WORKING PAPER

5-06

Linking household income to macro data to project poverty indicators

G. Dekkers, G. De Vil, P. Willemé

July 2006

Federal Planning Bureau
Economic analyses and forecasts

Kunstlaan/Avenue des Arts 47-49
B-1000 Brussel/Bruxelles
Tel.: (02)507.73.11
Fax: (02)507.73.73
E-mail: contact@plan.be
URL: http://www.plan.be

Linking household income to macro data to project poverty indicators

G. Dekkers, G. De Vil, P. Willemé

July 2006

Federal Planning Bureau

The Belgian Federal Planning Bureau (FPB) is a public agency under the authority of the Prime Minister and the Minister of Economic Affairs. The FPB has a legal status that gives it an autonomy and intellectual independence within the Belgian federal public sector.

FPB's activities are primarily focused on macroeconomic forecasting, analyzing and assessing policies in economic, social and environmental fields.

Internet

url: http://www.plan.be
E-mail: contact@plan.be

Publications

Publications:

The Economic Forecasts
The "Short Term Update"

Planning Papers (recent publications)

giving the results of policy analyses and assessments

99 *Participations publiques dans le secteur marchand en Belgique, 1997-2003 / Overheidsparticipaties in de marktsector in België, 1997-2003*
H. Spinnewyn - January 2006

100 *Les charges administratives en Belgique pour l'année 2005 / De administratieve lasten in België voor het jaar 2005*
L. Janssen, Ch. Kegels, F. Verschueren - February 2006

Working Papers (recent publications)

3-06 *Het Europese begrotingscompromis van december 2005 - Financiële impact voor België en de andere lidstaten / Le compromis budgétaire européen de décembre 2005 - Impacts financiers pour la Belgique et les autres Etats members*
F. Hennart (Ministère de la Région Wallonne - Direction de la Politique Economique), M. Saintrain (Bureau fédéral du Plan), T. Vergeynst (Studiedienst van de Vlaamse Regering) - April 2006

4-06 *Fiscal councils, independent forecasts and the budgetary process: lessons from the Belgian case*
H. Bogaert, L. Dobbelaere, B. Hertveldt, I. Lebrun

Citations should refer to the publication, mentioning the author(s) and the date of publication.

Responsible publisher: Henri Bogaert
Legal Deposit: D/2006/7433/22

Table of contents

I. Introduction

The Belgian Study Group on Ageing of the High Council of Finance, in its Annual Report, publishes the results of research on the budgetary and social effects of ageing. In this context, the Federal Planning Bureau, in its capacity as secretariat and main research body of the Committee, has in recent years been stepping up its efforts to develop models based on socioeconomic micro data. The results of one of these models, designed to make short-term projections of poverty indicators, are presented in this paper.

The need to project poverty indicators arose from the fact that the principal Belgian data source used to study poverty, the 'Panel Survey on Belgian Households' (PSBH), was discontinued in 2002. Its successor, the EU 'Statistics on Income and Living Conditions' (SILC), is available for 2002 and 2003, and is not sufficiently comparable with the PSBH to form a homogeneous panel data set. In order to gain insight into the recent evolution of the poverty risk of elderly people, a simple micro simulation model was developed that enabled us to estimate various poverty indicators for the entire 1994-2005 period.

Micro simulation in the social sciences typically involves the adjustment of a set of data on micro-entities, be it individuals or households, to meet exogenous information. The motivation behind this modelling on the micro level is that social and fiscal policy measures intervene on the micro level, and the simulation should therefore be carried out on this level. Furthermore, micro simulation models allow simulation of the distributional effects of these developments. Unfortunately, these models usually come at a significant cost in development and maintenance time. The goal of this paper is to show how a simple model can generate acceptable results. Secondly, this paper aims at reconsidering the use of the state that an individual occupies. Most micro simulation models to date have as a common characteristic that every individual occupies one of multiple mutually exclusive states in each period of time. For example, one is either working, unemployed or out of the labour market. This state may change over time or not, but alternative states within one period are not considered and therefore do not affect the simulation results. While this may be a realistic description of a person's situation at any moment in time (insofar as states cannot be cumulated, like being employed and unemployed, or employed and retired), it is less so over the course of the time span typically covered by socioeconomic surveys (which are conducted annually at best, and often biannually). Indeed, it is far from unusual that individuals report having been in multiple states over the course of the period covered in the interview. Obviously, if the states are mutually exclusive, this implies that the individual has spent a number of months in different states consecutively. This point can be made even more forcefully when the unit of observation is the household, as in the present study (for reasons to be discussed below): households, except when they consist of a single individual, usually combine several states over the observation period, and almost invariably combine several sources of income. Consequently, the concept of the 'state' of a household may not be very useful to describe its socioeconomic and demographic position, nor the associated income. In this paper, an alternative and very simple model of static simulation is proposed and applied to the Panel Set of Belgian Households dataset.

II. Household and individual incomes in the PSBH dataset

The Panel Study on Belgian Households is a panel dataset consisting of 11 waves to date. It covers the years from 1992 to 2002 and questions describe individuals and their households in terms of demographics, employment, health, incomes and earnings, professional and other activities, housing and living conditions, possession of durables, affordability of specific goods and services and perceived welfare, time use, social relations, migration, and so forth. On the whole, the PSBH contains about 800 variables at the individual level and 400 at the household level.

The PSBH started in 1992 with 8741 adult respondents in 4438 households and this decreased to 5789 in 1997. In 1998, due to the inclusion of a new set of respondents (an additional sample to counter panel attrition), it rose again to 7015 adults in 3773 households and ended in 2002 with 5362 adults in 2959 households.

Individual respondents were asked about their annual revenues, benefits and other incomes in the previous year. The household questionnaire however, also contains a question on the total disposable (i.e. after taxes and contributions) current monthly income. Furthermore, the respondents were asked about the main sources of their household income. They were offered a choice from one or more of the following sources:
1. earnings, salary: income from work
2. earnings, profits: income from an independent activity or farming
3. pension benefits
4. unemployment benefits or layoff-premiums
5. other social benefits
6. rentals, returns from investments or savings
7. other

For a good comprehension of what follows, it is important to emphasize some limitations of the data set. First, it should be noted that, despite the fact that respondents were allowed to choose more than one income category, it is possible that they failed to report minor income sources, or were uncertain about the kind of income they received. Elderly unemployed persons, for example, who are not expected to re-enter the labour force (both officially and subjectively), may consider themselves as retired although they are officially recorded as unemployed. Second, the reported household incomes do not take into account non-financial wealth and non-financial services (like in-kind help, government-subsidized services, etc.). This may have repercussions on the measurement of poverty. For example, a non-financial asset such as a house may keep a family out of poverty by providing a non-financial benefit to its owners. Moreover, this property is often taken into account when a social service analyzes the means of subsistence of a person or family applying for a subsistence allowance. Assuming that wealth is largely concentrated in the older age groups, not taking it into account would lead to overestimation of their poverty risks. Third, the data set does not contain information on elderly people living in nursing homes, possibly influencing the poverty rate of this age group. This potential source of selection bias is likely to affect the results of the oldest old in particular, though it is a priori unclear what the effect will be. Finally, like most socioeconomic surveys, the PSBH almost certainly fails to include a representative number of respondents who live in extreme poverty.

This paper adopts the above questions on the household income as the starting point of the analysis, for several reasons. First, poverty risks and other descriptions of welfare (or the lack of it), typically uses equivalent household information at the level of the individual, thereby assuming that the income of one member of a household adds to the welfare of all members of that household. Using information from the household questionnaire therefore seems the consistent thing to do. Second, though they might be more precise, the questions in the individual questionnaire refer to the income in the *previous* year. So, wave *t* of the PSBH dataset contains individual income information pertaining to the year *t-1*. This is problematic for the method used in this paper, which links income to individual and household characteristics and several macroeconomic time series. If individual income were to be used, this would mean that the income at time *t-1* would be regressed on individual, household and macroeconomic information at time *t*. An obvious solution would have been to link the individual income information from wave *t* (pertaining to *t-1*) to the individual and household characteristics of the previous wave *t-1*. However, since macroeconomic time series vary over time but are constant in the cross-sectional dimension of the data set, giving up the final year's observations was expected to weaken the statistical link between the micro data and the macro variables.

The third and final reason for using the income data from the household questionnaire is that the reported income also includes the incomes of those household members which were not interviewed, a feature which is especially relevant for the last years of the PSBH.

