were completed in July/August 2003 and May/June 2004 respectively. Thus, the data spans a period of at least three years.

I the survey 331 households and 1 173 individuals constitute a balanced cohort with data available at each wave. The survey has two main characteristics: the selection process of affected and non-affected households, and the large heterogeneity between rural and urban households.

Affected households are those in which, at the time of the interview, someone 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 ther serostatus. This choice has been respected by fieldworkers. 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

[8]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 such a present case the algorithm performs well without any instrument, but it could be possible that with smaller time-dimension and low levels of within variance such an instrumentation becomes useful.

[9]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".

thanks to the selection process[10]. 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[11].

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 Rands. It can be defined as an urban setting[12]. In contrast, the rural magisterial district of Witsieshoek, which is within the boundaries of the former Qwaqwa, is the poorest in the Free State province and is ranked among the poorest in the country. The headcount poverty ratio in this district is 0.69, while average monthly household expenditure amounts to 807 Rands. 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.

To conclude this section, in 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[13]. Yet, the results reported in these pages, albeit based on a relatively small sample, do present a telling picture of the socio-economic impact of the HIV/AIDS epidemic.

4.2 Descriptive statistics

The sample is composed of 167 urban households and 166 rural ones, among which the proportion of affected households is by construction almost similar and about 55%. First, there are important differences between urban and rural populations depicted on Table 1. On average urban households have one year of schooling more than rural ones and earn 40% more. 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 smaller poverty rate[14]: 23% versus 33% for a 250 Rands threshold. Looking at unemployment over the whole period confirms the vulnerability of the rural population: in terms of percentage of total observations, the unemployment rate amounts 37.3% over the six waves in the urban population and 50.3% in the rural one.

Second, HIV/AIDS splits clearly each group of population into two separate subgroups.

[10]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.

[11]households originally classified as non-affected but who declared in the second wave having experienced deaths or illness in the recent past were reclassified as "affected".

[12]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.

[13]see for instance Booysen-Arntz (2003)

[14]for the headcount ofchronic poverty, see below.

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 versus 54% for the non-affected. Most of this increase can be attributed to grants (public aid and charity). 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 earnings 49% lower. Both subgroups are comparable regarding transfers, and total income is 40% lower among affected households.
To sum up, HIV/AIDS seemingly entails a fall of earnings partly compensated by transfers among urban households, and a large drop of earnings with unchanged transfers among rural population. Of course, these are only correlations, and next section is going to show that some of them reflect a causality, with some others being spuriously driven by endogeneity.

4.3 Procedure and results

As a starting point it is important to tackle the issue of simultaneity: potentially, earnings and non-labour income are simultaneously determined if households are rational and have expectations on both sources of income, so that earnings may enter the non-labour equation as a regressor, and on the contrary non-labour income may enter the earnings equation. We argue here that the latter case is not plausible.
First, in the empirical literature many authors have made a similar assumption: studying the determinants of remittances and earnings, Maytra and Ray (2004) exclude remittances from the earnings equation.
Second and more importantly, this assumption is consistent with the mechanism we have in mind: HIV/AIDS is an income shock, which consequences are tackled through various coping strategies including remittances and grants. In other words, the increase of non-labour income is posterior to the decrease of earnings following the emergence of the illness. Of course, one could argue that if correctly expected, the latter increase may have a negative impact on earnings via a substitution effect. According to us, this view neglects the fact that most of households dwell with poverty, and would hardly diminish their earnings because they receive some money in support of the ill: in fact, their labour supply is plausibly constrained.
To sum up, we are interested in the direct impacts of HIV/AIDS on both earnings and non-labour earnings, and in the indirect impact channeled by the income substitution from earnings to non-labour income, namely that non-labour income increases when earnings fall, which is interpreted as a coping strategy. We neglect the indirect negative impact of HIV/AIDS on earnings channeled by the substitution from non-labour income to earnings[15], and exclude non-labour income from the earnings equations.
Then, the latter Bayesian procedure is used to assess the impact of HIV/AIDS on earnings, while specifying an unique endogenous variable, the HIV/AIDS dummy variable[16]. Results are in the first part of Table 3. For the urban population, we find that the illness does not have any impact on the level of earnings, but do decrease participation.

[15] If it exists, it is encompassed in the direct effect
[16] in this paper one is only interested in the causal impact of this particular variable on poverty.

