ISSN 1977-0375

eurostat
Methodologies and Working papers

Macro determinants of individual income poverty in 93 regions of Europe

2010 edition

eurostat
EUROPEAN COMMISSION

eurostat
Methodologies and Working papers

Macro determinants of individual income poverty in 93 regions of Europe

2010 edition

eurostat
EUROPEAN COMMISSION

More information on the European Union is available on the Internet (http://europa.eu).

Cataloguing data can be found at the end of this publication

Luxembourg: Publications Office of the European Union, 2010

ISBN 978-92-79-16764-1
ISSN 1977-0375
Doi:10.2785/53721
Cat. No. KS-RA-10-012-EN-N

Theme: Population and social conditions
Collection: Methodologies and working papers

Eurostat is the Statistical Office of the European Union (EU). Its mission is to provide the EU with high-quality statistical information. To that end, it gathers and analyses data from the National Statistical Institutes (NSIs) across Europe and provides comparable and harmonised data for the EU to use in the definition, implementation and analysis of EU policies. Its statistical products and services are also of great value to Europe's business community, professional organisations, academics, librarians, NGOs, the media and citizens. In the social field, the EU Statistics on Income and Living Conditions (EU-SILC) instrument is the main source for statistics on income, poverty, social exclusion and living conditions.

Over the last years, important progress has been achieved in relation to EU-SILC. This is the result of the coordinated work of Eurostat and the NSIs, *inter alia* in the context of the EU 'Living Conditions' Working Group and various thematic Task-Forces. Despite these significant achievements, EU-SILC data are still insufficiently analysed and used.

It is in this context that Eurostat launched in 2008 a call for applications with the following aims:

(1) develop methodology for advanced analysis of EU-SILC data;
(2) discuss analytical and methodological papers at an international conference;
(3) produce several publications presenting methodological and analytical results.

The 'Network for the Analysis of EU-SILC' (Net-SILC), an ambitious 18-partner Network bringing together expertise from both data producers and data users, was set up as in response to this call. The initial Net-SILC findings were presented at the international conference on 'Comparative EU Statistics on Income and Living Conditions' (Warsaw, 25-26 March 2010), which was organised jointly by Eurostat and the Net-SILC network and hosted by the Central Statistical Office of Poland. A major output from Net-SILC is a book to be published by the EU Publications Office at the end of 2010 and edited by Anthony B. Atkinson (Nuffield College and London School of Economics, United Kingdom) and Eric Marlier (CEPS/INSTEAD Research Institute, Luxembourg).

The present methodological paper is also an output from Net-SILC. It has been prepared by Anne Reinstadler (CEPS/INSTEAD, Luxembourg) and Jean-Claude Ray (Nancy University and CNRS UMR 7522, France). Gara Rojas González was responsible at Eurostat for coordinating the publication of the methodological papers produced by Net-SILC members.

It should be stressed that this methodological paper does not in any way represent the views of Eurostat, the European Commission or the European Union. The authors have contributed in a strictly personal capacity and not as representatives of any Government or official body. Thus they have been free to express their own views and to take full responsibility both for the judgments made about past and current policy and for the recommendations for future policy.

This document is part of Eurostat's *Methodologies and working papers* collection which are technical publications for statistical experts working in a particular field. All publications are downloadable free of charge in PDF format from the Eurostat website (http://epp.eurostat.ec.europa.eu/portal/page/portal/income_social_inclusion_living_conditions/publications/Methodologies_and_working_papers). Furthermore, Eurostat databases are freely available at this address, as are tables with the most frequently used and requested short- and long-term indicators.

Table of contents

Macro determinants of individual income poverty in 93 regions of Europe

Anne REINSTADLER[1]
CEPS/INSTEAD, Luxembourg

Jean-Claude RAY[2]
Nancy University and CNRS UMR 7522, France

Abstract: The analysis of the at-risk-of-poverty determinants can be improved by taking into account factors at macro (regional) level. This hypothesis has already been made in previous research, most often at country level, on cross-sectional data. We use longitudinal data in this analysis in order to obtain more accurate estimated parameters, and we test whether the regional unemployment rate and the regional GDP affect the individual at-risk-of-poverty status. The countries taken into account are those included in the Statistics on Income and Living Conditions (EU-SILC) dataset.

Key words**:** income poverty, macro determinants, EU-SILC, multilevel models, longitudinal data

JEL: I32

---

[1] Corresponding author: anne.reinstadler@ceps.lu
[2] jean-claude.ray@univ-nancy2.fr
We would like to thank David Brady, Jacques Brosius, Alessio Fusco, Tony Atkinson, Eric Marlier and Philippe Van Kerm for useful discussions. This paper has also been presented at the international conference on ‘Comparative EU Statistics on Income and Living Conditions’, 25-26 March 2010, Warsaw. It has benefited there from useful suggestions from Brian Nolan. Of course, remaining errors are ours.

1. Introduction

Tackling poverty by 2010 was one of the objectives defined by the Lisbon European Council in 2000. Ten years on, 2010 is the European Year for Combating Poverty and Social Exclusion, and new objectives are being defined that have to be achieved by 2020. Poverty therefore remains at the heart of social policy in most Member States. Ideally, public policies aimed at reducing poverty need to be based on an in-depth understanding of the underlying processes at work. A first step towards such an understanding consists in shedding some light on the main determinants of poverty.

Early descriptive studies have checked for relationships between poverty status and different characteristics taken in turn (Bradshaw, 1999; UNICEF, 2000; Mejer *et al*, 2000; Bradbury *et al*, 2001). This has given some insight into the factors involved, but these studies have provided only a partial understanding of how these factors work. Other researches (see for example Cappellari and Jenkins, 2002; Fertig and Tamm, 2007) have widened out this initial approach by reasoning all other things being equal, checking the effect on poverty of factors such as educational attainment, age, employment status, family structure – all factors that have been calculated at the individual level. Simultaneously, another stream of studies (see Moller *et al*, 2003; Wiepking and Maas, 2005; Tai and Treas, 2008; Callens and Croux, 2009; Brady *et al*, 2009) has emphasized the analysis of the role of macro characteristics in a cross-national context. These analyses have shown that the macro factors could well have an effect on the probability of poverty. Indeed, the generosity of social benefits (and especially of family benefits) is proving to have a significant negative effect on the odds of poverty (cf. Brady *et al*, 2009, Moller *et al*, 2003).

