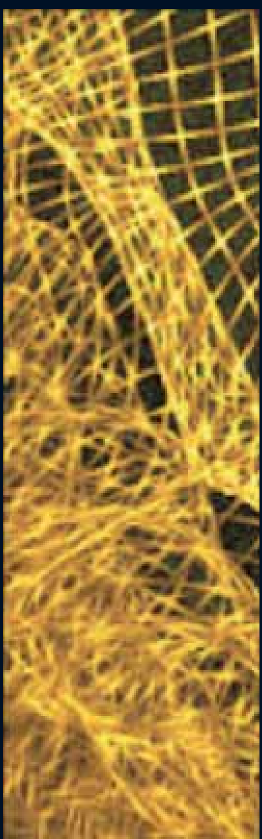
imdea ciencias sociales



working papers series
in Economics and Social Sciences
2007/17

Education and income inequality in the regions of the European Union

by Andrés Rodríguez-Pose and Vassilis Tselios

September 2007

Ctra. Colmenar Viejo km. 14, 28049 Madrid, Spain

http://www.imdea.org/socialsciences
http://repec.imdea.org
contacto.cienciassociales@imdea.org

www.imdea.org

Education and income inequality in the regions of the European Union

by

Andrés Rodríguez-Pose and **Vassilis Tselios**

Department of Geography and Environment
London School of Economics
Houghton St
London WC2A 2AE, UK
Tel: +44-(0)20-7955 7971
Fax: +44-(0)20-7955 7412
E-mail: a.rodriguez-pose@lse.ac.uk, v.tselios@lse.ac.uk

Education and income inequality in the regions of the European Union

Abstract

This paper provides an empirical study of the determinants of income inequality across regions of the EU. Using the European Community Household Panel dataset for 102 regions over the period 1995-2000, it analyses how microeconomic changes in human capital distribution affect income inequality for the population as a whole and for normally working people. The different static and dynamic panel data analyses conducted reveal that, while the relationship between income inequality and income per capita is positive, the relationship between income inequality and educational attainment is not clear. Across European regions high levels of inequality in educational attainment are associated with higher income inequality. This may be interpreted as the responsiveness of the EU labour market to differences in qualifications and skills. The above results are robust to changes in the definition of income distribution. Other results indicate that population ageing and inactivity are sensitive to the specification model, while work access and latitude are negatively associated to income inequality. Urbanisation has a negative impact on inequality, but for the population as a whole only, and the relationship between unemployment and income inequality is positive. Female participation in the labour force is negatively associated with inequality and explains a major part of the variation in inequality. Finally, income inequality is lower in social-democratic welfare states, in Protestant areas, and in regions with Nordic family structures.

Keywords: Income inequality, educational attainment, educational inequality, regions, Europe

1. Introduction

It is often claimed that greater educational attainment makes societies more egalitarian, and that income and educational inequalities are perfectly correlated (Checchi 2000). But, in spite of these claims, the influence of education on inequalities is still a long way from being perfectly understood, especially at a regional level. This paper addresses the questions of the supposed negative relationship between educational attainment and income inequality and of the positive correlation between inequality in education and in income for the regions of the EU. Our methodology is based on the estimation of various specification models (both static and dynamic) in order to assess the sensitivity of the relationships.

This aim of the paper is to analyse how microeconomic changes in human capital distribution affect income inequality, not only for the population as a whole, but also for normally working people. In this paper human capital distribution is measured in terms of both average education of the population and inequality in educational attainment. By analysing the microeconomic processes underpinning the relationship between individual educational endowments and income inequality, we also expect to draw greater light on whether government education policies contribute to a more equal income distribution and whether the EU labour market is responsive to differences in qualifications, knowledge, and skills.

The paper is organised in five additional sections. The next section reviews the existing debate over the determinants of income inequality, putting greater emphasis on the relationship between income and educational distribution. The empirical regression model and the relevant static and dynamic estimation methods are discussed in Section 3. Section 4 describes the data and the construction of variables. Section 5 reports and discusses the regression results and, finally, Section 6 concludes with policy recommendations and some suggestions for further research.

2. Education and Income Inequality: Theoretical Considerations

Given the vast body of literature on the determinants of income inequality, the aim of this section is not to review this vast array of sources, but simply to focus on how the impact of income per capita, as well as of average education levels and inequality in that area, on income inequality is perceived by the literature. To achieve that aim, we first review the link between income and inequality, before going on to analyse the impact of educational attainment and inequality on income inequality. We also consider the dynamic structure of inequalities.

Changes in the distribution of income take place at a very slow pace. There are several reasons for this. First, people are often reluctant to change jobs for psychological and institutional reasons (Gujarati 2003). Additionally, income levels are often perpetuated from one generation to another by means of inheritance, cultural background, and, more generally, the characteristics of the community (Bourguignon and Morrisson 1990; Cooper, Durlauf et al. 1994; Durlauf 1996; Checchi 2000). This allows for intergenerational stability in income, indicating the existence of a positive autocorrelation in inequalities. Cooper (1998), for instance, has pointed out that poorer or more wealthy families tend to exhibit a greater degree of intergenerational income stability than middle income families. Hence, it is often the case that a proportion of the population remains trapped at the same level of income for more than one generation. Income persistence is often viewed (i.e. Lane 1971) as an essential characteristic in rewarding achievement and, particularly, in ensuring that the most suitable people are allocated the most suitable roles. The presence of inequalities in income provides an additional incentive for achievement and innovation, which are an integral part of modern society. Some degree of inequality is generally perceived as a necessary constituent of a healthily functioning economy (Champernowne and Cowell 1998: 14). Yet, according to Aghion and Bolton (1992) and Galor and Zeira (1993), the persistence of income inequalities across generations is possible only if capital markets are imperfect. High intergenerational correlations imply less mobility in the distribution of income. The key question is whether the persistence of inequality has an impact on economic performance. Do unequal societies perform better than more equal ones or is it vice versa?

This relationship has been most famously addressed by Kuznets (1955). According to Kuznets (1955) income per capita was found to have an inverted U-curve effect on income inequality. Income inequality increases as nations begin to industrialise and, then, declines at later stages of industrialisation. This relationship is known as ‘Kuznets curve’ and was formalised by Knight (1976a; 1976b), Robinson (1976), and Fields (1979). The Kuznets curve shows that in the early stages of industrialisation, the labour force is primarily engaged in agriculture. As industrialisation takes hold, workers move from the larger agricultural sector to the smaller industrial one and, since wages are usually higher in the industrial sector, this migration boosts further income inequality (Firebaugh 2003). Income distribution thus becomes more unequal as income increases. Moreover, as the agricultural sector shrinks and industry increases in size, further transfers from agriculture to manufacturing reduce, rather than increase, income inequality.