For the rural population, HIV/AIDS does not have any impact on the level of earnings, nor on participation. This is a surprise because descriptive statistics showed that affected households had much lower earnings and participated less. This is explained by other observed and unobserved characteristcs correlated with HIV/AIDS and having a negative impact on earnings. The negative correlation between fixed-effects and HIV/AIDS in the level equation explains the seemingly negative correlation between earnings and the illness. For participation, the dummy for households with a female head explains the same result.

On a second step, one estimates the non-labour income equations. We specify two endogenous variables, HIV/AIDS and the latent earnings variable estimated before. As a result, fixed-effects are a linear combination of those two variables' within averages and another independant component. Table 3 depicts the results in its second part. It turns out that the dummy HIV/AIDS has a positive impact on participation for the urban population. For the rural population, it has a negative impact on transfers' level, 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 their status, the causal impact of the illness was hidden. The higher participation of affected households in the transfers networks can be interpreted as the feedback of earnings on participation. But again, this is not a consequence of HIV/AIDS, which has no impact on rural earnings, but of the fact that affected households are poorer than non-affected households for observed reasons - education, female head - and unobserved ones.

A comparison with Table 4 is enlightning on 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 very close, and both non significant for urban and rural populations. Therefore the selection bias is negligeable. 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 have to be estimated as emphasized before.

In sum, HIV/AIDS affects urban households by decreasing their participation on the labour market, but this income shock is compensated by an increase of participation to the transfers network. In particular, they receive more grants as shown by descriptive statistics. On the other hand, HIV/AIDS has no impact on rural earnings, but decrease the amount of tranfers received by affected households. The null impact on earnings is interpretated as a high degree of substituability of labour force between the different members ofthe household. Then, the impact on transfers is likely a disruption of informal mechanisms of risk-sharing caused by the increase of risk associated to the illness. In a rural setting households partly mutualize income as shown by Townsend (1995), but this system is subject to moral hazard as emphasized by Coate-Ravallion (1996), Ligon-Worrall (2000) and Genicot-Ray (2002). Because non-affected households have low gain expectations in an insurance system set up with affected households that might disappear in a close future, they rationnally reduce their transfers. In a companion paper, the authors analyse the probability of being part of a transfers network with the help of a representative two-waves Zambian panel. They find that the illness nega-

tively affects the likelihood of a transfer[17].

4.4 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 households headed by females. Taking those variables as exogenous ones, the findings above describe the direct impact of HIV/AIDS, as well as the potential indirect impact on transfers going through earnings - at last found to be irrelevant for both populations since earnings levels are never affected. 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, then households may be more likely headed by a female in case the ill people have passed away. Similarly, affected households may ask to some relatives to join the household in order to assist the ill. Those relatives are likely to be old people out of the labour force, who may lower the dependency ratio and diminish the equivalent income.
To test these ideas, we regressed the dependency ratio on the HIV/AIDS variable, quadrics in education and age and time dummies in a fixed-effects model for different groups: the urban population, the urban population with positive earnings, with null earnings, with positive transfers, with null transfers, and similarly for the rural population. For none of these ten groups we found that the change of serologic status entailed a significant change in the dependency ratio. So we rule out this potential indirect impact. Conducting the same experience with the sex of household head variable, we found the same result excepted for the group with stricty positive earnings in urban settings. Running a fixed-effects model on earnings for this group while excluding the female head dummy, we found a smaller coefficient for HIV/AIDS, but still not significant[18]. In that case, there is no reason of changing the main conclusions of the paper, in particular that HIV/AIDS has a null impact on earnings levels.
Next, the small sample size (about 170 households in each group) could be somewhat disturbing and entail large standard errors in estimates, maybe turning down the significance of some effects[19]. This is indeed a limit of the analysis. On the other hand, it is to our knowledge one of the very few African dataset with such a time span where the affected persons are well identified. It is important to maximize the time span in order to obtain the maximum of credible transitions from one state to another. Some might argue that estimates based on a 10% transition rate are not credible given the small sample size. We invite them to have another look at the test of the algorithm, which delivers exact estimates in the same conditions.

[17] on the determinants of risk sharing, see Fafchamps (2002)
[18] for this regression bayesian and frequentists estimates were very close.
[19] see for instance the almost significant positive effect of HIV/AIDS on rural earnings

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

In this section we address the impact of HIV/AIDS on poverty, which can be decomposed into the respective impact of the illness on participations and levels of both earnings and non-labour income.