For all of these three types of analyses, one major improvement has been to take into account the longitudinal feature of poverty, using panel data.[3] An indicator of persistent poverty, for example, enables analysts to determine whether poverty is a temporary phenomenon or a long-term one. Furthermore, developments in the econometrics of panel data have allowed researchers to further investigate important topics, such as poverty duration or unobserved heterogeneity.

[3] See Ray and Jeandidier (2003) for a comprehensive review of the French literature on this subject.

Brady *et al* (2009) study the effect of macro-determinants on the probability of being poor, using a GEE model[4] applied to 15 EU (plus some non-EU) countries, but with cross-sectional data. In fact, to our knowledge, only one study has so far dealt simultaneously with all EU-countries, longitudinal data and factors at both individual and macro levels using a relevant specification (Callens and Croux, 2009). In this paper we will extend this kind of analysis to 93 European-regions (26 countries[5]), using the EU-SILC longitudinal dataset (2005 and 2006).

The aim of this paper is to disentangle the role of micro and macro factors in explaining the poverty status, by using detailed information about different regions in Europe. Indeed, we would like to test whether there is a genuine effect of macro factors, such as the regional unemployment rate, on the poverty probability, and especially if these factors can affect the impact of individual characteristics, such as the education level, on this probability.

In Section 2 we present the definition of income poverty that we intend to apply. In Section 3 we then develop the different econometric methods available to deal with the issues and data at hand. Section 4 gives a detailed description of our dataset. The results and comments of our own model are then presented in Section 5. Final conclusions are to be found in Section 6.

[4] A Generalized Estimated Equations model can be used to estimate marginal or population-averaged effects taking into account the dependence among units nested in clusters (Rabe-Hesketh and Skrondal, 2008).
[5] Belgium, Czech Republic, Denmark, Germany, Estonia, Ireland, Greece, Spain, France, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Hungary, Netherlands, Austria, Poland, Portugal, Slovenia, Slovakia, Finland, Sweden, United Kingdom, Iceland and Norway

2. The definition of income poverty in Europe

In Europe, poverty is officially defined in relative terms as the percentage of individuals living in a household whose equivalent income is below the poverty threshold. This threshold is defined in each country (equal to 60% of the national median equivalent income), in order to take the national income inequalities. As a consequence, two countries with very different standards of living (and thus very different median equivalent incomes and different poverty thresholds) may have the same poverty rate (meaning that they actually have the same level of income inequality).

Seemingly contradictory results due to this definition are immaterial, as long as one is aware of the conventions on which they are based, and when the at-risk-of-poverty rates are interpreted together with the threshold values. However, in our case, the main objective is to establish to what extent certain macro factors, such as the GDP (Gross Domestic Product) or the unemployment rate, can explain the fact of being poor. Using the official at-risk-of-poverty thresholds could actually allow the kind of comment such as: the macro characteristic of the countries has a particular effect on national inequality levels. Interesting though this question may be, it is not the one that we plan to address in this paper, where poverty and not inequality is at stake. What we need, therefore, is a measure that ranks individuals in terms of poverty and not of inequality. As a consequence, we have chosen to continue defining poverty in relative terms (60% of a certain threshold), but to calculate a new single European poverty threshold by considering that all individuals belong to the same big country, which is Europe. Note that this definition of a European at-risk-of-poverty threshold is supported by Marlier *et al* (2007).[6]

We then determine, for each country, the proportion of individuals situated below this new European threshold (cf. Appendix 1, where it can be seen that, in 2006, the new poverty rates range from 2%-3% in Denmark, the Netherlands, Norway or Luxembourg, to 72%-75% in Slovakia, Poland, Latvia and Lithuania). These figures concern individuals of working age only (aged 25-55), because the behaviours - and thus the factors at work - can be very different for the other two groups (children and retired people).

[6] Note also that the results we get using the official definition of the at-risk-of-poverty threshold do not differ much from those we present in this paper (a fact that could explain why the other authors adding macro factors in their analysis keep with this usual definition – cf. for example Callens and Croux, 2009; Brady *et al* 2009).

3. Methodology

Two main approaches have been used to study the determinants of income poverty. One consists in explaining the transitions into and out of poverty (probability of staying poor, and probability of entering poverty). The second approach focuses on the poverty status at a specific point in time.

The first approach takes into account the issue of initial conditions by using longitudinal data. This issue refers to the fact that the poverty status during the first period may not be exogenous because of observed and unobserved characteristics, which would subsequently affect the probability of being poor. Some authors also control for the retention probability (see, for example, Cappellari, 2002; Cappellari and Jenkins, 2002, 2004; Ayllon, 2008). The idea is that the probability of being observed during two consecutive years could depend on unobserved characteristics of the individuals which should thus be controlled for. However, papers running that kind of analysis do not introduce macro factors into the analysis. This is either because they are interested in only one country (see Cappellari and Jenkins, 2002, 2004; Van Kerm, 2004; Buddelmeyer and Verick, 2008; Ayllon, 2008[7]), or because the various countries are treated separately, with as many models as there are countries (Andriopoulou *et al*, 2008). One exception is the analysis of 63 European regions by Callens and Croux (2009), although they have adopted another approach: using longitudinal data, they estimate a recurrent discrete-time hazard model. More precisely, they estimate two equations, the first for poverty entry and the second one for poverty exit, using the discrete logit model (cf. Allison, 1982).