The key factors underlying the inverted U-curve effect of income per capita on inequality are industrialisation and labour migration. The additional factors behind this association include market and government failures, government social expenditures, and the development of financial services. For example, De Gregorio and Lee (2002) show that income inequalities are negatively correlated with government social expenditure. Schultz (1962) indicated that modifications in income transfers and in progressive taxation are relatively weak factors in altering the distribution of income. Motonishi (2000; 2006) argues that the effect of financial service development on income inequalities is not straightforward. On the one hand, more developed financial services enable the poor to borrow from the rich and this leads to a decrease in income inequality; while, on the other hand, financial services are often not available to the poor due to constraints on the credit market arising from information asymmetries. Finally, market failures, such as credit constraints and monopsony or monopoly power and government failures, often increase income inequalities (Graham 2002).

Despite the significant amount of research that has set out to test the Kuznets curve at the national level, the results are ambiguous (i.e. Ahluwalia 1976; Papanek and Kyn 1986; Anand and Kanbur 1993; Bourguignon and Morrisson 1998; Checchi 2000; Motonishi 2006). Ahluwalia (1976), for instance, finds for a cross-section of counties evidence to support the inverted U-curve, while Anand and Kanbur (1993) report that the Kuznets curve is not inverse at all. Overall, the literature seems unable to provide conclusive empirical results on the relationship between income inequality and per capita income, since social structures, such as historical heritage, religion, ethnic composition, and cultural traditions, evolve differently across countries (Checchi 2000). In this paper, we do not expect to test the validity of the Kuznets curve for two reasons. Firstly, the majority of the relevant empirical studies focus not only on European but also on less economically advanced countries (i.e. African countries). Secondly, the studies in question show that the declining segment of the Kuznets curve begins around 1970 (Nielsen and Alderson 1997). However, we use Kuznets' theory in order to assume a linear association between income per capita and income inequality for developed countries over a relatively limited period of time. We therefore expect to find that over the period 1995–2000 income per capita was negatively associated with income inequality.

The notion of education as an underlying factor in income differences also has a long history, dating back to the work of Adam Smith (Griliches 1997). Based on the work of Schultz (1961; 1962; 1963), Becker (1962; 1964) and Mincer (1958; 1962; 1974), income inequality is generally considered to be affected by educational attainment, in a process which is sometimes referred to as 'skills deepening' (Williamson 1991). A higher level of educational attainment is achieved through improvements in access to education (i.e. lower tuition fees, better education financing, improved vocational training), a higher quality of education (i.e. better services from teachers, librarians, and administrators), and greater investment in physical capital for education. Improved access to education, for example, is likely to increase the earning opportunity of the lowest strata, leading to a reduction in earning inequality (Checchi 2000).[1] Furthermore, more widespread access to education allows for a more informed participation in the market economy, reducing the lobbying ability of the rich, while simultaneously increasing the social and job opportunities of the poor, implying lower inequality. Education is thus regarded as one of the most powerful instruments known for reducing income inequality (World Bank 2002).

According to Knight and Sabot (1983), the impact of educational attainment on income inequalities depends on the balance between the 'composition' and the 'wage compression'

[1] Income inequality, at least in industrialised countries, is explained by a rise in earning inequality (Gottschalk and Smeeding 1997; Cornia, Addison et al. 2001). Hence inequality in pay is an important component of total income inequality (Blinder 1974; Brown 1977).

effect. Concerning the ‘composition’ effect, an increase in the levels of education of the population tends, at least initially, to increase income inequality. With respect to the ‘wage compression’ effect, over time education leads to decreased income inequality. An increase in the level of education reduces the wages of highly-educated workers, because their supply goes up, and simultaneously raises the wages of the less-educated workers, because their supply goes down. Hence, an increase in the educated labour supply is likely to increase competition for positions requiring advanced educational credentials and thereby should reduce the income differential between the educated and uneducated people (Tinbergen 1975; Lecaillon 1984). Moreover, an increased proportion of the population attaining a higher level of education leads to inflation in the value of educational credentials and, in the long run, to decreasing wages for highly-educated workers. Thus, the effect of education on income inequality is based on a balance of supply and demand.

Spence’s (1973; 1974; 1976) signalling model offers a different perspective on the relationship between income and education. This model demonstrates that education has no direct effect on income distribution, because education acts as a ‘label’ or ‘signal’. More specifically, his model posits a situation in which the possibility of higher pay for more educated people has little to do with academic and vocational skills, because formal education is seen as an elaborate device for detecting and labelling those who have skills (Champernowne and Cowell 1998; Wolf 2004). The individual’s education level is more closely related to innate ability and to psychological and personality traits, such as diligence, and these are what employers reward, rather than regarding education as a means of instilling or enhancing skills (Wolf 2004). Differences in educational attainment may arise as a consequence of heterogeneity in ability. Galor and Tsiddon (1997b) and Hassler and Mora (2000), for example, support the idea that individuals with a higher level of innate cognitive ability can fare better with less knowledge than others do. For them, genetic characteristics are highly correlated with the education that children receive and their skills. In contrast, López, Thomas et al. (1998) support the notion that education levels are not necessarily correlated with abilities. Nevertheless, education still works as a marker for achieving better jobs.

To sum up, given the complexity of the relationship between education and income, it is difficult to predict a priori the sign and the significance of the relationship between educational attainment and income inequality.

On the relationship between educational inequality and income inequality most theoretical analyses tend to report that both factors are positively correlated (Jacobs 1985; Glomm and Ravikumar 1992; Saint-Paul and Verdier 1993; Galor and Tsiddon 1997a; Chakraborty and Das 2005). More explicitly, Thorbecke and Charumilind (2002: 1488) have pointed out that, with regard to the supply side of skilled labour education, a greater share of highly-educated workers within a cohort may signal to employers that those with less education have less ability, and hence the latter's earnings may be reduced accordingly, which may also lead to a greater wage inequality between workers with high and low levels of education. With respect to the demand side of skilled labour education, if the demand for unskilled labour is either contracting or growing at a slower rate than the demand for skilled labour, then earning inequalities will increase. Finally, the empirical studies of Becker and Chiswick (1966) and Park (1996) show that a higher level of educational attainment among the labour force has an equalising effect on income distribution, and that the greater the inequality in educational attainment, the greater the income inequality.