III. An empirical model of household income

In order to project the evolution of household incomes over time, and (possibly) to evaluate the impact of social policy measures on poverty, one needs a model that links the observed household situation to the evolution of exogenous macro variables. Given that the socioeconomic and demographic 'state' of a household is difficult to define, a feasible alternative is to explain total household income by a set of macro income indicators that reflect the different 'micro' sources of income such as wages, pensions, unemployment benefits etc. We tried to identify macro indicators that match the seven income sources available in the PSBH, and found four that were suitable for our purposes. These are the evolution of the average labour income, the average income of the self-employed, the average pension income and the average unemployment benefits. The evolution of the remaining three income sources was modelled using a linear time trend [1].

Table 1 displays the evolution over time of the various macroeconomic per capita incomes, expressed in average annual growth rates. They have been computed by the Federal Planning Bureau, based on data from the National Accounts, the National Employment Office, the National Pension Office ('Office National des Pensions'), and the Pension Office for the Public Sector ('Service des Pensions du Secteur Public'). It should be noted that, while the survey data report net household incomes, their macroeconomic counterparts refer to gross incomes. This inconsistency is unfortunate, but probably less harmful than any attempt to convert the macro data to after-tax equivalents.

Table 1 - Development over time of nominal per capita incomes, annual growth rates *(in percent)*

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Pension benefit	2.6	2.7	1.8	1.4	2.3	1.6	3.3	3.1	3.7	4.3	1.8	4.2
Unemployment benefit	-4.0	-3.9	1.9	1.0	4.2	3.5	1.3	4.9	8.5	0.9	1.7	1.0
Gross earnings of employees	3.9	2.4	1.5	9.4	0.0	3.7	2.4	3.4	2.7	1.2	2.4	2.6
Gross income of independent workers	4.1	-0.1	0.1	4.3	2.5	2.1	6.0	1.6	-0.9	-0.0	-0.2	4.2

Source: Federal Planning Bureau.

1 The macro indicators cover the main income sources. The evolution of capital income (item 6) was modeled with a time trend, despite the fact that a macro proxy could be used. The main reason is its relatively minor importance, both as a percentage of total household income and in terms of the number of households reporting it as a source of income. Introducing an indicator for average capital income could be a future refinement of the model, although we do not expect this to affect the results fundamentally.

The proposed model can be specified as a panel data model of the following form:

$$y_{it} = \mu + \mu_i + \Sigma\, \alpha_j D_{ijt} + \Sigma\, \beta_j D_{ijt}\, x_{jt} + \varepsilon_{it} \qquad (1)$$

where
y_{it} = the logarithm of total net monthly income of household i in year t
μ_i = individual effects, also containing household-level information.
D_{ijt} = dummy variables indicating household income sources (1 if income source j is reported, 0 otherwise)
x_{jt} = the logarithm of the nominal macroeconomic aggregates per capita for j = 1, … , 4 (earnings and salaries, self-employment incomes, pension-incomes, unemployment benefits) or a linear trend for j = 5, 6, 7 (other social benefits, capital incomes, other incomes).
α_j = the intercept coefficients for each source of income (j = 1, …, 7).
β_j = the slope coefficients for each source of income (j = 1, …, 7).

The strategy to link micro income data to macro per capita averages, although appealing because of its simplicity, has some inherent limitations that should be kept in mind. In addition to the usual data limitations that affect any empirical model based on survey data, the specification of a fixed-coefficients panel model imposes specific restrictions on the kind of effects that can be detected, and on their interpretation. First, the estimates of the slope coefficients 1 to 4 ($\hat{\beta}_j$) are in fact elasticities of the individual income with respect to the associated macroeconomic per capita aggregates. One might a priori think that these elasticities should be equal to one. However, due to the fact that households typically combine several income sources, the income source dummies D_{ijt} are not mutually exclusive. As a result, the estimated elasticities with respect to the macro variables are expected to be below one. Second, since the macro variables themselves are weighted averages of different income subtypes (like different pension schemes, for instance), and since they are identical across individuals, the estimated elasticities are average effects, unable to reflect specific measures taken to improve the situation of targeted income groups. An increase in the minimum pension benefit, for example, even if it is reflected in the overall average amount, will only marginally change the incomes of retirees with extremely low pensions. Indeed, the estimated slope coefficient will ‘smear out’ the effect of the increased minimum over all beneficiaries instead of allocating it to the targeted subgroup.

Equation 1 was estimated using both the fixed effects and the random effects estimators for the individual-specific level parameters μ_i . The choice between the two estimators was based on the familiar Hausman test, which strongly rejected the null hypothesis of random effects. The fixed effects estimation results are presented in Table 2.

Table 2 - Estimation results of the fixed effects panel model of household income (PSBH, waves 3 through 11, 1994-2002)

Coefficient	Estimate	Std. Err
α1	-1.822 [a]	0.108
α2	0.211	0.316
α3	-0.772 [a]	0.139
α4	-0.361 [b]	0.154
α5	0.005	0.008
α6	0.067 [a]	0.010
α7	-0.022	0.014
β1	0.469 [a]	0.024
β2	-0.020	0.070
β3	0.243 [a]	0.040
β4	0.107 [b]	0.049
β5	0.007 [a]	0.001
β6	-0.002	0.001
β7	0.006 [a]	0.002
Const	10.922 [a]	0.006
$\sigma(\mu_i)$	0.413	
$\sigma(\varepsilon_{it})$	0.199	
ρ	0.811	(fraction of variance due to μ_i)
Diagnostic statistics:		
Number of observations	36314	
Number of households	8299	
R^2 (overall)	0.396	
Model test F(14,28054)	261.31 [a]	
Fixed effects F(8313, 28054)	13.33 [a]	
Hausman test $X^2(14)$	3141.85 [a]	

[Note: [a, b] denote significance at 1% and 5% respectively, two-sided tests.]

The results in Table 2 confirm the link between the micro-level household incomes and the evolution of the corresponding macroeconomic averages. The estimated slope coefficients are positive and below unity as expected, except for the self-employed (*β2*) and capital income components (*β6*). The former is not significantly different from zero, while the latter (insignificant as well by the usual standards) possibly reflects declining capital income due to the general decrease in interest rates over the observation period. Most of the coefficients are strongly significant. Taking into account that one-sided tests would be appropriate to test for positive effects, most null hypotheses of zero slopes can be rejected at the 5% level or better (note that the table reports two-sided significance levels). Some of the coefficients appear to be rather low, especially the coefficient of the unemployment benefit indicator (β_4). One explanation is that the fourth income component in the PSBH includes lay-off premiums, which are unlikely to be correlated with the average unemployment benefit. Another explanation is that unemployment spells in most households usually last less than a year, and almost always less than two years. As a result, the link with the evolution of macro unemployment benefits gets rather weak. If, for example, four households each experience a one year spell of unemployment at different moments (in, say, the first, third, fifth and seventh year of the observation period respectively), any effect of a change in the average unemployment benefit over time would be captured by their household-specific level effects (μ_i), rather than by the macro indicator (β_4). Technically, this means that a time series effect would be captured by a set of cross-section parameters. As for β_3, the explanation for its low value might be that the macroeconomic variable reflects three different pension benefits (corresponding with the employees, self-employed and civil servants' schemes).

IV. The income simulation model

The main purpose of this paper is to project poverty indicators from 2003 up to 2005, based on the household income model estimated in the previous paragraph. To explain how this can be achieved, let us start from the following general formulation of the panel model:

$$y_{it} = X_{it}\,\beta + \mu_i + \varepsilon_{it} \qquad (2)$$

where the overall mean μ is contained in the β-vector. Clearly, estimates of the dependent variable can be obtained from

$$\hat{y}_{it} = X_{it}\hat{\beta} + \hat{\mu}_i \qquad (3)$$

using the exogenous variables X_{it}. A drawback of this method, however, is that the estimated household incomes have substantially lower variance than the observed incomes due to the omitted variance of the error term ε_{it}. While this is inherent in any simulation based on regression analysis, it poses a serious problem to study aspects of income distribution. Indeed, the variance and higher moments of the distribution play a key role in determining the prevalence and severity indicators of poverty and inequality. As such, it is important for the distribution of the estimated or simulated incomes to reflect the observed income distribution as faithfully as possible. A natural and simple way to achieve this goal is to add an error term to the estimated equation, randomly drawn from a normal distribution with zero mean and variance equal to the estimated variance from equation (1):

$$\hat{y}_{it} = X_{it}\hat{\beta} + \hat{\mu}_i + \tilde{\varepsilon}_{it} \qquad (4)$$

where $\tilde{\varepsilon}_{it} \approx N\left[0, \hat{\sigma}^2(\varepsilon_{it})\right]$.