A Monte-Carlo experiment is run to generate the distribution of poverty indicators conditionnally on suffering from HIV/AIDS or not. We draw in the empirical distribution of residuals, which cumulative distribution functions are computed with a Kaplan-Meyer procedure. That enables us to reconstruct 1000 counterfactual levels and participations of earnings and non-labour income, from which total income then poverty measures are deduced. We use Headcount indices respectively denoted Hc and H for chronic and transient poverty, as well as those defined by Ravallion (1998) taking into account poverty intensity and denoted respectively Ic and It: chronic and transient poverty are respectively

$$\begin{aligned} Hc &= \int 1_{\hat{y_i}<z} di \\ Ht &= \int 1_{y_i<z} di - Hc \\ Ic &= \int (1-\frac{\hat{y_i}}{z})^2 . 1_{\hat{y_i}<z} di \\ It &= \int (1-\frac{y_i}{z})^2 . 1_{y_i<z} di - Ic \end{aligned}$$

where y_i represents total simulated income, $\hat{y_i}$ its average and z the poverty threshold. This procedure allows us to decompose the economic consequences of HIV/AIDS vehiculed through the distinct direct channels - earnings levels, earnings participation, and their counterparts for non-labour income. In each case one compares the outcomes of three counterfactual populations, one having a HIV/AIDS variable equal to 1 in both earnings and transfers equations, a second were this variable takes value 1 only in earnings equations, and a reference group whre it is 0 everywhere.

Table 5 and last figure sum up the main results: HIV/AIDS decrease employment by around 10% in the urban setting, and increase monetary poverty by about 6%. Hopefully, participation in transfers networks rise by about the same proportion and cancels half of the latter poverty increase. In contrast, HIV/AIDS does not have any impact on earnings in a rural setting and the first two counterfactual groups are the same. The impact on transfers, diminishing average transfers by 37%, has huge consequences on poverty: it increases its level from 46% to about 64%. This stresses the importance of negative externalities of HIV/AIDS, and the unexpected roads from the illness to poverty.

Importantly, Table 5 shows that most of poverty increases are permanent: the increase of transient poverty is only marginal for the urban population, and even decreases for the rural population. Considering intensity poverty measures give the same kind of conclusion: the increase of poverty due to HIV/AIDS is permanent, hence calls for a

permanent public policy.
Last, it is worth mentionning that those results would be somewhat modified if one had accounted for the distribution of ressources 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 gauge the influence of income repartition within the household. We leave this question opened for further research.

6 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 of regressors. Hence, a causal analysis becmes feasible. We find much heterogeneity in the process between urban and rural populations. HIV/AIDS affects participation to the labour market for the urban population but not for the rural one, and increases (resp. decreases) the amount of transfers received by the urban (resp. rural) population. We interpret these distinct mechanisms as an evidence for a public safety-net in urban settings and the fragility of risk-sharing mechanisms in rural settings.

A 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.

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.

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

	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 Earnings[1]	709	593	817	535	363	702
Participation Rate	63.7	58.8	69.0	49.7	45.7	54.3
Average Grants[1]	250	251	249	193	197	189
Participation Rate	47.2	56.7	36.8	53.2	54.1	52.2
Average Non-labour Income[1]	278	266	294	238.8	223.6	257
Participation Rate	62.6	70.0	54.4	75.2	75.8	74.6
Average Total Income[2]	625	535	724	445.2	335.5	571.6
Participation Rate	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

Table 2 Test of the Hybrid Gibbs Sampling

$\beta^{(1)}$	$\beta^{(2)}$	$\rho_{b_i^{(1)},X}$	$\rho_{b_i^{(2)},X}$	s^2	$D^{1,1}$	$D^{1,2}$	$D^{2,2}$
-1	-1	0.54	0.52	1	1	0.5	
-1.25 (0.24)	-	0	-	6.30	3.88	-	-
-0.89 (0.12)	-	0.38	-	0.81	0.64	-	-
-1.05 (0.10)	-0.99 (0.16)	0.63 (0.04)	0.69 (0.05)	0.98 (0.03)	0.99 (0.12)	0.45 (0.15)	1.22 (0.26)