3. Econometric Approach

As a means to test the relationship between education and income inequality in a European regional context, we use microeconomic data as a means to estimate income inequality as a linear function of per capita income, educational attainment, and educational inequality. We use different empirical specifications in order to assess the robustness of the econometric models and to examine the impact of adding control variables, such as population ageing, work access, unemployment, and inactivity. The methodology incorporates variability both across regions (N) and over time (T) in a pooled cross-sections analysis. Our emphasis is on the case where $N \rightarrow \infty$ with T fixed and on the one-way error component model, due to the limited number of observations. Different panel data analyses are conducted in order to reduce measurement error on inequalities and minimise potential problems of omitted-variable bias. Panel data also allow for greater degrees of freedom than with time-series or cross-regional data and improve the accuracy of parameter estimates (Hsiao 2003; Baltagi 2005).

This study deals with two methods of panel regression analysis: *static* and *dynamic models*. These models are increasingly popular for panel data analysis among regional scientists. With repeated observations for 102 regions, panel analysis permits us to study the dynamics of change with short-time series. The static models endow regression analysis with both a spatial and temporal dimension. The first dimension pertains to a set of cross-regional units of observation, while the second one pertains to periodic observations of a set of variables characterising these cross-regional units over a particular time span. We use a pooled regression model as the baseline for our comparison. As the surveys of the European Community Household Panel (ECHP) dataset were conducted regularly at approximately one-year interval, the error terms of inequality regressions are expected to be correlated with the regional-specific effect. This can be dealt with the fixed effects (FEs) panel data analyses in which the error terms may be correlated with the regional-specific effects. This study also includes dynamic models due to the short-time period of analysis. For instance, the equilibrium may be constrained in the short-run because of supply rigidities or factor immobilities that are removed in the longer-run (Combes, Duranton et al. 2005). The dynamic models test for the existence of autocorrelation. In these models, finally, we can obtain both short-run and long-run parameters. To sum up, in order to examine the impact of education on income inequality and to evaluate the robustness of the results, we experiment with a number of alternative specifications and include additional determinants to our equations.

Our econometric analysis starts with a static panel data model of the form:

$$y_{it} = \beta' x_{it} + \gamma' z_i + \nu_i + \varepsilon_{it}$$

with i denoting regions ($i = 1,..., N$) and t time ($t = 1,...,6$)[2]. y_{it} is income inequality, x_{it} is a vector of explanatory variables, z_i is a vector of time-invariant explanatory variables

[2] $t = 1$ denotes 1995, …, $t = 6$ denotes 2000.

(i.e. urbanisation and latitude), β and γ are coefficients, ν_i is an unobserved regional-specific effect (unobserved heterogeneity) and ε_{it} is the disturbance term with $E[\varepsilon_{it}]=0$ and $Var[\varepsilon_{it}]=\sigma_\varepsilon^2$ (idiosyncratic error). The term $\nu_i+\varepsilon_{it}$ is the composite error.

When the dependent variable is income inequality for the population as a whole, we consider population ageing, work access, unemployment, and inactivity as time-variant explanatory variables, while when the dependent variable is income inequality for normally working people, we consider only population ageing as time-variant explanatory variable.

We then analyse the role of welfare state, religion, and family structure on income inequality. These are explanatory variables, represented by dummies in the static panel data model. Our analysis takes on the following form:

$$y_{it}=\beta' x_{it}+\eta' d_{\lambda i}+\nu_i+\varepsilon_{it},$$

where η are coefficients and $d_{\lambda i}$ is a vector of dummy variables with λ denoting categories ($\lambda=2,...,m$). If a qualitative variable has m categories, we introduce $m-1$ dummy variables (categories). Category d_{1i} is referred to as the base category. Comparisons are made with that category (Gujarati 2003).

This static model is characterised by one source of persistence over time due to the presence of unobserved regional-specific effects. The presented static methods of panel estimation are pooled ordinary least squares (OLS) and FEs (random effects (REs) in appendix). To evaluate which technique is optimal, it is necessary to consider the relationship between the regional-specific effects and the regressors, among others. Both FEs and REs estimators are based on the strict exogeneity assumption. Hence the vector of the explanatory variables (x_{it} and z_i) is strictly exogenous. The usual diagnostic tests are also presented. Hausman's (1978) chi-squared statistic tests whether the REs estimator is an appropriate alternative to the FEs estimator. Breusch and Pagan's (1980) Lagrange multiplier (LM) statistic tests the REs model against OLS model. LM test is a test for regional effects. Large values of LM statistic favour the REs model.

In the static models, we assume that the regression disturbances are homoskedastic with the same variance across time and regions. However, heteroskedasticity potentially causes problems for inferences based on least squares. Assuming homoskedastic disturbances in the FEs model, for example, might be a restrictive assumption for panels (Baltagi 2005). Thus when heteroskedasticity is present, the consistent estimates are not efficient. If every ε_{it} has a different variance, the robust estimation of the covariance matrix is presented following the White estimator for unspecified heteroskedasticity (White 1980).

There are a variety of different techniques that can be used to estimate a dynamic model of the form:

$$y_{it} = \delta y_{i,t-1} + \beta' x_{it} + \zeta' x_{i,t-1} + \gamma' z_i + \nu_i + \varepsilon_{it} \qquad (3)$$

with i denoting regions ($i = 1,...,N$) and t time ($t = 2,...,6$)[3]. y_{it} is income inequality, $y_{i,t-1}$ is the first lagged income inequality, x_{it} is a vector of explanatory variables, $x_{i,t-1}$ is a vector of first lagged explanatory variables, z_i is a vector of time-invariant explanatory variables (urbanisation and latitude), δ, β, ζ and γ are coefficients, ν_i are the random effects (unobserved regional-specific effects) that are independent and identically distributed over the panels and ε_{it} is the disturbance term with $E[\varepsilon_{it}] = 0$ and $Var[\varepsilon_{it}] = \sigma_\varepsilon^2$ (idiosyncratic error). It is assumed that the ν_i and the ε_{it} are independent for each i over all t.

This dynamic model is characterised by two sources of persistence over time: autocorrelation due to the presence of a lagged dependent variable among the regressors and unobserved regional-specific effects (Baltagi 2005). Pooled OLS, FEs and REs estimators are now biased and inconsistent, because the econometric model contains a lagged endogenous variable (Baltagi 2005).

The dynamic panel structure of our data is exploited by a generalised method of moments (GMM) estimation suggested by Arellano and Bond (1991) (Arellano-Bond estimation).

[3] $t = 2$ denotes 1996, …, $t = 6$ denotes 2000.