This procedure ensures that the first two moments of the estimated income distribution correspond to those of the observed distribution. Unfortunately, drawing error terms from a normal distribution does not guarantee that the higher moments coincide as well. Specifically, when the observed income distribution is skewed and leptokurtic (as is often the case), the assumption of normality may be untenable. We considered two alternative methods to tackle this problem. First, we tried to fit a known probability distribution to the estimated residuals, with the first four moments matching their empirical counterparts. The random errors could then be drawn from this fitted distribution. Unfortunately, finding a theoretical distribution with a reasonable fit to the estimated errors turned out to be rather difficult. The second alternative, which led to more satisfactory results, was to estimate the empirical distribution nonparametrically, using an Epanechnikov kernel density estimator evaluated at 400 equally-spaced intervals. Transforming the kernel density estimates into a cumulative distribution then allowed us to generate the random errors from this empirical distribution function.

Before presenting the simulation results of the model, it can be integrated in the well-known classification of micro simulation models by Harding (1996). As a key variable is time, this model clearly is dynamic. Furthermore, as individuals do not change as such, it is a model with static ageing. However, by contrast with most models of this type, we do not reweigh the individual units (the individual characteristics are indeed assumed to remain constant after 2002), but instead concentrate on uprating household incomes using macroeconomic information on various incomes and benefits. As such, our model seems best suited for making short-term projections designed to show the effect on poverty of the possibly diverging development of various time series on incomes and social benefits (see Table 1). These developments, which can be influenced by changes in social policy such as increased per capita unemployment spending, will affect the incomes (and poverty and inequality indicators) of the actual beneficiaries only, taking into account their unobserved individual and household characteristics. Obviously, only the distributional effect of policy interventions with a measurable impact on the macroeconomic aggregates can be analyzed with the current model.

This concludes the description of the model. The next section will discuss the simulation results of the model, concentrating notably on poverty measures.

V. Simulation results: poverty indicators 2003-2005

In view of the data limitations discussed in section 2 and the additional implicit or explicit restrictions implied by the model specification discussed in sections 3 and 4, the simulation results presented in this section must be interpreted with caution. Specifically, the estimated poverty measures are to be interpreted as indicative numbers rather than precise estimates. They are internally comparable (over time or between subgroups of the sample), but do not necessarily coincide with estimates based on other data sources or models.

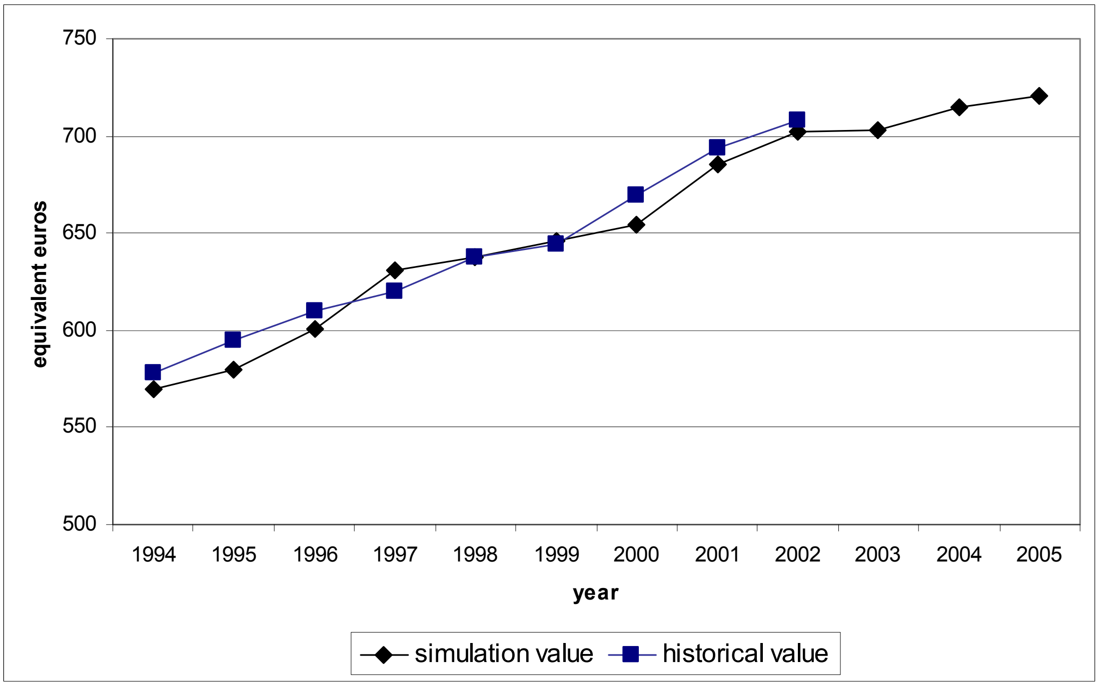
A. Poverty

The financial dimension of poverty is usually approximated by indicators based on poverty thresholds. An individual or family with an income below the threshold level is considered to be at risk of poverty. The 'relative method' for estimating financial poverty risk calculates the threshold or 'poverty line' as a percentage of a standard of living indicator. A widely used criterion sets the poverty line at 60 percent of median equivalent income [2], and this is the criterion adopted in the following paragraphs.

A first and most straightforward way of considering the fit of the model is by comparing the simulated and historical income means for all periods under consideration.

[2] The equivalent income takes household composition into account to determine the financial well-being of its members. Each family member is assigned a weight, and individual income is defined as household income divided by the sum of the weights. The weights are assigned according to the EU equivalence scale: 1 for the first adult, 0.5 for each additional adult, and 0.3 for each child.

Figure 1 - Simulated and historical poverty lines



The difference between a simulated and historical value may be caused by the simulation error in the model describing income. Figure 1 shows that the simulation error seems limited. We will therefore reconsider the fit of the model when discussing sample-wide poverty measures in Figure 2 and the inequality measures in Figure 7.

In this section, some exemplary measures of poverty will be discussed. Consider a poverty line z and a sample of individuals i with income y_i. As the measures discussed pertain to separate years, the subscript for time is omitted. The standard analysis of poverty starts with the headcount ratio $HC=p/n$, the ratio of the number of poor $p=\#(y_i \leq z)$ in the entire population. The headcount reflects the number of poor, but is insensitive to the depth of poverty. An alternative measure, the Income Gap Ratio (IGR), is $\frac{z-\bar{y}}{z}$, the relative shortfall of the average income of the poor[3]. This measure can be rewritten as $\frac{1}{p}\sum_{i=1}^{p}\left(\frac{z-y_i}{z}\right)$, for $y_i \leq z$, but, although it captures the depth of poverty, it is completely insensitive to the number of poor. An obvious measure that combines both aspects of poverty is the product of the HC and IGR. This results in the Poverty Gap Ratio or PGR , which equals $\frac{1}{n}\sum_{i=1}^{p}\left(\frac{z-y_i}{z}\right)$ for $y_i \leq z$. This measure also has the useful advantage over both the HC and IGR that it is 'additively separable': it is the weighted sum of relative poverty gaps of subgroups, the weights being the proportional sizes of these subgroups. Also, like the HC (except in rare cases) but unlike the IGR, the PGR will decrease as a result of economic growth (defined as an increase in mean income which is accompanied by no change in the relative inequality; see Kakwani (1999, 603).

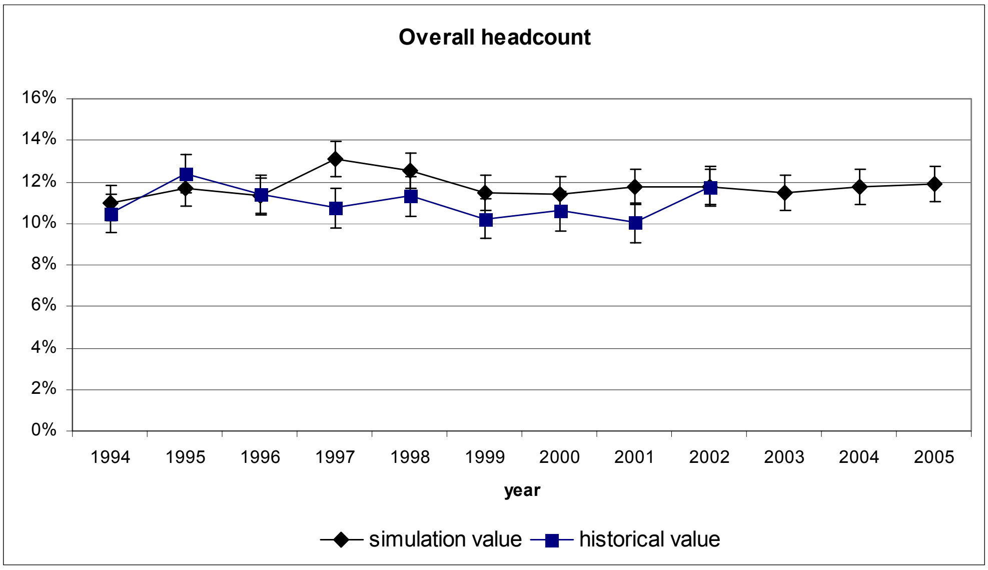
Unfortunately, the PGR as well as the IGR and HC are insensitive to transfers among the poor. In other words, if a poorer individual transfers money to a less poor individual, neither of the above measures reflect the fact that income inequality between the poor will increase. The Squared Poverty