Table 3 - Bayesian Estimation of HIV/AIDS Impact

	Urban		Rural	
	y^*	e^*	y^*	e^*
Earnings				
HIV/AIDS	0.029 (0.258)	-0.606 (0.413)	0.319 (0.214)	-0.055 (0.315)
Dependency Ratio	-0.143^{**} (0.066)	-0.016 (0.134)	-0.149 (0.078)	-0.449^{**} (0.139)
Education of Head	0.050 (0.036)	0.092 (0.078)	-0.035 (0.033)	-0.027 (0.067)
Squared Education	0.002 (0.002)	-0.003 (0.006)	0.008^{**} (0.002)	0.006 (0.005)
Age of Head	0.035^{**} (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	-0.352^{**} (0.091)	-0.644 (0.189)	-0.353 (0.102)	-0.407^{**} (0.202)
$\rho_{b_i,HIV/AIDS}$	-0.18 (0.17)	0.088 (0.136)	-0.46^{**} (0.11)	-0.12 (0.10)
Non-labour Income				
HIV/AIDS	-0.085 (0.249)	0.751 (0.576)	-0.458^{**} (0.156)	-0.038 (0.451)
Earnings	-0.105 (0.054)	-0.063 (0.101)	0.044 (0.044)	-0.185 (0.112)
Dependency Ratio	0.108 (0.069)	0.563 (0.152)	0.036 (0.050)	0.580^{**} (0.180)
Education of Head	-0.059 (0.041)	0.005 (0.079)	-0.016 (0.027)	-0.007 (0.079)
Squared Education	0.005 (0.003)	0.000 (0.006)	0.003 (0.002)	0.001 (0.006)
Age of Head	0.041^{**} (0.020)	-0.150^{**} (0.041)	0.019 (0.013)	-0.029 (0.039)
Squared Age	-0.000 (0.00)	0.002^{**} (0.000)	0.000 (0.000)	0.001^{*} (0.000)
Female Head	0.241^{*} (0.109)	-0.401 (0.100)	-0.048 (0.088)	-0.043 (0.212)
$\rho_{b_i,HIV/AIDS}$	-0.04 (0.15)	-0.09 (0.23)	0.33^{**} (0.10)	0.02 (0.15)

Table 4 - Fixed-effects Estimates of HIV/AIDS Impact

	Urban	Rural
	y^*	y^*
Earnings		
HIV/AIDS	-0.009	0.367^*
	(0.312)	(0.230)
Dependency Ratio	-0.170^{**}	-0.027
	(0.081)	(0.104)
Education of Head	0.071	0.019
	(0.044)	(0.040)
Squared Education	-0.002	0.000
	(0.003)	(0.003)
Age of Head	0.016	0.005
	(0.022)	(0.022)
Squared Age	-0.000	0.000
	(0.000)	(0.000)
Female Head	-0.258^{**}	-0.153
	(0.126)	(0.156)
$\rho_{b_i,HIV/AIDS}$	-0.15	-0.54
Non-labour Income		
HIV/AIDS	-0.287	-0.450^{**}
	(0.279)	(0.172)
Dependency Ratio	0.096	0.047
	(0.081)	(0.064)
Education of Head	0.024	-0.025
	(0.053)	(0.035)
Squared Education	-0.005	0.004^*
	(0.004)	(0.002)
Age of Head	0.051^{**}	0.062^{**}
	(0.027)	(0.019)
Squared Age	-0.000	-0.001^*
	(0.00)	(0.000)
Female Head	0.267^*	-0.091
	(0.160)	(0.122)
$\rho_{b_i,HIV/AIDS}$	0.09	0.30

Table 5 - Fixed-effects Estimates of HIV/AIDS Impact

	Urban		Rural	
	Average Income	Participation Rate	Average Income	Participation Rate
Earnings				
Reference Group	501	74	196	45
Affected Level	501	74	196	45
Affected Level and Participation	455	63	196	45
Transfers				
Reference Group	228	59	159	76
Affected Level	228	59	101	76
Affected Level and Participation	254	72	101	76
Total Income				
Reference Group	729	96	355	94.2
Affected Level	682	92	296.6	94.2
Affected Level and Participation	708	96	296.6	94.2
Headcount Poverty Rates	Hc	Ht	Hc	Ht
Reference Group	15.5	10.5	46.0	7.4
Affected Level	20.4	10.8	46.0	7.4
Affected Level and Participation	18.1	11.1	63.5	0.1
Intensity Poverty Rates	Ic	It	Ic	It
Reference Group	14.6	15.6	28.0	15.0
Affected Level	19.2	17.5	28.0	15.0
Affected Level and Participation	16.5	16.2	38.1	12.5