The main idea behind GMM estimation is to establish population moment conditions and then use sample analogs of these moment conditions to compute parameter estimates (Greene 2003; Wooldridge 2003; Baltagi 2005). Arellano and Bond first transform the model to eliminate the regional-specific effect (ν_i). The observed time-invariant explanatory variables (z_i) are eliminated as well. The first-differencing transformation is:

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2}) + \beta'(x_{it} - x_{i,t-1}) + \zeta'(x_{i,t-1} - x_{i,t-2}) + (\varepsilon_{it} - \varepsilon_{i,t-1}), \qquad (4)$$

where all variables are expressed as deviations from period means. Models in first differences usually face the problems arising from the non-stationarity of the data. The correlation between the explanatory variables and the error is handled by instrument variables (IVs). In Arellano-Bond estimations, the predetermined and endogenous variables in first differences are instrumented with suitable lags of their own levels, while the strictly exogenous regressors can enter the instrument matrix in first differences. For instance, for 1997 $(t=3)$, $y_{i,1}$ is an instrument for $(y_{i,2} - y_{i,1})$ and not correlated with $(\varepsilon_{i3} - \varepsilon_{i2})$ as long as the ε_{it} themselves are not serially correlated; for 1998 $(t=4)$, $y_{i,1}$ and $y_{i,2}$ are instruments for $(y_{i,3} - y_{i,2})$, and so on. This procedure is more efficient than the Anderson and Hsiao (1981; 1982) two stage least squares estimator which does not make use of all of the available moment conditions (Ahn and Schmidt 1995).

In our model, we assume that the explanatory variables might be:

a. strictly exogenous, if $E[x_{it}\varepsilon_{is}] = 0$ for all t and s,

b. predetermined, if $E[x_{it}\varepsilon_{is}] \neq 0$ for $s < t$, but $E[x_{it}\varepsilon_{is}] = 0$ for all $s \geq t$, and

c. endogenous, if $E[x_{it}\varepsilon_{is}] \neq 0$ for $s \leq t$, but $E[x_{it}\varepsilon_{is}] = 0$ for all $s > t$;

except for population ageing which is definitely a strictly exogenous variable.

The GMM methodology is based on a set of diagnostics. First of all, it assumes that there is no second-order autocorrelation in the first-differenced idiosyncratic errors.[4] Additionally, Arellano and Bond (1991) developed Sargan's test (1958) of over-identifying restrictions. The Sargan test has an asymptotic chi-squared distribution in the case of homoskedastic error term only. Both the homoskedastic one-step and the robust one-step GMM estimators are presented.

The dynamic model is also used in order to obtain short-run and long-run parameters. The short-run effect of an independent variable is the first year effect of a change in this variable, whereas the long-run effect is the effect obtained after full adjustment of income inequality. The short-run effect of the variable x is β and its long-run effect is $\beta + \zeta / 1 - \delta$. Long-run standard errors are calculating using the Delta method (Greene 2003).

Broadly speaking, the advantage of dynamic over static models is that the former correct the inconsistentcy introduced by lagged endogenous variables and, also, permits a certain degree of endogeneity in the regressors.

4. Data and Variables

As in other recent studies dealing with human capital variables across European regions (Rodríguez-Pose and Vilalta-Bufí 2005; Ezcurra *et al.* 2008), the data used to estimate the econometric models come from the ECHP data survey conducted by the EU during the period 1994-2001 (wave2-wave8) and the Eurostat's Regio dataset. In the surveys individuals were interviewed about their socioeconomic status. Data stemming from the ECHP can be aggregated regionally at NUTS I or II level for the EU15. Unfortunately there are no data available for the Netherlands. Finnish regions also had to be dropped from the sample because of discrepancies between the regional division included in the ECHP and those in the Regio databank. The resulting database includes 102 NUTS I or II regions from

[4] The consistency of the GMM estimator relies upon the fact that $E[\Delta\varepsilon_{it}\Delta\varepsilon_{i,t-2}] = 0$ (Arellano and Bond 1991: 282).

13 countries in the EU.[5] On average 116,574 individuals were surveyed, with a maximum of 124,759 in 1997 and a minimum of 105,079 in 2001.

The variable '*Total net personal income (detailed, NC, total year prior to the survey)*' from the ECHP is used as the main source for the average income and the income inequality for the population as a whole. This variable is regionalised. Income is collected not only for each individual in the household so as to measure income per capita (IMN) and income inequality for the population as a whole $(IGE1)$, but also for each normally working (15+ hours/week) individual[6] in the household in order to measure income per capita (NMN) and income inequality for normally working people $(NGE1)$. Income per capita is transformed for the same level of prices using the harmonised indices for consumer prices and then is divided by 1,000. Income inequality is calculated using the generalised Theil entropy index (Theil 1967). This index considers a region's population of individuals $i \in \{1,2,..., N\}$ where each person is associated with a unique value of the measured income. The total net personal income is the sum of wages and salaries, income from self employment or farming, pensions, unemployment, and redundancy benefits or any other social benefits or grants, and private income. Income inequality within a region is defined as $IGE1 = \sum_{i=1}^{N} y_i \log(Ny_i)$, where y_i is income share that is individual i's total income as a proportion of total income for the entire regional population. This index varies from 0 for perfect equality to $\log N$ for perfect inequality.

The average and inequality in education are calculated using the microeconomic variable '*Highest level of general or higher education completed*' which is also extracted from the ECHP data survey. Individuals are classified into three educational categories: recognised third level education completed, second stage of secondary education level completed, and less than second stage of secondary education level completed. These categories, which are

---

[5] NUTS I data for Austria, Belgium, Denmark, France, Greece, Ireland, Italy, Luxembourg, Spain, Sweden. NUTS II data for Germany, Portugal, and the UK.

[6] It is extracted from the variable '*Main activity status-Self defined (regrouped)*'.

mutually exclusive, allow for international comparisons, because they are defined by the International Standard Classification of Education.

The average education level completed was first has been defined by Psacharopoulos and Arriagada (1986) and Ram (1990). It corresponds to the educational attainment (or educational achievement) and is given by the index $EMN = \sum_{j=1}^{3} L_j S_j$, where L_j is the proportion of the respondents who belong in the j^{th} category and S_j denotes an assessment of each category. At the risk of some oversimplification, we assume $S_1 = 2$ for recognised third level education completed, $S_2 = 1$ for second stage of secondary education level completed, and $S_3 = 0$ for less than second stage of secondary education level completed.

Following the work of Thomas, Wang et al. (2001), we calculate the inequalities in educational attainment using an education Theil index $(EGE1)$. This is defined as $EGE1 = \sum_{i=1}^{N} z_i \log(Nz_i)$, where z_i is human capital share, that is, individual i's higher education level completed as a proportion of total human capital for the entire regional population. As in the case for income inequality the index has a minimum value of 0 when the entire population is concentrated in a single educational category, and a maximum of $\log N$.