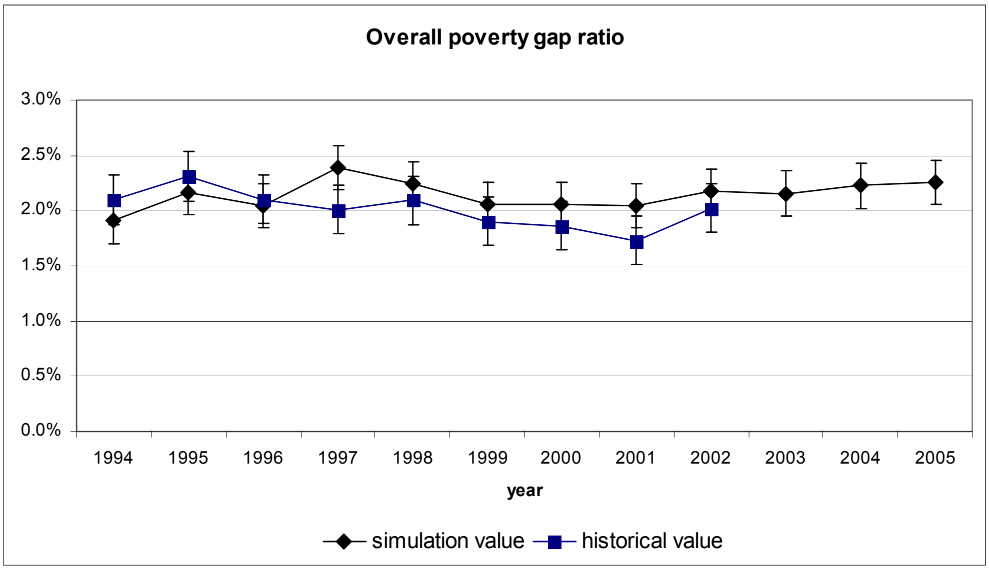
[3] The IGR is the ratio of the per person minimum amount of money that should be given to the poor to bring their income to the poverty line, i.e. to bring them out of poverty (Kakwani, 1999, 609). Note that this is only the case if the poverty line z is fixed or a function of median income. It does not apply if it is a function of mean income.

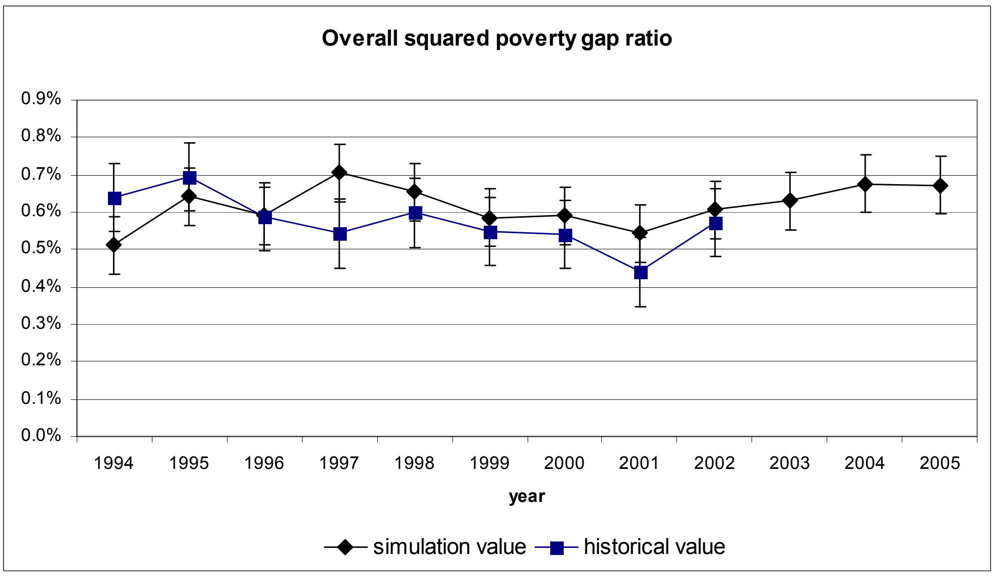
Gap Ratio (SPGR) equals $\frac{1}{n}\sum_{i=1}^{p}\left(\frac{z-y_i}{z}\right)^2$, and has the same characteristics as the PGR, but also takes into account this between-poor inequality. Note, finally, that the above headcount ratio, PGR and SPGR are all special cases of the decomposable general measure by Foster, Greer and Thorbecke.

Figure 2 contains the HC, PGR and SPGR, each with its confidence interval, using the Huber/White sandwich estimate of variance (Williams, 2000).

Figure 2 - Overall poverty indicators







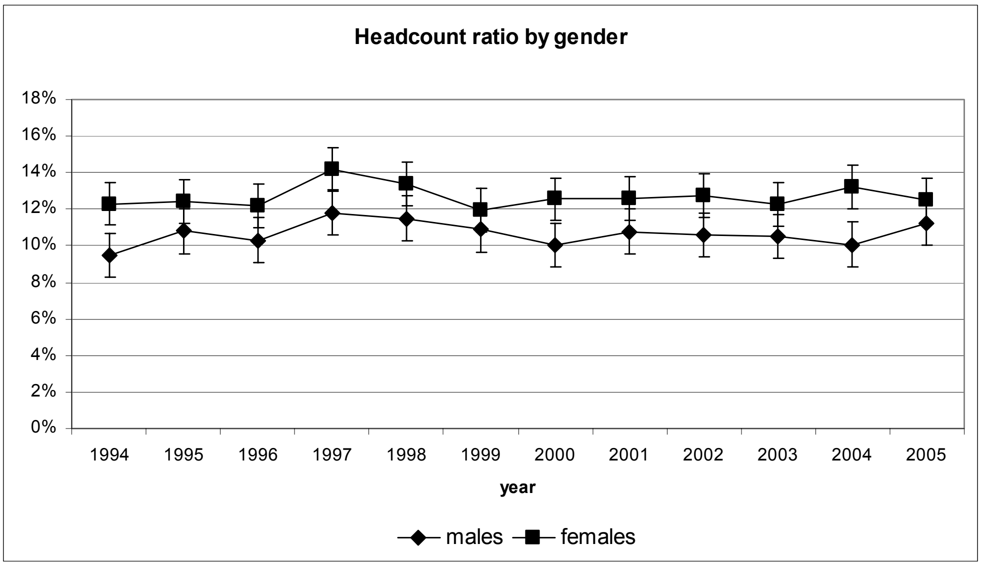
The first pane of Figure 2 contains the historical and simulated headcount ratio for the years under consideration. The historical value is based on the waves of the PSBH between 1994 and 2002. The simulated values cover the years until 2005. Again, the difference between the two seems to be limited and not statistically significant, lending support to the ability of the model to reproduce the salient features of Belgian (financial) poverty indicators.