Some authors, on the other hand, estimate the probability of being poor at a specific point in time (cf. Wiepking and Maas, 2005; Tai and Treas, 2008; Brady *et al*, 2009). All of these authors use cross-sectional data of 22 countries from the Luxembourg Income Study (LIS) and integrate macro variables (such as the unemployment rate or the welfare generosity) in the analysis, stressing that the welfare system could play a role in allowing individuals to escape from poverty.[8]

[7] Other authors have used the same kind of models, on related subjects but not poverty: Stewart and Swaffield (1999) and Cappellari (2004) on earnings, Poggi (2007) on the social exclusion persistence.
[8] Moller *et al* (2003) work on these data as well, but at the macro level. Indeed, they link the national at-risk-of-poverty rate to macro variables such as the GDP or the employment rate in the agricultural sector. Their sample is quite small (61 observations, nested in 14 countries).

In this paper, we focus on the poverty status and integrate macro factors as well. In a similar way to Callens and Croux (2009), we use longitudinal data and we take into account the fact that some variability can be found at the regional level; unlike them, however, we do not model the poverty duration but the poverty probability. In fact, two reasons have led us to conduct the analysis at the regional rather than at the country level: firstly because the situation the individuals face (in terms of unemployment rate, for example) could be very different from one region to another within the same country, and secondly because there are more regions than countries (93 versus 26), which is better from a statistical point of view. [9]

In other words, we estimate a model of poverty probability, using two years of observations for each individual (in order to increase the accuracy of the estimates). Some of these individuals live in the same region. From an econometric point of view, this data setup causes a problem as far as the independence of observations is concerned: individuals being observed over two years and/or living in the same region share their own time-invariant characteristics and/or the characteristics of the area, and can therefore no longer be considered to be independent. As a consequence, using traditional techniques would give consistent estimates, but heavily[10] under-estimated standard errors (cf. Angrist and Pischke, 2009).

In order to cope with that statistical problem, we have chosen one of the several techniques available: we run a multilevel model, which treats the upper levels (the individuals and the regions here) not as unique entities but as units primarily characterised by factors calculated at their level (e.g. characteristics of the individual, or the unemployment rate of the region). These models explicitly take into account the hierarchical structure of the data, thereby allowing us to analyse — first to measure, then to explain — the proportion of the variability of the poverty rate which is attached to each (nested) level. Unlike the fixed effects models[11], multilevel models make use of the between variance, and are therefore especially useful when this variance is quite high. Some authors have already stressed that the use of this kind of model would be relevant in this context (Cappellari and Jenkins, 2004; Brady *et al*, 2009), but they have underlined the complexity of these models, whose convergence status is often out of reach. We could also consider that some individuals live in the same household and that the regions are nested within the countries. This would suggest moving from a three-level analysis (observations over time nested

[9] Note that some countries do not have a sample design which allows to calculate indicators at the regional level. However, this is not our purpose here, as we are only interested in shedding light on the factors which affect the outcome (and not in calculating this outcome), and as the measure of the role of each factor rests on the usual hypothesis that the individual behaviours are homogenous, whatever the representativity of the sample.

[10] See Reinstadler and Ray (2010) for more details.

[11] Yet these models have a strong advantage: they control for group-invariant factors, measured or unmeasured. But this advantage has a price: the inability to estimate regression coefficients for these group-invariant factors, and thus to allow the analyst to conclude in terms of the effect of these factors.

within individuals, who themselves are nested within regions) to a five-level analysis (adding the household level, and the country-level at the top of the hierarchy). But models with five levels and so many observations suffer from convergence problems and thus cannot be estimated. We have therefore chosen to restrict our analysis to three levels, and we introduce the country variable as an explanatory variable.

The model we estimate is a binary logistic regression, where the probability of being at risk of poverty is explained. This multilevel model takes three levels into account: time (measured in years), individuals and regions. It can be written as follows in its structural form (using Snijders and Bosker's notation):

$$logit(P_{ijk}) = \beta_{0k} + \beta_{1k}\, x_{1ijk} + \beta_2\, x_{2ijk}$$

$$\beta_{0k} = \gamma_0 + \gamma_1\, z_{0k} + U_{0k}$$

$$\beta_{1k} = \delta_0 + \delta_1\, z_{1k} + U_{1k}$$

where:

i indexes time
j indexes the individuals
k indexes the regions

P_{ijk} is the probability of being at risk of poverty

x_{1ijk} is a vector of independent factors defined at the individual/household level whose effects are assumed to be random

x_{2ijk} is a vector of independent factors defined at the individual/household level whose effects are assumed to be fixed

z_{0k} is a vector of independent factors defined at the regional level, which are supposed to have an impact on the average P_{ijk} in region k

z_{1k} is a vector of independent factors defined at the regional level, which are supposed to moderate the effect of the x_{1ijk} on P_{ijk}

β_{0k} is a random intercept

β_{1k} is a vector of random slopes

β_2 is a vector of fixed slopes

γ_0 measures the average value of P_{ijk} across regions, when each independent variable is 0

γ_1 measures the impact of z_{0k} on P_{ijk}

δ_0 measures the average impact, across regions, of x_{1ijk} on P_{ijk}, when each z_{1k} is 0

δ_1 measures the impact of z_{1k} on the effect of the x_{1ijk} on P_{ijk}

U_{0k} and U_{1k} are error terms assumed to follow a multinormal distribution $\mathbf{N}(0,0;\Omega)$, Ω being the variance-covariance matrix. [12]

[12] In our model, we specify an unstructured form of the variance-covariance matrix (allowing the covariance between random effects to be non zero) because the covariance between the error terms of the intercept and the existence of upper educated people in the household appears to be significant – see below.

The reduced form is thus:

$$logit(P_{ijk}) = \gamma_0 + \gamma_1 z_{0k} + \delta_0 x_{1ijk} + \delta_1 z_{1k} x_{1ijk} + \beta_2 x_{2ijk} + U_{0k} + U_{1k} x_{1ijk}$$

This formula refers to a random slope model, meaning that the intercept and at least one of the explanatory variables have a random coefficient.