As a way of controlling for the impact of additional factors, we also examine the impact of additional quantitative time-variant variables on income inequality: the average age of people (AGE), the percentage of normally working (15+ hours/week) respondents $(LFSTOCK)$, the percentage of unemployed respondents $(UNEM)$, and the percentage of inactive respondents $(INACTIVE)$ within a region. The source of these variables is again the ECHP dataset. Other controls include the economic activity rate of the population $(ECACRA)$ and female activity rate $(ECACRF)$ from the Eurostat's Regio dataset. These are also time-variant variables. The urbanisation ratio of a region $(URBANDPA)$ is constructed as the percentage of respondents who live in a densely populated area. Data for this variable are only available for 2000 and 2001, and not for all countries. We assume that the urbanisation ratio from 1995 to 2001 remains constant. This variable, therefore, introduces observed time-invariant effects. A second time-invariant variable is latitude (LAT).

The transformed dataset with means, standard deviation, and minimum and maximum value for each of the variable is reported in Table 1.[7] The descriptive statistics show that the dataset is unbalanced, which is amenable to estimation methods that manage potential heterogeneity bias. Table 1 also depicts that income inequality both for the population as a whole and for normally working people has decreased slightly between 1995 and 2000. Educational inequalities followed a similar declining trend over the period of analysis.

Table 1: Summary Statistics

Variable	Dedinition	Year	Source	Obs	Mean or %	Std. Dev.	Min	Max
IGE1	Income inequality for the population as a whole (Theil index)	1995	ECHP	94	0.42	0.16	0.18	0.83
		2000		102	0.36	0.14	0.11	0.74
		1995-00		604	0.38	0.15	0.11	0.83
IMN	Income per capita for the population as a whole (/1000)	1995	ECHP	94	9.76	3.54	3.40	18.93
		2000		102	12.81	4.55	4.05	21.14
		1995-00		604	11.30	3.96	3.40	21.14
NGE1	Income inequality for normally working people (Theil index)	1995	ECHP	94	0.24	0.08	0.13	0.49
		2000		102	0.21	0.07	0.06	0.41
		1995-00		604	14.83	4.56	4.94	29.35
NMN	Income per capita for normally working people (/1000)	1995	ECHP	94	13.19	4.32	4.94	28.42
		2000		102	16.62	5.21	5.80	29.31
		1995-00		604	0.22	0.08	0.06	0.49
EMN	Average education level completed	1995	ECHP	94	0.66	0.24	0.12	1.17
		2000		102	0.80	0.27	0.19	1.23
		1995-00		596	0.75	0.28	0.12	1.34
EGE1	Inequality in education level completed (Theil index)	1995	ECHP	94	0.90	0.45	0.21	2.38
		2000		102	0.72	0.39	0.17	2.02
		1995-00		596	0.79	0.44	0.17	2.42
AGE	Average age of respondents	1995	ECHP	94	45.19	2.29	39.76	51.39
		2000		102	45.96	1.86	42.32	51.35
		1995-00		596	45.40	1.95	39.76	51.61
LFSTOCK	Percentage of normally working (15+ hours/week) respondents (self-defined)	1995	ECHP	94	52.27	7.24	33.59	67.78
		2000		102	53.79	6.97	36.56	67.55
		1995-00		596	52.78	7.39	31.20	72.88
ECACRA	Percentage of economic acrivity rate of total population	1995	Eurostat	65	54.90	7.47	42.00	74.80
		2000		94	57.89	6.61	42.90	74.50
		1995-00		525	57.10	6.85	41.50	74.80

[7] Appendix A.1 shows the descriptive statistics of the ECHP quantitative and qualitative variables.

UNEM	Percentage of unemployed respondents (self-defined)	1995	ECHP	94	5.80	3.29	0.00	16.54
		2000		102	4.46	2.80	0.59	14.85
		1995-00		596	5.28	3.17	0.00	16.54
INACTIVE	Percentage of inactive respondents (self-defined)	1995	ECHP	94	41.92	5.96	29.21	55.49
		2000		102	41.74	5.86	29.53	55.42
		1995-00		596	41.94	6.05	27.12	56.72
ECACRF	Percentage of female's economic activity rate	1995	Eurostat	65	44.78	10.82	24.00	72.20
		2000		94	49.15	9.14	26.70	72.90
		1995-00		525	47.79	9.52	23.40	72.90

Source: ECHP dataset and Eurostat's Regio dataset

The qualitative explanatory variables (time-invariant) classify regions into categories that are hypothesised to have some underlying similarity concerning welfare regimes, religion, and family structure.

- Welfare regime: Following the work of Esping-Andersen (1990), Ferrera (1996), and Berthoud and Iacovou (2004), we use four welfare state categories: social-democratic (Sweden, Denmark), liberal (UK, Ireland), corporatist or conservatism (Luxembourg, Belgium, France, Germany, Austria) and residual or 'Southern' (Portugal, Spain, Italy, Greece). The hypothesis is that a country's welfare policy has an important effect on income redistribution and thus on income inequalities. The above classification assumes that a country belongs to only one welfare state regime. In reality, there is no single pure case because the Scandinavian countries, for instance, may be predominantly social democratic, but they are not free of liberal elements (Esping-Andersen 1990: 28).

- Religion: European regions' religious affiliation is classified into four categories[8]: mainly Protestant (Sweden, Denmark, northern Germany, Scotland), mainly Catholic (France, Ireland, Luxembourg, Portugal, Spain, Italy, Austria, parts of southern Germany, Belgium), mainly Anglican (England) and mainly Orthodox (Greece). It is hypothesised that regions with the same religion have close social links so at to have similar income inequality levels within-groups of religion, but different inequality between-groups.

[8] Sources: http://www.cia.gov/cia/publications/factbook;
http://commons.wikimidia.org/wiki/Image:Europe_religion_map_de.png;
http://csi-int.org/world_map_europa_religion.php

- Family structure: Following the work of Berthoud and Iacovou (2004), we use three groups of countries in the study of living arrangement: Nordic (Sweden, Denmark), North/Central (UK, Belgium, Luxembourg, France, Germany, Austria) and Southern/Catholic (Ireland, Portugal, Spain, Italy, Greece). The hypothesis is that a country's family structure plays a significant role in income inequality.

5. Regression Results

The empirical analysis exploits the panel structure of the dataset, for the 102 EU regions included in the analysis over the period 1995-2000, using pooled OLS, FEs, and REs estimation of the static models and by GMM estimation of the dynamic models taking into account the unobserved regional-specific effects. We first report the static regression models, followed by the dynamic ones.