C Figures

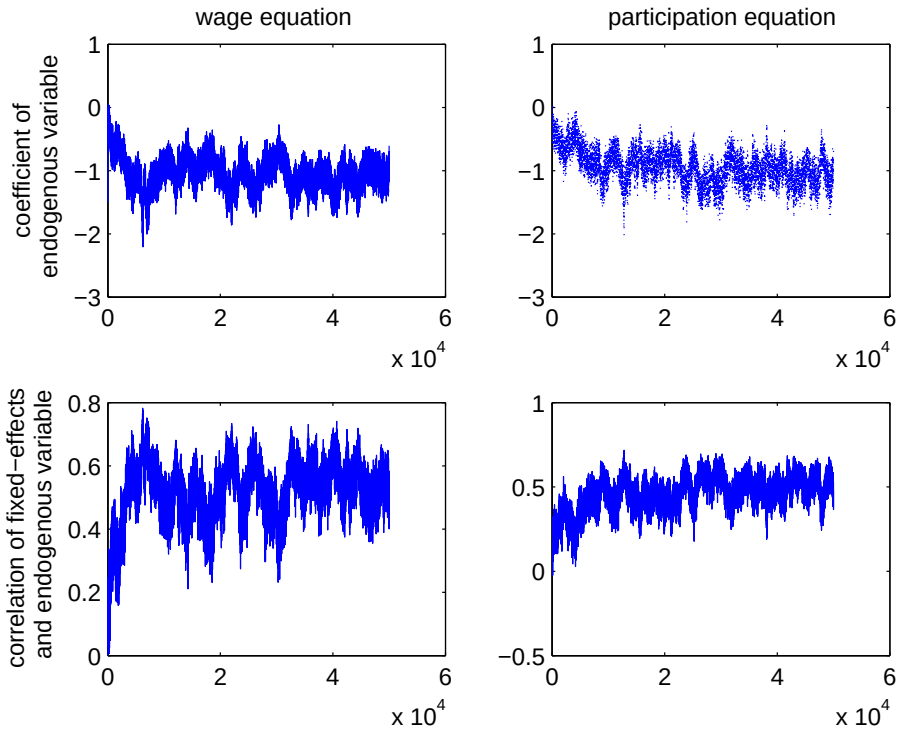

Figure 1: Convergence of the Hybrid Gibbs Sampling on Simulated Data

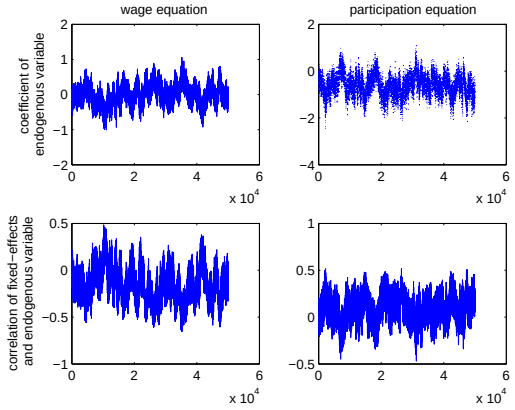

Figure 2: Bayesian Estimation of Earnings - Urban Population

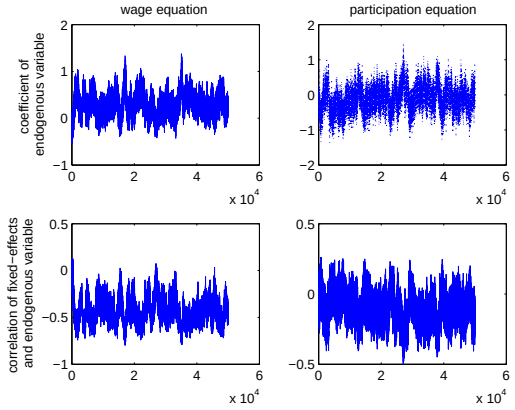

Figure 3: Bayesian Estimation of Earnings - Rural Population

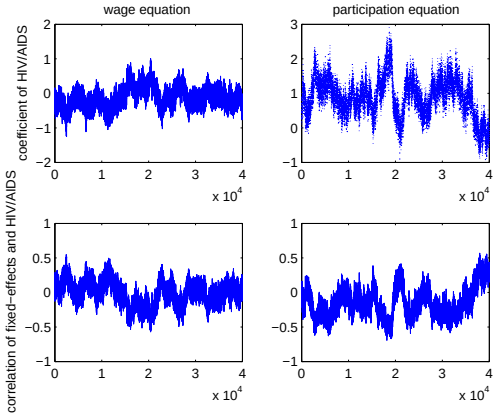

Figure 4: Bayesian Estimation of Non-Labour Income, Level Equation - Urban Population