4. Data

The EU-SILC longitudinal dataset provides information at both individual and household levels. Data are available for the years 2003 to 2007 in some countries[13], but are not available for some countries in 2003, 2004 and 2007. Had we used all five waves to calculate the annual poverty threshold as we define it (i.e. at the European level), it would have risen or fallen over time simply because some countries (e.g. Germany) are absent for some years –i.e. without any link to the economic situation. As a consequence, we work with data from two waves only, namely 2005 and 2006, where all 26 countries are present.

The unit of analysis is the individual: as stated by an OECD[14] report (2001), this is the usual choice for poverty analysis with longitudinal data, because individuals can be followed over time whereas households cannot. The sample contains 131 891 working age adults (25-55) for the first wave, and 166 379 for the second wave[15], split between 26 countries (see Appendix 2). These countries[16] are in turn divided into 93 regions.

The explanatory variables have been chosen in order to control for different determinants of the poverty status. Some are micro factors, related to the demographic characteristics of the household (number of children and number of adults), others are related to the labour market (presence of at least one adult with an upper level of education, number of employed people), others still to the health status (presence of at least one adult with chronic disease, or hampered by illness in his/her daily activities). In fact, several variables were originally defined at the individual level (e.g. having a chronic disease). But, as the same poverty status is, according to the European definition, allocated to all individuals belonging to the same household, we have decided to define the micro factors exclusively at that level[17] (as other authors do – see Andriopoulou

[13] Note that EU-SILC is a rotational panel, meaning that the individuals are followed-up no longer than four years. As a consequence, an individual can be observed at the most 4 times when the sample is pooled.
[14] Organisation for Economic Cooperation and Development.
[15] Note that the sample is not balanced: 41 % of the individuals have only one observation (only 5% in Denmark, up to 59% in the Czech Republic). The 93 regions are not listed in this paper but can be found in Reinstadler and Ray (2010).
[16] The 2009 release of EU-SILC longitudinal data (August) contains only 22 countries, Germany, Ireland, Greece and Denmark being absent. But we wanted both to work on this release, the data of which had been cleared of previous problems, and to keep these four countries in the analysis. We have therefore added to these 22 countries the other 4 from the previous release (March 2009).
[17] Gender and age are the only exceptions to this rule. We have added these two variables to our model for comparison's sake, as almost all studies do this as well. As far as age is concerned, it could be argued that, even though age is, like gender, a factor measured at the individual level, it might have some signification as a household characteristic: due to frequent

et al, 2008). In order to have all variables defined at that same (household) level, we have tried to build aggregated variables, such as the number of adults in the household suffering from a chronic disease. Unfortunately, this information was available only for one individual per household in some countries (the register countries, which use information from administrative datasets when available and interview only one individual per household for the remaining questions to be asked). Keeping this kind of definition would therefore have resulted in a considerable loss of information. As a consequence, we have defined a much less precise indicator, such as 'presence in the household of at least one adult suffering from a chronic disease'. Note that even this imprecise indicator could be difficult to measure in those register countries, as only one household member is interviewed (and the construction of the variable would then rest only on that member). The level of education is defined for all household members aged 16 or over, but a lot of values are missing in some countries (13% in Portugal, 14% in Spain and up to 16% in the United Kingdom in 2005). We have thus adopted the same definition in order to construct a variable at household level.

Two additional variables are measured at the (macro) regional level: the GDP and the unemployment rate.[18] Some descriptive statistics for the whole sample can be found in Appendix 3 (and descriptive statistics for each country are to be found in Reinstadler and Ray, 2010).

endogamy, the age of one adult belonging to the household offers some clue about the age of other adults in the household, at least for non single adult households.

[18] Unfortunately, we were not able to integrate in our main model potential important macro factors such as the social expenditures (expressed as percentages of the GDP): unemployment compensation, public health expenditures, and expenditures with respect to inclusion. Indeed, they were not yet known at regional level.

Beyond the usual hypotheses concerning all the control variables[19], we make two further hypotheses concerning our two variables of interest: firstly we assume that the negative effect of the level of education on the probability of being at risk of poverty could be weaker in richer areas (where the probability of being poor is quite low, whatever the level of education). In order to test this hypothesis, we introduce interaction terms in the model between the level of education and the regional GDP. Secondly, we assume that the regional unemployment rate could affect the probability of being at risk of poverty.

[19] Note that income (and thus the poverty status) and individual/household demographic characteristics have not been measured at the same time: income refers to the year prior to the survey, whereas demographic characteristics to the time of the survey. We could have dealt with this by lagging all non-income variables (e.g. demographic characteristics given for year 2006 should be linked to the income declared in 2007), yet at the price of losing some countries since not all of them have available data for 2007.

5. Results and comments

The results of the model are shown in Table 1. All our analyses (descriptive and econometric) use weighted data[20].

Table 1: Probability of being at risk of poverty in 93 European regions. Estimation with a multilevel model[21] [22]

Variables	Parameter estimate	Standard error	Odds ratios
intercept	-0,4991	0,4061	0,6071
woman	0,0686 ***	0,0133	10,710
age centered (around the average : 41.27)	0,0058 ***	0,0008	10,058
age centered squared	0,0011 ***	0,0001	10,011
chronic disease in the household	-0,0021	0,0153	0,9979
activity hampered by disease in the household	0,3780 ***	0,0157	14,597
upper education level in the household	-1,5480 ***	0,1553	0,2127
number of children in the household	0,2389 ***	0,0146	12,699
number of children in the household squared	0,0292 ***	0,0040	10,296
number of adults in the household	-0,8989 ***	0,0257	0,4070
number of adults in the household squared	0,1309 ***	0,0042	11,399
regional annual GDP per capita (in 10^3 Euros)	-0,0439 ***	0,0075	0,9570
regional unemployment rate (expressed in %)	0,0442 ***	0,0062	10,452
upper education level * regional GDP per capita	0,0180 **	0,0066	
wave 2005	-0,1132 ***	0,0138	0,8930
wave 2006	ref.	ref.	ref.
country BE	-0,0209	0,4062	0,9793
country CZ	1,6501 ***	0,3725	52,075
country DK	-0,8197	0,4989	0,4406
country DE	-0,2769	0,4864	0,7581
country EE	2,9504 ***	0,5030	191,136
country IE	0,1860	0,4964	12,044
country EL	1,0960 **	0,3902	29,922
country ES	0,9450 **	0,3557	25,728
country FR	-0,2432	0,3544	0,7841

[20] Moon and Stotsky (1993) state that, if the data come from a stratified and clustered random sampling, it is reasonable to treat the sample as a simple random sample, thus ignoring weights. And Poggi (2007) states that it is more efficient, from an econometrical point of view, not to weigh the data. That way of doing has been adopted by studies on panel data (as stated by Ayllon, 2008, or Andriopoulou *et al*, 2008). But the question of weighting the data is still open. We have decided to weigh them, as the EU-SILC dataset for most countries is not representative of the population due to the over-sampling of some subpopulation and/or the selective non-response.