5.1 Estimations of the Static Model

In all the regressions of income inequality for the population as a whole, the p-values of Breusch and Pagan's Lagrange multiplier test strongly reject the validity of the pooled OLS models, and the p-values of Hausman's test reject the GLS estimator as an appropriate alternative to the FEs estimator. According to the specification tests, the FEs models are the most appropriate. There is also not much difference between the significance of the homoskedasticity and the heteroskedasticity consistent covariance matrix estimator. Thus, the determinants of income inequality are not sensitive to the model specification of the error term. Tables 2 and 3 display the FEs and OLS regression results, respectively, while the REs results are displayed in Appendix A.2.

Table 2: FEs: Dependent variable is income inequality for the population as a whole (IGE1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IMN	-0.0001 (0.0011) *(0.0013)*	0.0016 (0.0014) *(0.0016)*	0.0026 (0.0014)* *(0.0017)*	0.0033 (0.0014)** *(0.0017)**	0.0029 (0.0016)* *(0.0017)**	0.0046 (0.0016)*** *(0.0017)****	0.0039 (0.0016)** *(0.0018)***	0.0110 (0.0025)*** *(0.0027)****	0.0111 (0.0019)*** *(0.0021)****
EMN		0.0396 (0.0305) *(0.0316)*	0.0394 (0.0303) *(0.0318)*	0.0466 (0.0301) *(0.0309)*	0.0018 (0.0306) *(0.0293)*	0.0136 (0.0298) *(0.0276)*	0.0101 (0.0305) *(0.0285)*	0.0222 (0.0396) *(0.0415)*	0.0103 (0.0314) *(0.0277)*
EGE1		0.0723 (0.0230)*** *(0.0231)****	0.0732 (0.0229)*** *(0.0232)****	0.0685 (0.0227)*** *(0.0223)****	0.0313 (0.0224) *(0.0197)*	0.0330 (0.0218) *(0.0184)**	0.0361 (0.0222) *(0.0188)**	0.0831 (0.0302)*** *(0.0374)***	0.0424 (0.0211)** *(0.0163)****
AGE			-0.0057 (0.0022)** *(0.0024)***	-0.0059 (0.0022)*** *(0.0026)***	-0.0082 (0.0022)*** *(0.0025)****	-0.0053 (0.0022)** *(0.0025)***	-0.0073 (0.0022)*** *(0.0024)****	-0.0073 (0.0027)*** *(0.0026)****	-0.0030 (0.0022) *(0.0023)*
LFSTOCK				-0.2765 (0.0837)*** *(0.0981)****					
ECACRA					-0.0089 (0.0014)*** *(0.0016)****				
UNEM						0.5541 (0.1404)*** *(0.1515)****		0.4594 (0.2069)** *(0.2305)***	0.3783 (0.1378)*** *(0.1511)***
INACTIVE							0.0084 (0.0933) *(0.1080)*		
ECACRF						-0.0068 (0.0012)*** *(0.0013)****	-0.0079 (0.0012)*** *(0.0013)****	-0.0020 (0.0017) *(0.0017)*	-0.0042 (0.0012)*** *(0.0014)****
YR96*UR BANDPAV								-0.0290 (0.0148)* *(0.0151)**	
YR97*UR BANDPAV								-0.0453 (0.0150)*** *(0.0136)****	
YR98*UR BANDPAV								-0.0136 (0.0163) *(0.0147)*	
YR99*UR BANDPAV								-0.0374 (0.0174)** *(0.0170)***	
YR00*UR BANDPAV								-0.0743 (0.0184)*** *(0.0171)****	
YR96*LAT									-0.0002 (0.0001) *(0.0001)*
YR97*LAT									-0.0005 (0.0001)*** *(0.0001)****
YR98*LAT									-0.0003 (0.0001)*** *(0.0001)****
YR99*LAT									-0.0006 (0.0001)*** *(0.0001)****
YR00*LAT									-0.0009 (0.0001)*** *(0.0002)****
CONSTANT	0.3821 (0.0121)*** *(0.0151)****	0.2787 (0.0382)*** *(0.0396)****	0.5255 (0.1022)*** *(0.1072)****	0.6732 (0.1106)*** *(0.1220)****	1.2128 (0.1333)*** *(0.1438)****	0.8348 (0.1195)*** *(0.1213)****	1.0108 (0.1153)*** *(0.1182)****	0.6300 (0.1611)*** *(0.1640)****	0.5593 (0.1288)*** *(0.1337)****
ADJ R-SQ	0.0000	0.0313	0.0445	0.0654	0.1343	0.1743	0.1432	0.2704	0.2601
OBS.	604	596	596	596	513	513	513	299	513
LM TEST (p-value)	916.46 (0.0000)	715.20 (0.0000)	645.03 (0.0000)	634.09 (0.0000)	715.68 (0.0000)	676.43 (0.0000)	630.60 (0.0000)	322.72 (0.0000)	694.28 (0.0000)
HAUSMAN TEST (p-value)	71.46 (0.0000)	289.07 (0.0000)	35.86 (0.0000)	87.27 (0.0000)	46.71 (0.0000)	54.24 (0.0000)	73.32 (0.0000)		

NOTES: (*), (**), and (***) indicates significance at the 10%, 5% and 1% level, respectively. *(*), (**),* and *(***)* denotes the significance of the White (1980) estimator (robust standard errors). LM TEST is the Lagrange multiplier test for the random effects model based on the OLS residuals (Breusch and Pagan 1980). HAUSMAN TEST is the Hausman (1978) test for fixed or random effects.