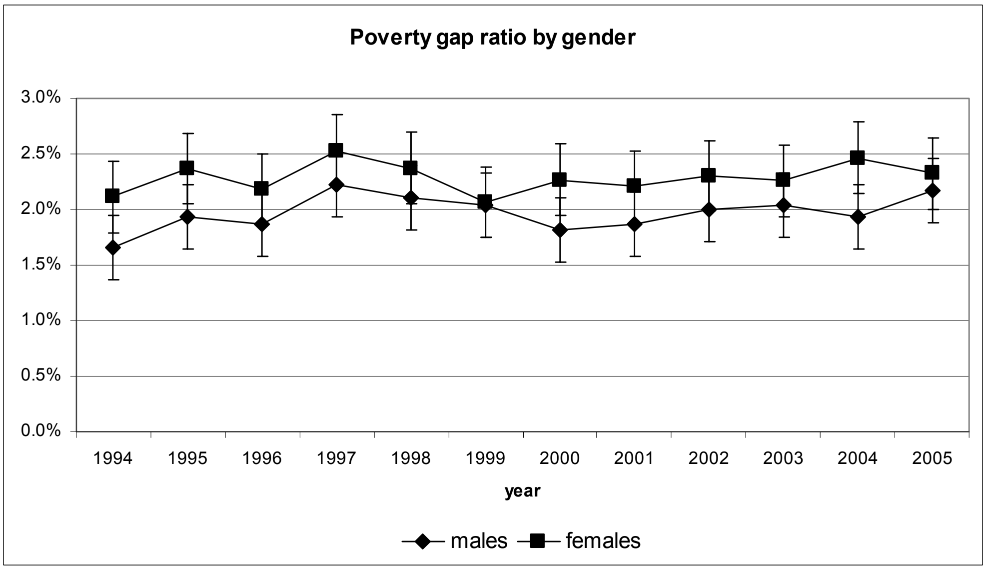
The discussion now turns to the simulation results as such. The overall simulated headcount lies between 11 and 13 percent. The vertical lines in each year reflect the 95% confidence interval. The headcount ratio shows an increase between the years 1994 and 1997, followed by a decrease until 2000. It then remains stable until 2003 after which a slow increase seems to set in. The pattern in the headcount ratio is more pronounced in the course of the PGR and the SPGR. The decrease of the number of poor seems to go simultaneously with a decrease in the average depth of poverty and the inequality between the poor. However, the increase of the headcount setting in from 2000 onward goes along with a sharp increase of the (squared) poverty ratio.

Next, Figure 3 shows the three indicators of poverty by gender. The headcount ratio of females in all years is higher than that of males. The same conclusion holds for the PGR and SPGR, save for the year 1999. However, the confidence intervals show that the difference is statistically insignificant except for the headcount in some years. This may seem somewhat unexpected, for common knowledge dictates that the poverty risk is higher for females than for males. However, this study considers household income. When two partners (more often than not of the opposite sex) live in the same household, they are assigned the same equivalent household income. The observed difference therefore comes from male and female singles, and their different patterns in terms of poverty risk is blurred by the sample of married or cohabiting males and females.

Figure 4 shows the poverty indicators, subdividing the sample into two age groups. These are individuals younger than 65, and those 65 and older.

Figure 3 - Poverty indicators by gender





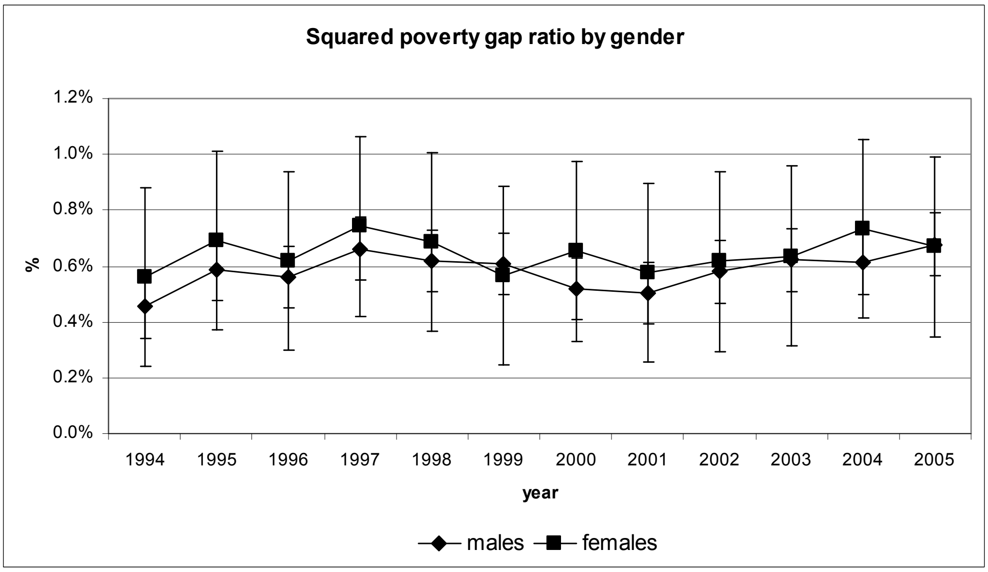
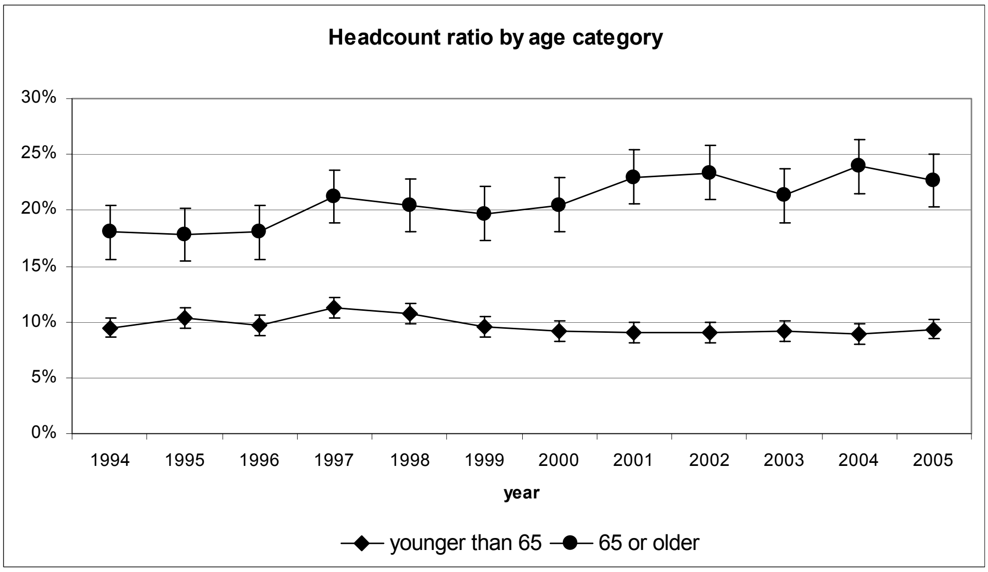
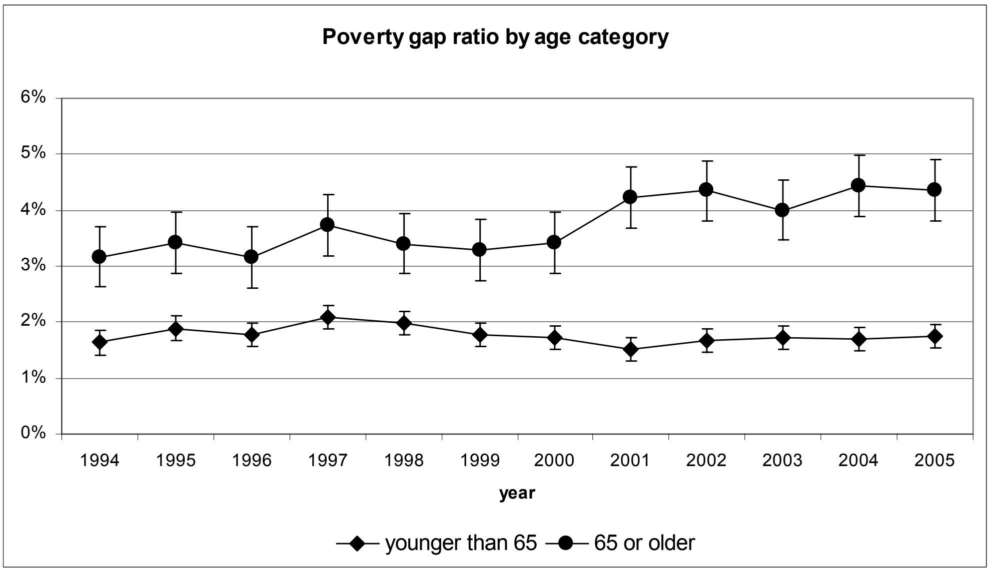
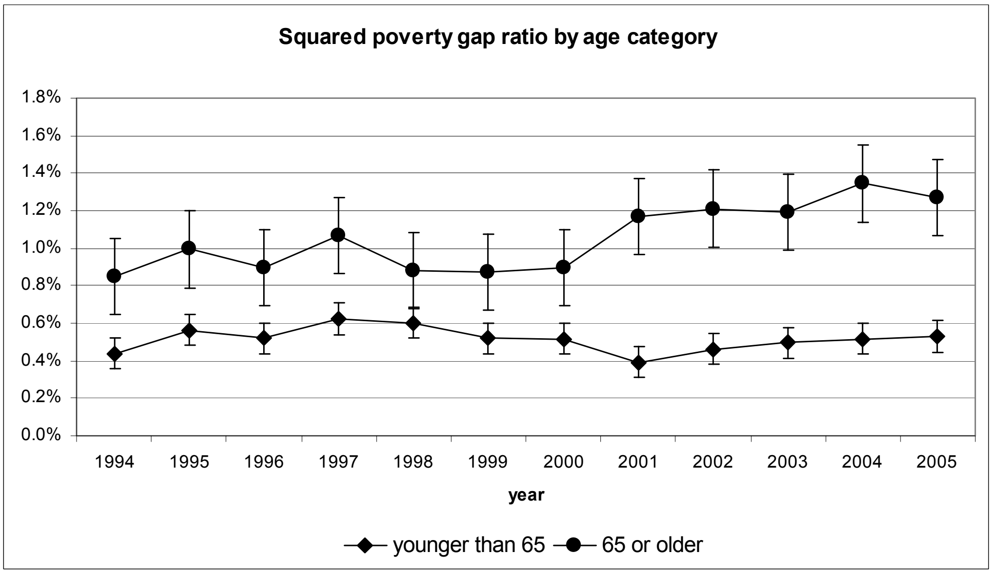