[21] We have used the SAS GLIMMIX command. Useful SAS code examples can be found in Allison (2008) and were kindly made available to us by David Brady.

[22] See appendix 3 for a description of the variables.

	Parameter estimate	Standard error	Odds ratios
country IT	0,4480	0,3782	15,652
country CY	-0,3412	0,5341	0,7109
country LV	3,3363 ***	0,5047	281,149
country LT	3,7109 ***	0,5033	408,906
country LU	-0,3582	0,7582	0,6989
country HU	3,0490 ***	0,4074	210,942
country NL	-0,3064	0,4879	0,7361
country AT	-0,4327	0,4038	0,6488
country PL	2,7027 ***	0,3930	149,200
country PT	1,4412 **	0,4920	42,258
country SI	0,2115	0,5011	12,355
country SK	2,8819 ***	0,4995	178,482
country FI	-0,4769	0,3964	0,6207
country SE	-0,6395	0,4914	0,5276
country UK	ref.	ref.	ref.
country IS	-0,6700	0,6945	0,5117
country NO	-0,4789	0,5146	0,6195
Regional-level error terms variances			
intercept	0,1377	0,0312	
upper education level in the household	0,2067	0,0438	
Regional-level error terms covariance			
COV (intercept, upper education level in the household)	-0,0650	0,0313	
Other parameters			
Rho coefficient of AR(1)	0,4192	0,0025	
Residual	0,9533	0,0027	
Fit measure: -2 Log Pseudo Likelihood	1731151		

Source: EU-SILC Users' database, longitudinal file, 1.08.2009 release, 22 countries from this release plus 4 countries from the March 2009 release (see above), authors' computations.
Level of significance for independent variable coefficients: *: p-value < 0.05; **: p-value < 0.01; ***: p-value < 0.001

Unfortunately, the SAS/PROC GLIMMIX command does not offer a statistical test indicating the level of significance of the variances and covariances of the error terms. However, compared to their standard errors, the estimated variances are quite high, which suggests their high level of significance. This in turn justifies, on the one hand, the choice of the multilevel model and, on the other hand, our choice to allow the intercept and the variable in question to have random rather than fixed coefficients. Looking at the empty model (see Appendix 4), we can see that the intra-class correlation (calculated according to the second formula given by Snijders and Bosker, 1999, page 224) is equal to 0.55, meaning that the between-variance is substantial.

Let us now examine the effects of our variables of interest. Remember that our objective is to measure the specific effect of the regional GDP per inhabitant and the regional unemployment rate on the probability of being at risk of poverty. For the first of these two macro variables, this effect could be either direct or indirect since it can act through one individual variable (the education level) on the probability of being poor. For the second variable (the regional unemployment rate), the potential effect is direct only.

As expected, the regional GDP per capita has a strong (and highly significant) direct negative effect on the risk of poverty: for individuals living in households where nobody has a higher level of education, the odds of being poor (probability of being poor divided by probability of not being poor) decreases by 4.3% (1-0.957=0.043) for an increase of 1000 Euros in annual GDP per capita. This direct effect is supplemented by an indirect effect: the regional GDP per capita moderates the negative impact of higher education on the poverty risk. In fact, in the average region in terms of GDP per capita (about 24260 Euros/year), the presence of an adult with a higher education level decreases the poverty odds by 67% (odds ratio = 0.33[23]). In a rich region such as Luxembourg (GDP per capita = 60150 Euros/year), it decreases the odds by 37%; in a quite disadvantaged region like Estonia (GDP per capita = 14547 Euros/year) it decreases the odds by 72 %. In other words, the moderating effect of the regional GDP per capita on the impact of the presence of highly educated people on the poverty risk is quite large.

As for the regional unemployment rate, an additional percentage point increases the poverty odds by 5% ceteris paribus (especially when GDP per capita is controlled for). The sign of this effect was expected, but knowing its extent is interesting too.

Besides, and not surprisingly (given the sample size), almost all control variables - even gender - have an effect on the poverty probability: women have a slightly higher risk of being poor than men. One interesting aspect is the extent to which these control variables have an impact on the poverty risk (even if our study focuses on the possible impact of macro determinants on the effect that some factors of interest can have on the poverty risk):

- ceteris paribus, the poverty risk first decreases more than linearly with age, reaching a minimum at age 39, and then increases more than linearly. For example, at age 25, an additional year results in a decrease of the poverty risk by 3.3%; and at age 50, an additional year of age results in an increase of 2.6%

- if the activity of at least one household member is hampered by disease, the odds of being poor increase by 46%

[23] Odds ratio = exp(-1.5480 + 0.01803*24.260) = 0.33

- the effect of the number of children in the household is not linear: the first child increases the poverty odds by 31% while the fourth child increases them by 58%

- the number of adults in the household also proves to have a non linear effect on the poverty risk: ceteris paribus, the odds of being poor decrease by 40% with the second adult in the household, are virtually the same for two or three adults in the household, and then increase (for example by 32% with the fifth adult).

6. Conclusion

Many studies have attempted to analyse the determinants of the monetary poverty probability. However, few of them have simultaneously used panel data, considered factors at the macro level, and used multilevel modelling with three levels to deal with all these elements.