Table 3: OLS: Dependent variable is income inequality for the population as a whole (IGE1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
IMN	-0.0253 (0.0012)*** _(0.0014)_***	-0.0140 (0.0018)*** _(0.0021)_***	-0.0129 (0.0018)*** _(0.0020)_***	-0.0017 (0.0016) _(0.0018)_	-0.0065 (0.0015)*** _(0.0016)_***	-0.0033 (0.0015)** _(0.0017)_*	-0.0043 (0.0014)*** _(0.0014)_***	-0.0076 (0.0024)*** _(0.0028)_***	0.0020 (0.0014) _(0.0015)_	0.0072 (0.0018)*** _(0.0021)_***	0.0005 (0.0015) _(0.0017)_	0.0084 (0.0018)*** _(0.0020)_***
EMN		-0.2817 (0.0355)*** _(0.0304)_***	-0.2906 (0.0347)*** _(0.0285)_***	-0.0800 (0.0312)** _(0.0263)_***	0.0097 (0.0331) _(0.0315)_	0.0498 (0.0298)* _(0.0288)_*	0.0652 (0.0295)** _(0.0286)_**	0.0710 (0.0375)* _(0.0381)_*	0.0354 (0.0263) _(0.0237)_	0.0309 (0.0338) _(0.0358)_	0.1064 (0.0340)*** _(0.0372)_***	0.0381 (0.0283) _(0.0296)_
EGE1		-0.0556 (0.0210)*** _(0.0199)_***	-0.0412 (0.0206)*** _(0.0179)_***	0.0719 (0.0183)*** _(0.0167)_***	0.0961 (0.0189)*** _(0.0181)_***	0.1074 (0.0175)*** _(0.0166)_***	0.0926 (0.0166)*** _(0.0152)_***	0.0700 (0.0217)*** _(0.0185)_***	0.0582 (0.0160)*** _(0.0141)_***	0.0887 (0.0187)*** _(0.0192)_***	0.1483 (0.0188)*** _(0.0198)_***	0.0935 (0.0164)*** _(0.0173)_***
AGE			-0.0130 (0.0023)*** _(0.0024)_***	-0.0170 (0.0019)*** _(0.0019)_***	-0.0138 (0.0019)*** _(0.0018)_***	-0.0078 (0.0018)*** _(0.0018)_***	-0.0134 (0.0020)*** _(0.0022)_***	-0.0041 (0.0023)* _(0.0022)_*	-0.0076 (0.0016)*** _(0.0015)_***	-0.0082 (0.0017)*** _(0.0017)_***	-0.0113 (0.0018)*** _(0.0017)_***	-0.0077 (0.0017)*** _(0.0016)_***
LFSTOCK				-1.1632 (0.0693)*** _(0.0676)_***								
ECACRA					-0.0134 (0.0008)*** _(0.0007)_***							
UNEM						0.2519 (0.1304)* _(0.1352)_*		0.8557 (0.2080)*** _(0.1794)_***	0.2375 (0.1150)** _(0.1190)_**	0.4602 (0.1410)*** _(0.1380)_***	0.3112 (0.1384)** _(0.1431)_**	0.5367 (0.1264)*** _(0.1362)_***
INACTIVE							0.4937 (0.1052)*** _(0.1141)_***					
ECACRF						-0.0116 (0.0006)*** _(0.0005)_***	-0.0096 (0.0007)*** _(0.0008)_***	-0.0083 (0.0010)*** _(0.0009)_***	-0.0084 (0.0006)*** _(0.0006)_***	-0.0085 (0.0007)*** _(0.0007)_***	-0.0104 (0.0006)*** _(0.0006)_***	-0.0082 (0.0007)*** _(0.0007)_***
URBANDP AV (fixed)								-0.0736 (0.0215)*** _(0.0211)_***				
LAT (fixed)									-0.0102 (0.0008)*** _(0.0009)_***			
DWSLIB										0.0356 (0.0185)* _(0.0166)_**		
DWSCORP										0.0374 (0.0169)** _(0.0154)_**		
DWSRES										0.1814 (0.0261)*** _(0.0291)_***		
DRLCATH											0.0408 (0.0109)*** _(0.0112)_***	
DRLORTH											0.1584 (0.0196)*** _(0.0179)_***	
DRLANGL											-0.0104 (0.0122) _(0.0127)_	
DFNORD												-0.0402 (0.0163)** _(0.0145)_***
DFSC												0.1566 (0.0147)*** _(0.0179)_***
ADJ R-SQ	0.4233	0.4890	0.5144	0.6709	0.7139	0.7674	0.7755	0.7672	0.8192	0.8022	0.7978	0.8097
OBS.	604	596	596	596	513	513	513	299	513	513	513	513

In Regression 1 of Table 2, the impact of income per capita (IMN) on income inequality $(IGE1)$ is analysed. This equation is unconditioned by any other effects. The relationship between income per capita and inequality is negative, but it is not statistically significant. The adjusted R-squared shows that income per capita does not explain any variation in income inequality in the sample. In terms of goodness-of-fit, it is likely to indicate a poor unconditioned model. In the FEs conditional regressions (Regressions 3–9) income per capita is positively correlated with income inequality. The higher the income per capita, the higher the inequality within a region. A plausible explanation for this is that regional economic development seems to increase the occupational choices and the earning opportunities of the rich (Lydall 1979). In all the regressions, however, the coefficients on income per capita are very low. For instance, Regression 4 shows that an increase of one per cent in income per capita is associated with, on average, about 0.0033 per cent more income inequality, as measured by the Theil index. The findings also indicate that the effect of income per capita on inequality is robust as it is not sensitive to the model specification.

The next step in the analysis is the introduction of human capital distribution, as measured by educational attainment (EMN) and educational inequality $(EGE1)$. Regressions 2–9 indicate that regional educational achievement probably has no influence on the resulting income distribution, as the coefficients on educational attainment are not statistically significant. Thus, it is not clear whether higher educational attainment increases the occupational choices and the earning opportunities of the population as a whole so as to make societies more egalitarian. Additionally, it is not clear whether education improves the overall chances for individuals, because it reflects abilities, choices, and preferences (Hannum and Buchmann 2005). The insignificant correlation between income inequality and educational attainment also says nothing about the balance between the 'wage compression' effect and the 'composition' effect (Knight and Sabot 1983). Education does not seem to expose all economic agents to a common shift factor that affects each individual's income. The empirical results, nonetheless, show that a highly unequal distribution of education level completed is associated with higher income inequality. This relationship is robust and statistically significant (Regressions 2–4 and 6–9). A larger share of highly-educated workers within a region may signal to employers that those with less education have less ability, which may also lead to a larger wage differential between highly-educated and less-educated workers and thus to higher income inequality. An increase in the levels of education of the highly-educated tends to increase income inequality as the imperfect competition for positions requiring advanced educational credentials raises the wages of educated people even more. Another explanation is that the demand for unskilled labour grows at a slower rate than the demand for skilled labour. Hence, the positive relationship seems to indicate the responsiveness of the EU labour market to differences in qualifications and skills.

The remaining regressions include the control variables described earlier. Regressions 3–9 test for the influence of the average age of respondents (AGE). The fact that age matters for income inequality is hardly surprising, as regions with a younger population also tend to have a lower rate of participation in the labour force and young people in work earn less in a labour market that rewards seniority, increasing the inequality levels within a society (Higgins and Williamson 1999).

In order to capture the economic activity characteristics of the regions, the percentage of normally working respondents $(LFSTOCK)$, and the economic activity rate of the total population $(ECACRA)$ are included in Regressions 4 and 5, respectively. As expected, both variables are negatively associated with income inequality and are statistically significant. The higher the level of economic activity of a region, the lower the income inequality, reflecting that one of the main factors determining income inequality is access to work.