Figure 4 - Poverty by age category





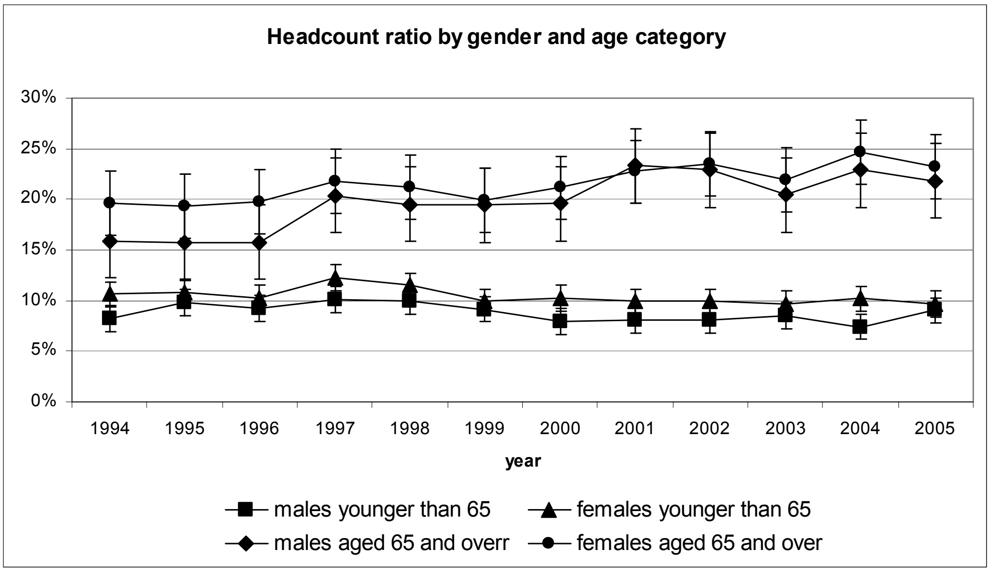


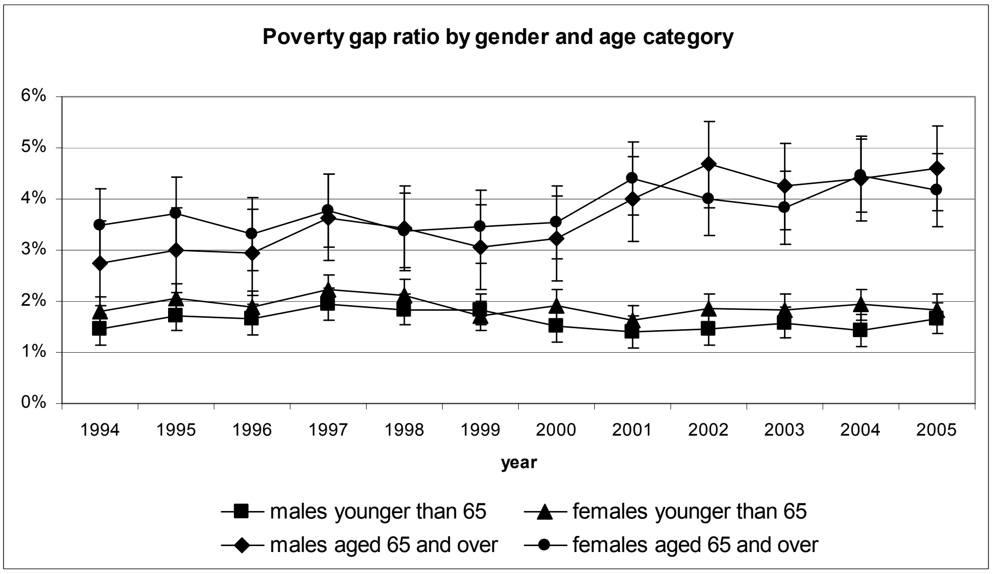
Roughly the same pattern again emerges in the figures of the HC, PGR and SPGR. The risk of poverty (the proportional number of poor) is significantly lower for those younger than 65 than for those of 65 and older, and these differences are always significant. The difference between the age groups is however much more outspoken in the HC and the PGR than it is in the SPGR. This suggests that the difference between those younger and older than 65 is less outspoken if average depth of poverty and inequality among the poor is taken into account. This is especially so in the first half of the simulation period. Over the simulation period, the divergence which has been going on during the years between 1999 and 2002 seems to become less outspoken, and the difference between those older and younger than 65 seems to stabilize and even decrease in 2003 and 2005.

Figure 5 shows the subdivision of the sample to gender and age category. For the same reason as argued in Figure 3, the difference between males and females within each age category is almost never significant and lines often cross. Furthermore, the pattern that emerged in Figure 4 again appears in Figure 5. When all household types are considered together, the conclusion is that age category is a stronger determinant of poverty than gender. The reason for this is obvious: individuals of opposite sex often join in households, which means that the effect of gender on poverty is obfuscated by the fact that a large part of males and females share an equal household income. Instead, partners in the same household often fall under the same age category.

Finally, Figure 6 contains the poverty measures for separate socio-economic categories. Before discussing the results, it should be reminded that the link between the micro incomes and their macroeconomic counterparts is rather weak, especially in the case of unemployment benefits and self-employment income. The category ‘other’ has a residual character. It is included only for completeness’ sake, but will not be discussed.