As for our results, they show that both the regional GDP per capita and the regional unemployment rate do have an effect on poverty risk.

In terms of economic and social policy implications, it means that:

- policies oriented towards higher economic growth rates in disadvantaged European regions are able to alleviate the risk of poverty, even if poverty is defined in relative terms;

- economic policies of this kind, if successful in their effort to sustain the economic well-being of families in poor regions, will, as a side-effect, diminish the anti-poverty effect of the presence of higher educated people in the household. We suspect that this indirect effect is associated with the choice of defining poverty as a relative concept – a European view, which is not shared by countries like the US;

- as for the regional unemployment rate, its direct positive impact on the poverty risk is essentially a confirmation of what was to be expected and of what is already known, even if the weakness of this effect is quite surprising.

However, our analysis faces two types of limitations. The first results from the methodological choices we have made, while the second is due to the data.

Firstly, because we needed an indicator to differentiate and thus to rank the 93 regions in terms of poverty rates and not in terms of inequality, we have made use of a European poverty threshold, which has proved to be quite relevant in terms of its ability to estimate the econometric model. However, precisely because the European regions are quite dispersed around the average at-risk-of-poverty rate, we were not able to check the consistency of the results by using alternative measures of the European threshold (such as 50% or 70% of the European median equivalent income). In fact, with a poverty threshold equal to 60% of the European median equivalent income, the at-risk-of-poverty rates of the different countries range between 1% and 82% in 2005, and between 2% and 75% in 2006. Changing this threshold for a lower (higher) one would lead to even lower (higher) rates in the richest (poorest) countries, making it impossible to run the analysis.

Concerning the data, we would have liked to test the effect of other macro characteristics at the regional level, such as the expenses in unemployment or social benefits (expressed in percentage of the GDP). However, these characteristics were not available at the regional level. In future analyses, we would therefore be interested in adding some variables of that kind, once they are available at the regional level. And, as for the regions, the variable defined in the EU-SILC dataset is missing in four countries (even in quite large countries such as Germany and the United Kingdom). In order to keep these countries in the analysis, we have defined each of them as a single – and quite large – region.

References

ALLISON Paul D. (1984), Event History Analysis, Sage University Paper Series on Quantitative Applications in the Social Sciences, n°07-046, Beverly Hills, CA, Sage, 87 pages.

ALLISON Paul D. (2008), Longitudinal Data Analysis Using SAS, October, 100 pages.

ANDRIOPOULOU Eirini, **TSAKLOGLOU Panos** (2008), 'Once poor, always poor? Do initial conditions matter? Evidence from the ECHP', IRISS Conference, Oct., 33 pages.

ANGRIST Joshua D., PISCHKE Jörn-Steffen (2009), 'Mostly Harmless Econometrics. An Empiricist's Companion', Princeton University Press, 373 p.

AYLLON Sara (2008), 'Modelling poverty transitions in Spain: Do attrition and initial conditions really matter?', IRISS Working Paper Series, 2008-08, Oct., 36 pages.

BRADBURY Bruce, **JÄNTTI Markus** (2001), 'Child Poverty across twenty-five countries', Bradbury B., Jenkins S., Mickelwright J. (eds), *The Dynamics of Child Poverty in Industrialized Countries*, chapter 4, pp. 62-91, Cambridge, Cambridge University Press.

BRADSHAW Jonathan (1999), 'Child Poverty in Comparative Perspective', *European Journal of Social Security*, Vol.1, pp. 383-406.

BRADY David, **FULLERTON Andrew S.**, **MOREN-CROSS Jennifer** (2009), 'Putting Poverty in Political Context: A Multi-Level Analysis of Adult Poverty Across 18 Affluent Countries', *Social Forces*, Vol. 88(1), September, pp. 271-300.

BUDDELMEYER Hielke, **VERICK Sher** (2008), 'Understanding the Drivers of Poverty Dynamics in Australian Households', *The Economic Record*, Vol. 84, n°266, pp. 310-321.

CALLENS Marc, **CROUX Christophe** (2009), 'Poverty Dynamics in Europe: A Multilevel Recurrent Discrete-Time Hazard Analysis', *International Sociology*, May, Vol. 24(3), pp. 368–396.

CAPPELLARI Lorenzo (2004), 'Earnings Mobility Among Italian Low Paid Workers', *Journal of Population Economics*, Vol. 20(3), pp. 465-482.

CAPPELLARI Lorenzo, **JENKINS Stephen P**. (2002a), 'Who Stays Poor? Who Becomes Poor? Evidence from the British Household Panel Survey', *The Economic Journal*, Vol. 112, No. 478, Conference Papers (Mar., 2002), pp. C60-C67.

CAPPELLARI Lorenzo, **JENKINS Stephen P.** (2002b), 'Modelling Low Income Transitions', *ISER Working Papers Number 2002-8*, 42 pages.

CAPPELLARI Lorenzo, **JENKINS Stephen P.** (2004), 'Modelling Low Income Transitions', *Journal of Applied Econometrics, Vol.* 19, pp**.** 593–610.

FERTIG Michael, TAMM Marcus (2007), 'Always Poor or Never Poor and Nothing in Between? Duration of Child Poverty in Germany', IZA Discussion Paper N°2645, February, 29 pages.

MARLIER Eric, **ATKINSON Anthony B.**, **CANTILLON Bea**, **NOLAN Brian** (2007), The EU and Social Inclusion – Facing the challenges, The Policy Press, The University of Bristol, 303 pages.

MEJER Lene, **SIERMANN Clemens** (2000), 'La pauvreté monétaire dans l'Union européenne : la situation des enfants, les différences entre les sexes et l'écart de pauvreté', Eurostat, *Statistiques en bref*, Population et Conditions sociales, 12/2000, 7 pages.

MOLLER Stephanie, **HUBER Evelyne**, **STEPHENS John D.**, **BRADLEY David**, **NIELSEN François** (2003), 'Determinants of Relative Poverty in Advanced Capitalist Democracies', *American Sociological Review*, Vol. 68, No. 1, February, pp. 22-51.