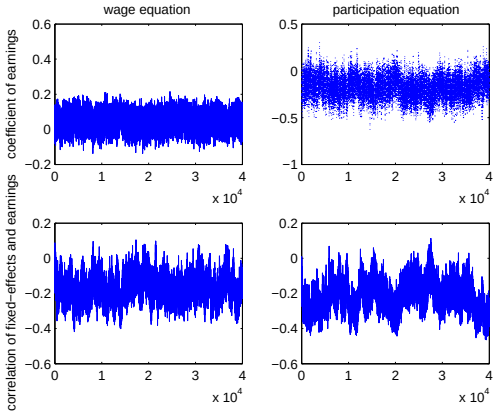

Figure 5: Bayesian Estimation of NonLabour Income, Participation Equation - Urban Population

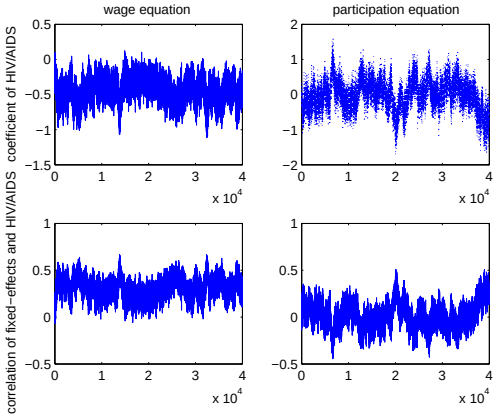

Figure 6: Bayesian Estimation of Non-Labour Income, Level Equation - Rural Population

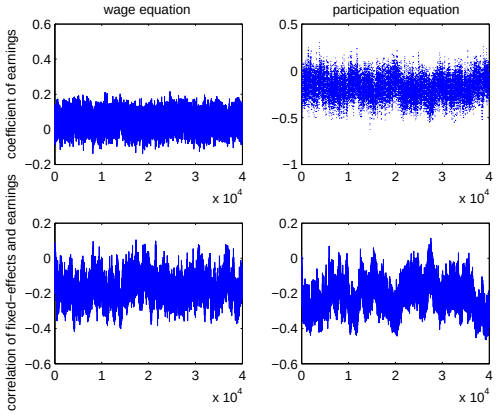

Figure 7: Bayesian Estimation of NonLabour Income, Participation Equation - Rural Population

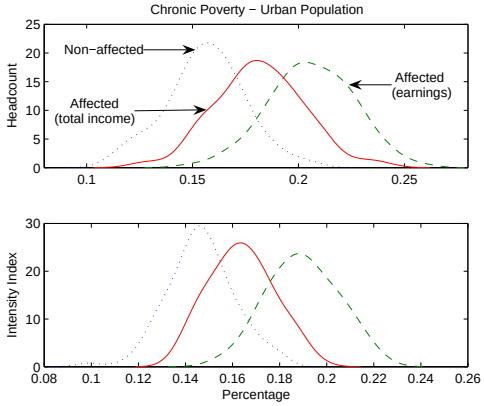

Figure 8: Microsimulation of Poverty Rates - Urban Population

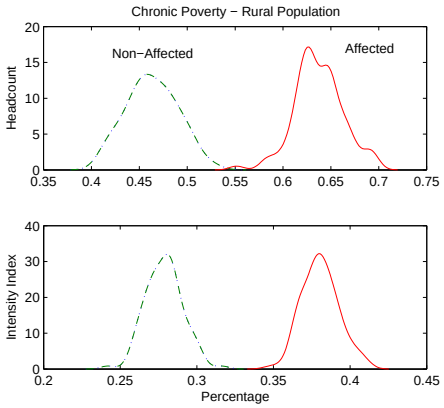

Figure 9: Microsimulation of Poverty Rates - Rural Population

D The algorithm

We use the following priors for Θ :

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

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