Figure 5 - Poverty by gender and age category





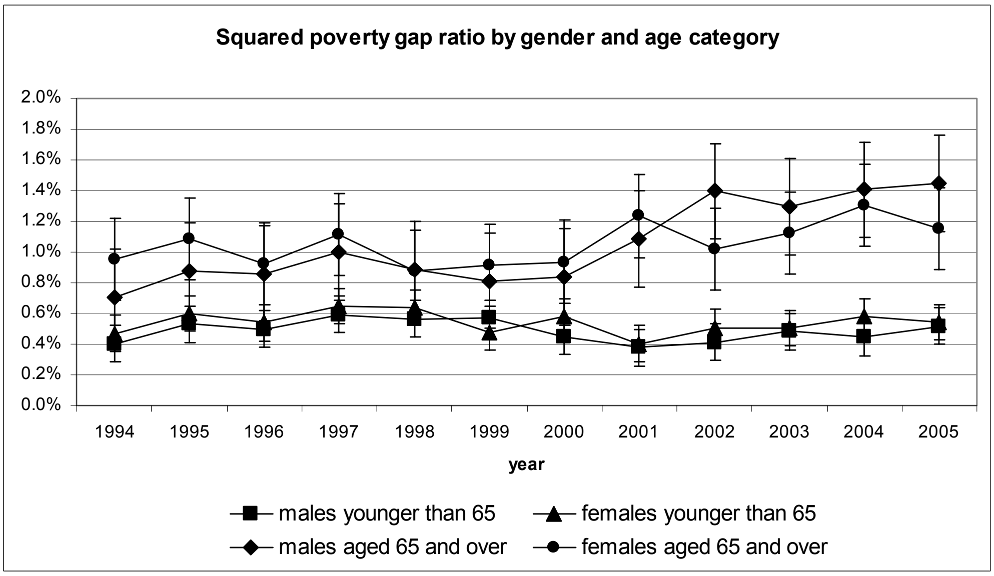
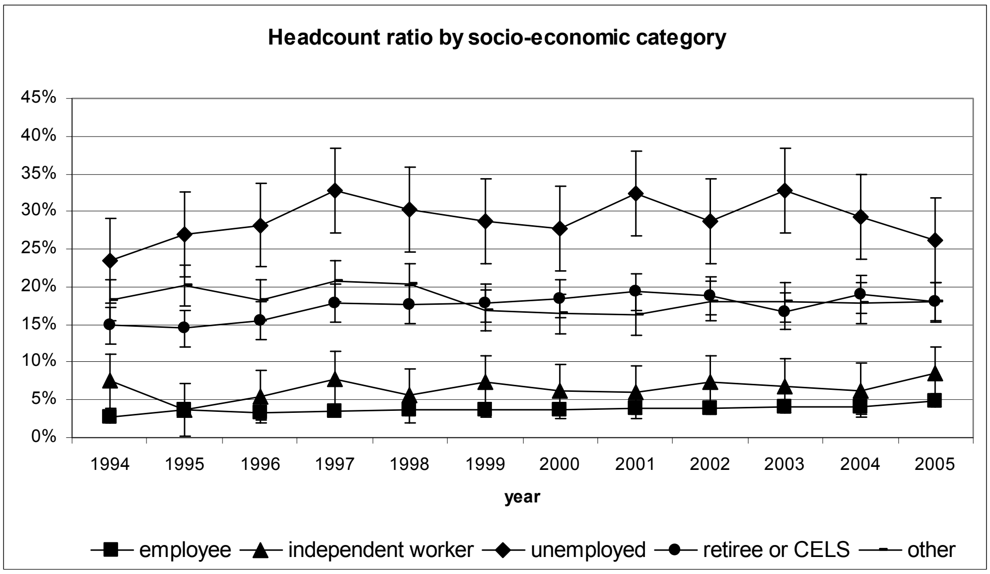
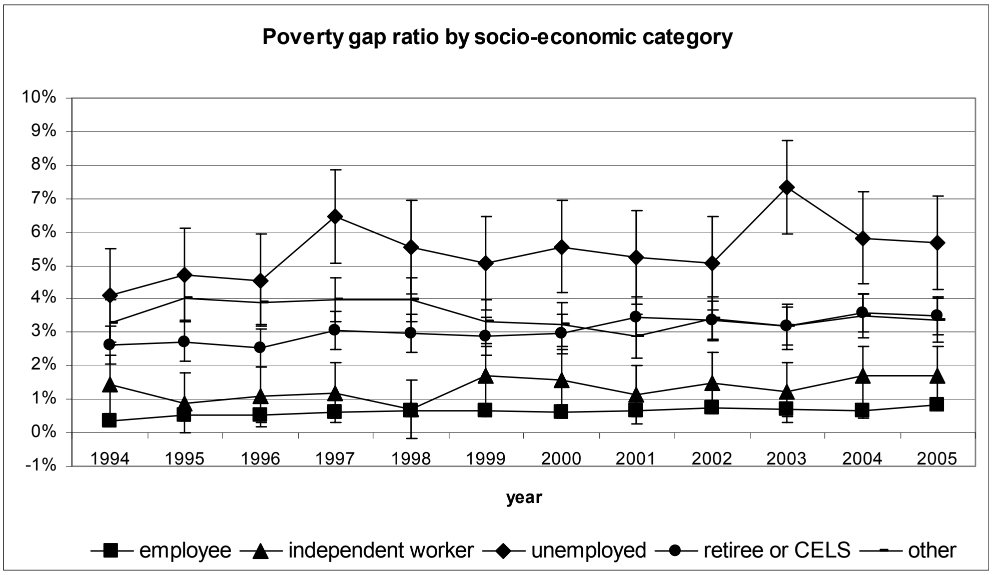


Figure 6 - Poverty by socio-economic category





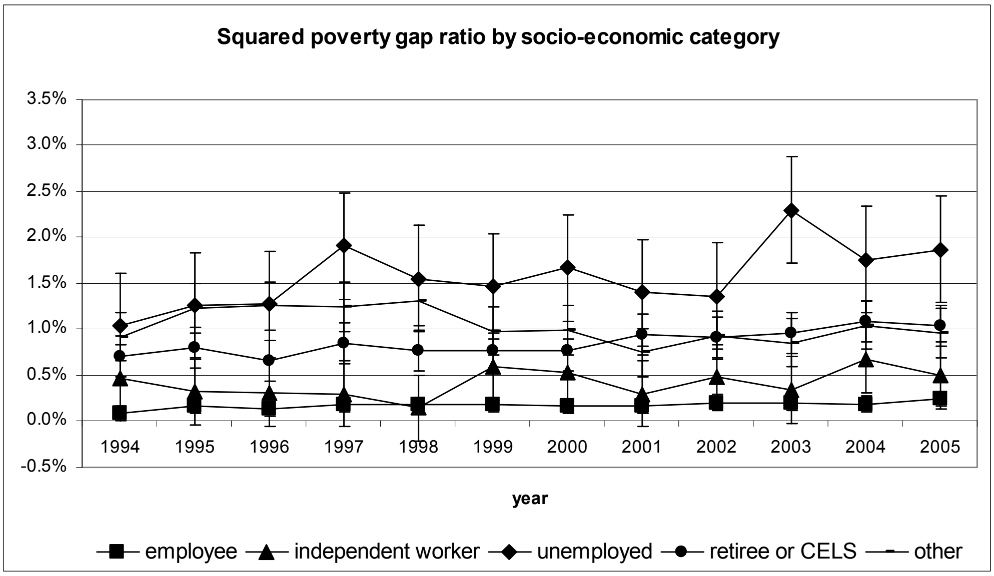


Figure 6 shows that employees run the lowest risk of poverty, with a headcount ratio situated between 2 and 5 percent. The poverty risk of the independent workers is somewhat higher, but the difference is never significant. The poverty risk of retirees and CELS-beneficiaries[4] is third in rank. The headcount lies between 14 and 19 percent. Finally, the class that shows to have the most precarious situation, in terms of the HC but also of the (S)PGR are the unemployed. Their risk of poverty lies between 23 and 33 percent and is in almost all years statistically significant from the poverty risk of the other categories of individuals. By contrast, the boundaries of the SPGR confidence intervals of the unemployed sometimes cross with those of the retirees. This might suggest that the within-poor inequality is lower for the unemployed than for the retirees, which partially neutralizes the higher poverty risk for the former as compared to the latter. The simulation results for the years 2003-2005 show no remarkable changes to the previous years. Even for the unemployed, where a distinct decrease in the headcount seems to set in, the annual changes remain small relative to the width of the confidence intervals. Furthermore, this decrease is not reflected in the PGR and SPGR.