MOON Choon-Geol, **STOTSKY Janet G.** (1993), 'The Effect of Rent Control on Housing Quality Change: A Longitudinal Analysis', *The Journal of Political Economy*, Vol. 101, No. 6, December, pp. 1114-1148.

OECD (2001), 'When Money is Tight: Poverty Dynamics in OECD Countries', in *Employment Outlook*, Chapter 2, pp. 37-87.

POGGI Ambra (2007), 'Does persistence of social exclusion exist in Spain? ', *Journal of Economic Inequality*, Vol. 5, pp. 53-72.

RAY Jean-Claude, **JEANDIDIER Bruno** (2003), 'Depuis une décennie, que nous ont appris les données longitudinales à propos de la pauvreté en France ? Une première synthèse', Communication aux Xes Journées du Longitudinal, Caen, mai, 61 pages.

REINSTADLER Anne, **RAY Jean-Claude** (2010), Macro Determinants of Individual Income Poverty in 93 Regions of Europe, CEPS Working Paper.

SNIJDERS Tom A.B., **BOSKER Roel J.** (2004), *Multilevel Analysis. An introduction to basic and advanced multilevel modelling*, Sage Publications, 266 pages.

STEWART Mark B., **SWAFFIELD Joanna K.** (1999), 'Low Pay Dynamics and Transition Probabilities', *Economica*, Vol. 66, pp. 23-42.

TAI Tsui-o, **TREAS Judith** (2008), 'Poverty, Household Composition, and Welfare States: A Multi-level Analysis of 22 Countries', Luxembourg Income Study Working Paper Series, Working Paper No. 492, September, 33 pages.

UNICEF (2000), 'A League Table of Child Poverty in Rich Nations', Innocenti Report Card, Issue n°1, June, 28 pages.

VAN KERM Philippe (2004), 'Une évaluation économétrique des flux vers et hors de la pauvreté en Belgique', IRISS Working Paper Series, 2004-04, 17 pages.

WIEPKING Pamala, **MAAS Ineke** (2005), 'Gender Differences in Poverty: A Cross-National Study', *European Sociological Review*, Vol. 21(3), July, pp. 187-200.

Appendices

Appendix 1: At-risk-of-poverty rates of individuals aged 25-55 in Europe with our definition

Table 2: At-risk-of-poverty rates of individuals aged 25-55 in Europe (European threshold = 60% of the European median equivalent income)

Country	2005	2006
Austria (AT)	5	6
Belgium (BE)	6	8
Cyprus (CY)	7	6
Czech Republic (CZ)	42	41
Germany (DE)	7	6
Denmark (DK)	3	3
Estonia (EE)	70	64
Spain (ES)	20	19
Finland (FI)	5	6
France (FR)	8	9
Greece (EL)	24	26
Hungary (HU)	73	69
Ireland (IE)	8	8
Iceland (IS)	3	3
Italy (IT)	14	17
Lithuania (LT)	79	75
Luxembourg (LU)	1	2
Latvia (LV)	81	74
Netherlands (NL)	6	3
Norway (NO)	3	3
Poland (PL)	74	73
Portugal (PT)	42	40
Sweden (SE)	5	5
Slovenia (SI)	13	13
Slovakia (SK)	77	72
United Kingdom (UK)	7	8

Source: EU-SILC Users' database, longitudinal file, 1.08.2009 release, authors' computations.

Reading note: with the European poverty threshold calculated for the whole population[24], 5% of individuals aged 25-55 in Austria were at risk of poverty in 2005.

We can note that the countries face very different situations in terms of the at-risk-of-poverty rate, a conclusion which cannot be drawn from the figures resulting from the official calculations.

[24] The European poverty threshold is calculated by taking into account all individuals living in the 26 countries covered by the study. In other words, children and elderly people are not excluded from this calculation, even if they are not subsequently kept in the analyses.

Appendix 2: Sample size, by country and year

Table 3: Number of individuals aged 25-55, in each country, for each year (sample size, unweighted cases)

Country	2005	2006
Austria (AT)	3882	5469
Belgium (BE)	2408	4002
Cyprus (CY)	2357	3373
Czech Republic (CZ)	4391	7367
Germany (DE)	10529	9285
Denmark (DK)	3549	3451
Estonia (EE)	2665	4269
Spain (ES)	8472	11501
Finland (FI)	4021	5528
France (FR)	5876	7324
Greece (EL)	4642	4211
Hungary (HU)	4353	6609
Ireland (IE)	2872	1992
Iceland (IS)	1584	2243
Italy (IT)	12955	18048
Lithuania (LT)	2533	3713
Luxembourg (LU)	4322	4601
Latvia (LV)	2704	3438
Netherlands (NL)	8574	9287
Norway (NO)	4088	3926
Poland (PL)	10627	14644
Portugal (PT)	2671	3662
Sweden (SE)	3541	4751
Slovenia (SI)	8671	11271
Slovakia (SK)	3447	4898
United Kingdom (UK)	6224	7594

Source: EU-SILC Users' database, longitudinal file, 1.08.2009 release, authors' computations.

Appendix 3: Description of the explanatory variables and descriptive statistics

Table 4: Description of the explanatory variables

Name of the variable	Label of the variable	Description of the variable
woman	*woman*	EU-SILC variable RB090; woman=1 if RB090=2
age_centered	*age centered (around the average: 41.27)*	EU-SILC variable RX020, centered (age in the year prior to the survey)
chronicdiseaseHH	*chronic disease in the household*	Authors' calculations using the EU-SILC variable PH020: is there at least one household member who suffers from a chronic disease?
activityhamperedHH	*activity hampered by disease in the household*	Authors' calculations using the EU-SILC variable PH030: is there at least one household member whose activities are hampered because of health problems?
uppereducHH	*upper education level in the household*	Authors' calculations using the EU-SILC variable PE040: is there at least one household member whose upper level of education is tertiary education (PE040=5)?
nbchildren	*number of children in the household*	number of children (age 0-14) in the year prior to the survey
nbadultsHH	*number of adults in the household*	number of adults (age 18 or more) in the year prior to the survey
nbemployedHH	*number of employed people in the household*	number of employed household members (authors' calculations using the EU-SILC variable PL030 – codes 1 or 2)
country	*country*	EU-SILC variable RB020
wave	*wave*	EU-SILC variable RB010
GDPhabnuts	*regional annual GDP per capita (in 10^3 Euros)*	Information from Eurostat
unempratenuts	*regional unemployment rate (expressed in %)*	Information from Eurostat