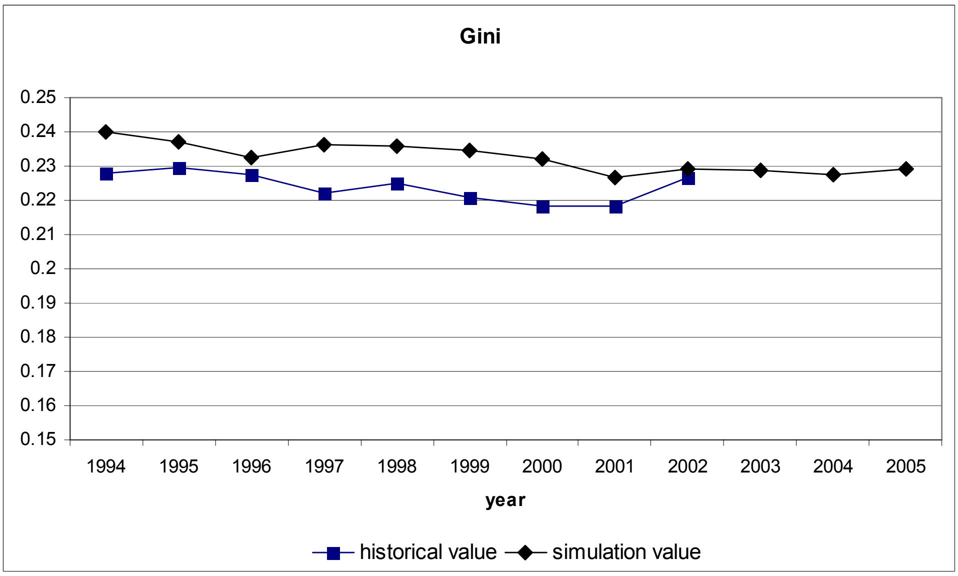
B. Income inequality

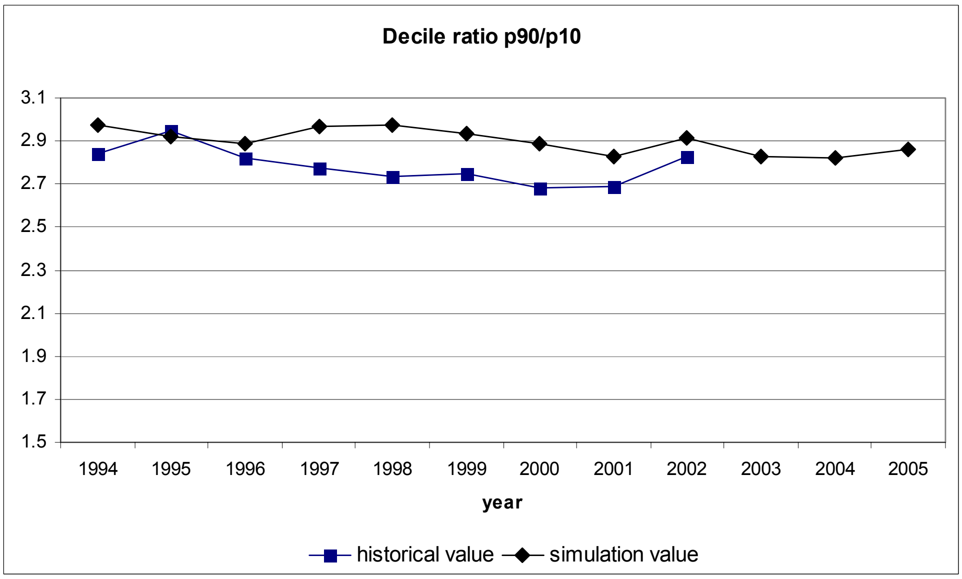
The main focus of this paper is on the simulation of poverty indicators. Before turning to the conclusions, however, we take a quick look at some indicators of income inequality. Figure 7 presents three indicators of income inequality. The Gini coefficient lies between 0 (perfect equality) and 1 and is sensitive to changes in the middle of the income distribution. The mean log deviation (MLD) is more sensitive to changes in the lower part of the income distribution and the percentile ratio (P90/P10) is sensitive to changes in the extremes of the income distribution.

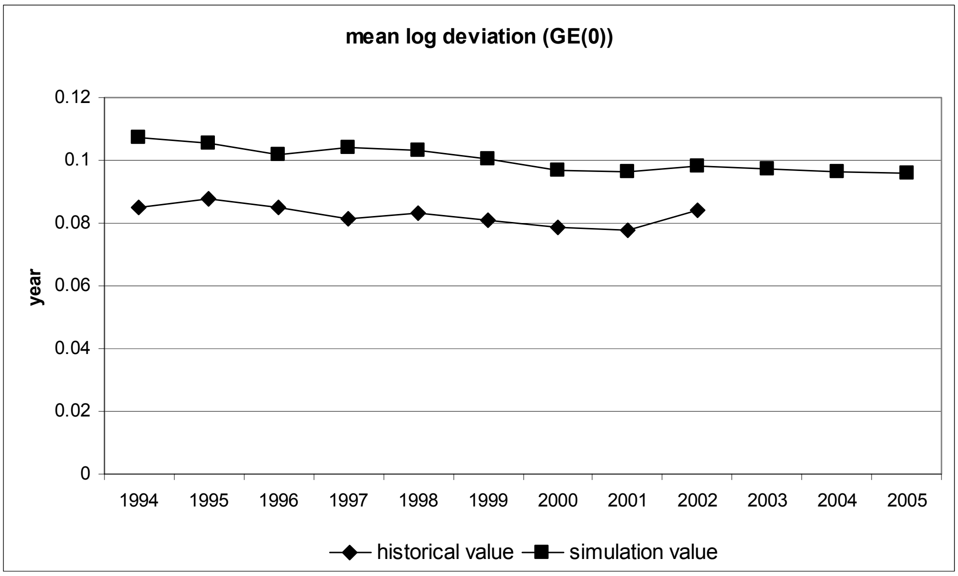
As could be expected from the results in Figure 2, the actual income inequality is somewhat overestimated, and the MLD suggest that the simulation error is mainly in the lower end of the income distribution. However, these simulation errors seem to be limited. The three indicators of inequality in Figure 7 show a continuous decrease in income inequality, only being interrupted by an increase between 1996 and 1997 and between 2001 and 2002.

[4] Conventional Early Leavers' Scheme ('Prépension') essentially is an unemployment scheme for older workers. It consists of an unemployment benefit with an additional benefit which is a function of the difference between the unemployment benefit and the last net wage before unemployment. Though it technically is an unemployment scheme, it is generally considered an early retirement scheme.and this is why it is classified as such.

Figure 7 - Overall income inequality







VI. Conclusions

In this paper a simple static microsimulation model with static ageing is presented. Unlike other models of its kind, it explicitly allows individuals to occupy several labour market states and to combine the associated income sources. This possibility seems more realistic than the usual assumption of mutually exclusive states, especially if household income is used to measure financial well-being. Given that the socioeconomic 'state' of a household is hard to define and, consequently, hard to link to the evolution of sociodemographic and macroeconomic variables, we have specified the link between individual income (defined in equivalence terms) and the macroeconomic environment directly. This was achieved by estimating a fixed effects panel regression model with macroeconomic averages for the main income sources as explanatory variables.

The model, estimated with the Belgian PSBH data set over the 1994-2002 period, produces plausible results both from a statistical point of view and based on in-sample simulation of the main poverty indicators. Its main purpose is to make short-term projections of poverty and inequality indicators, based on the evolution of the macroeconomic income variables. In order to obtain simulated incomes that match the distributional properties of the observations as closely as possible, we have drawn the random disturbances from the empirical (estimated) error distribution, using a kernel density estimator. The results, generated for the 2003-2005 period, confirm the ability of the model to produce sensible short-term poverty indicators.

While it is encouraging that a relatively simple model of individual equivalent income can be used to project poverty and inequality measures, a few limitations should be kept in mind. First, the fact that subjects are assumed to remain in their end-of-period states limits the use of the model to short-term projections. Indeed, labour market positions and sociodemographic characteristics can be assumed to remain constant only over short horizons. Second, the link between the observed micro incomes and the available macro variables is rather weak for some income categories, notably self-employment income and unemployment benefits. Two main factors may provide an explanation for this problem. One reason is probably that these incomes are often supplementary to the main source of family income, and, in the case of unemployment benefits, transitory in nature. A second is that the macroeconomic variables are probably too general to capture the evolution of the specific income sources of the households. A pension, for example, can be received from one of the three main pensions schemes (employees, self-employed, civil servants), which differ substantially from each other in terms of the average amounts and their evolution over time. Unfortunately, the available household income data do not allow to differentiate between the schemes. Consequently, the corresponding macro variable, the overall average per capita pension, is a rather crude measure to reflect changes in any of the underlying components. As a result, it is difficult to assess the impact of targeted social policy measures, aimed, for example, at alleviating the poverty of some subgroups of formerly self-employed people.

Despite its limitations, we believe the model is a useful tool to evaluate the impact of diverging evolutions of the main sources of personal income on poverty and inequality over short horizons. These divergent tendencies may be endogenous outcomes of the course of the macro economy (such as business cycles), or they may be the result of deliberate interventions by social policy makers. In the latter case, however, only the policy measures with a discernible impact on the macro averages will have a detectable influence on poverty measures simulated with the current model. Improving the sensitivity of the poverty indicators to targeted social policy interventions requires a more detailed model. One possibility, which we intend to investigate in future research, is to use individual instead of household data. We hope that the PSBH data set and its successor (the EU-SILC survey) will prove to be sufficiently rich and reliable to achieve that goal. If so, it should be possible to identify macroeconomic income variables that reflect the diverse sources of household income better than the aggregates we used in the current model.

VII. References

G.J.M. Dekkers, 2003, Socio-economic modelling, in Becker, H., F. Vanclay, 2003, *International Handbook of Social Impact Assessment: Conceptual and Methodological Advances*, Edward Elgar, Cheltenham, pp 179-194.

Harding, A., 1996, Introduction and Overview. in: Harding, A., *Microsimulation and Public Policy*, Contributions to Economic Analysis, Amsterdam: Elsevier Science, pp 1-20 .

Kakwani, N., 1999, Inequality, Welfare and Poverty: three interrelated phenomena. In: Silber, J. (ed.), 1999, *Handbook on Income Inequality Measurement*, Dordrecht: Kluwer Academic Publishers.

Mortelmans, D., M-T. Casman and R. Doutrelepont (Eds.), 2004, *Elf Jaar uit het Leven in België: Socio-Economische Analyses op het Gezinsdemografisch Panel PSBH*, Gent: Academia Press.

Panel Study on Belgian Households, 1992-2004, Universiteit Antwerpen-Université de Liège.

O'Donoghue C. 2001, Dynamic Microsimulation: A Survey, *Brazilian Electronic Journal of Economics*.

Williams, R., 2000, A Note on Robust Variance Estimation for Cluster-correlated Data, *Biometrics*, 56, 645-646.