Table 5: Descriptive statistics for the whole sample

Wave	N Obs	Variable	N	NMiss	Mean	Std Dev	Minimum	Maximum
2005	131891	pov_indicator	131891	0	0.1965	0.4259	0	1.0000
		woman	131887	4	0.5130	0.5357	0	1.0000
		age_centered	131891	0	-0.6425	9.1801	-16.2665	14.7335
		age_centered2	131891	0	73.7935	76.2358	0.0710	264.6
		chronicdiseaseHH	131891	0	0.3885	0.5223	0	1.0000
		activityhamperedHH	131891	0	0.2931	0.4878	0	1.0000
		uppereducHH	126795	5096	0.3962	0.5182	0	1.0000
		nbchildren	131891	0	0.9077	1.1426	0	11.0000
		nbadultsHH	131891	0	2.3903	1.0792	1.0000	10.0000
		nbemployedHH	129384	2507	1.5131	0.8971	0	8.0000
		GDPhabnuts	131891	0	23.7538	7.1961	8.2000	57.1000
		unempratenuts	131891	0	8.8728	4.4381	2.5000	21.4000
2006	166379	pov_indicator	166379	0	0.2000	0.3757	0	1.0000
		woman	166371	8	0.5128	0.4695	0	1.0000
		age_centered	166379	0	-0.2041	7.9962	-16.2665	14.7335
		age_centered2	166379	0	72.5178	66.5149	0.0710	264.6
		chronicdiseaseHH	166379	0	0.3911	0.4584	0	1.0000
		activityhamperedHH	166379	0	0.2926	0.4273	0	1.0000
		uppereducHH	158892	7487	0.4059	0.4560	0	1.0000
		nbchildren	166379	0	0.9288	0.9917	0	12.0000
		nbadultsHH	166379	0	2.4239	0.9506	1.0000	11.0000
		nbemployedHH	162220	4159	1.6016	0.7713	0	8.0000
		GDPhabnuts	166379	0	24.7924	6.6164	8.7000	63.1000
		unempratenuts	166379	0	8.1568	3.0558	2.8000	21.0000

Appendix 4: Empty model (weighted)

Response profile

Ordered value	Pov_indicator	Total frequency
1	1	76252
2	0	222018

The GLIMMIX procedure is modeling the probability that pov_indicator='1'.

Dimensions	
G-side cov. parameters	1
R-side cov. parameters	2
Columns in X	1
Columns in Z per subject	1
Subjects (blocks in V)	93
Max Obs per subject	19942

Optimization Information	
Optimization technique	Newton-Raphson
Parameters in optimization	2
Lower boundaries	2
Upper boundaries	1
Fixed effects	Profiled
Residual variance	Profiled
Starting from	GLM estimates

6 iterations

Convergence criterion (PCONV=1.11022E-8) satisfied

Fit statistics	
-2 Res Log Pseudo-Likelihood	1738618
Generalized Chi-Square	294644.1
Gener. Chi-Square / DF	0.99

Covariance parameter estimates

Cov parm	Subject	Estimate	Std error
Intercept	region	1.7720	0.2639
AR (1)	ID_unique_UE (region)	0.4571	0.002340
Residual		0.9878	0.002734

Asymptotic covariance matrix of covariance parameter estimates				
Cov Parm	Subject	CovP1	CovP2	CovP3
Intercept	region	0.06965	-2.32E-7	-2.12E-7
AR (1)	ID_unique_UE (region)	-2.32E-7	5.477E-6	2.254E-6
Residual		-2.12E-7	2.254E-6	7.473E-6

Asymptotic correlation matrix of covariance parameter estimates				
Cov Parm	Subject	CovP1	CovP2	CovP3
Intercept	region	1.0000	-0.00038	-0.00029
AR (1)	ID_unique_UE (region)	-0.00038	1.0000	0.3524
Residual		-0.00029	0.3524	1.0000

Solutions for fixed effects					
Effect	Estimate	Std error	DF	t - value	Pr > ltl
Intercept	-1.4902	0.1387	92	-10.74	<.0001

The empty model contains only a random intercept. The dependent variable is thus explained by the overall mean, a random term at group level and a random term at individual level. The empty model enables the relative parts of the variance between groups to be identified (regions in this case) and the variance within groups. It is thus possible to calculate the intraclass correlation coefficient (ICC) (here: 55.1% = 1.772/3.2169, i.e. the intercept variance divided by the total variance).

European Commission

Macro determinants of individual income poverty in 93 regions of Europe

Luxembourg: Publications Office of the European Union

2010 — 30 pp. — 21 x 29.7 cm

ISBN 978-92-79-16746-1
ISSN 1977-0375
Doi:10.2785/53721
Cat. No KS-RA-10-012-EN-N

Theme: Population and social conditions

Collection: Methodologies and working papers

How to obtain EU publications

Free publications:

- via EU Bookshop (http://bookshop.europa.eu);
- at the European Union's representations or delegations. You can obtain their contact details on the Internet (http://ec.europa.eu) or by sending a fax to +352 2929-42758.

Priced publications:

- via EU Bookshop (http://bookshop.europa.eu).

Priced subscriptions (e.g. annual series of the *Official Journal of the European Union* and reports of cases before the Court of Justice of the European Union):

- via one of the sales agents of the Publications Office of the European Union (http://publications.europa.eu/others/agents/index_en.htm).

KS-RA-10-012-EN-N

ISBN 978-92-79-16746-1

9 789279 167461

Publications Office