This point is further confirmed by the introduction of unemployment $(UNEM)$ and inactivity levels $(INACTIVE)$ within a region, as well as the participation in labour market by sex $(ECACRF)$ in Regressions 6 and 7, respectively. The results indicate that high unemployment is associated with higher income inequality. Increases in unemployment aggravate the relative position of low-income groups, because marginal workers with relatively low skills are at the bottom of the income distribution and their jobs are at greater risk during an economic downturn (Mocan 1999). Additionally, unemployment insurance, welfare benefits, and other forms of income support are not enough to offset the loss in income due to transitory unemployment. European labour conditions, such as the degree of centralisation in wage bargaining, the existence of a minimum wage, the differences among countries with regard to recruitment and dismissal legislation, and the differences among the European countries concerning unemployment benefits, job-creation policies, and vocational training programmes (Ayala, Martinez et al. 2002), represent an important factor in determining the differences observed in income inequality across European regions. The coefficients on the female economic activity rate in all regressions are negative and significant. The impact of the increase in women's access to work has been to lessen the trend toward greater income inequality caused by aspects of social change during the period of analysis (Ryscavage, Green et al. 1992).

In Regressions 8 and 9 we introduce a year dummy variable for urbanisation $(URBANDPAV)$ and latitude (LAT), respectively, in order to see whether the effects of urbanisation and latitude on income inequality have changed over the period 1995–2000. The effect of urbanisation and latitude is lower in 2000 (Regressions 8 and 9, respectively). The OLS (Table 3) and REs (Appendix A.2) results show the negative correlation between urbanisation and inequality. Considering Kuznets' assumption that urbanisation is a measure of economic development, the negative relationship highlights the fact that European societies are located in the declining segment of the Kuznets curve. However, this disproves Estudillo's (1997) hypothesis that the heterogeneity of urban areas enhances, rather than lowers, inequality. Highly-urbanised regions seem not only to be more economically prosperous — the correlation between income per capita and urbanisation is positive (0.46) — but also less unequal. Notably, the OLS and REs results show that the latitude variable has the 'right' sign and is significant. The northern regions exhibit the lowest income inequality levels. On the one hand, an analysis involving latitude is likely to highlight the EU north-south divide in terms of income inequality. On the other hand, bearing in mind that latitude is a good proxy for the effect of a region's climate on its level of productive efficiency, it is likely to account for a large proportion of the differences in regional inequality levels.

Finally, the impact of the qualitative explanatory variables on income inequality (Regressions 10–12) is presented in Table 3 (OLS results) and Appendix A.2 (REs results).[9] The FEs estimator is not provided because there is no within-group variation in the dummy variables.

In Regression 10 of Table 3, the omitted category is social-democratic welfare states. The regression results show that all welfare regimes are important determinants of income inequality. Social-democratic welfare states, which in theory promote a higher standard of equality, indeed have lower levels of income inequality than corporatist welfare states, in which private insurance and occupational benefits play a truly marginal role and corporatism displaces the market as a provider of welfare (Esping-Andersen 1990). In addition, social-democratic welfare states are more egalitarian than corporatist ones because, in the former, the welfare state minimises dependence on the family and allows women greater freedom to choose work rather than to stay at home, while in the latter state intervention is more modest and comes into effect mainly when the family's capacity to service its members becomes exhausted (Esping-Andersen 1990). Corporatist welfare states in turn have higher levels of income inequality than liberal welfare states. However, both regimes are more egalitarian than 'Southern' (or 'residual') ones.

Regression 11 introduces religion as an explanatory variable. Mainly Protestant regions, which are the base category, have a lower level of income inequality than Catholic ones. Orthodox regions have the most inegalitarian societies. Finally, it is interesting to note that all categories of family structure and living arrangements affect income inequality significantly (Regression 12). Regions with a Nordic family structure have the most egalitarian societies and Southern/Catholic regions have the highest inequality.

Considering the standardised coefficients for the above regressions (Appendix A.4),[10] women's access to work explains the largest variation in income inequality. The impact of both approaches to economic activity (work access of total population) on income inequality is high. In contrast, population ageing, unemployment, and urbanisation explain only a relatively small part of the total variation in income inequality.

---

[9] See Appendix A.3 for dummy variable definition.

[10] The standardised coefficient is the standard deviation change in the dependent variable caused by one standard deviation change in each explanatory variable.

The static regression results of income inequality for normally working people ($NGE1$) are similar to the regression results of income inequality for the population as a whole.[11] More specifically, income per capita is positively associated with income inequality. This relationship is statistically significant and robust. This behaviour rejects the declining segment of the Kuznets' curve. Once more, the impact of educational achievement on income inequality is not clear, as the coefficients on educational attainment are not statistically significant; while the results are consistent with the current belief that educational inequality is positively correlated with income inequality. The latter relationship is also robust. The influence of population ageing is not statistically significant. The results display the negative impact of the female participation in labour force on inequalities. The impact of urbanisation and latitude on inequalities is stronger in 2000 than in 1995. Finally, income inequality for normally working people is lower in social-democratic welfare states, in mainly Protestant areas and in regions with Nordic family structures; while, it is higher in 'Southern' welfare states, in Orthodox areas and in Southern/Catholic regions.

5.2 Estimations of the Dynamic Model

Table 4 presents the long-run results for the dynamic income inequality for the population as a whole (Arellano-Bond estimator). The first column of each model specification assumes that the explanatory variables are strictly exogenous. The last two columns show the GMM results for the same model specification regarding whether the explanatory variables are predetermined (column b) or endogenous (column c). The short-run parameters and the specification tests (the tests regarding serial correlation and the Sargan tests) are presented in Appendix A.8.

[11] The FEs, OLS and REs results of income inequality for normally working people are reported in Appendices A.5, A.6 and A.7, respectively.

Table 4: Long Run GMM: Dependent variable is income inequality for the population as a whole (IGE1)

	REGRESSION (1)			REGRESSION (2)			REGRESSION (3)			REGRESSION (4)		
	(a) x_{it} strictly exogenous	(b) x_{it} predetermine d	(c) x_{it} endogenous	(a) x_{it} strictly exogenous	(b) x_{it} predetermine d	(c) x_{it} endogenous	(a) x_{it} strictly exogenous	(b) x_{it} predetermine d	(c) x_{it} endogenous	(a) x_{it} strictly exogenous	(b) x_{it} predetermine d	(c) x_{it} endogenous
IMN	0.0331 (0.0137)** *(0.0143)***	0.0266 (0.0200) *(0.0189)*	0.0377 (0.0136)*** *(0.0151)***	0.0654 (0.0890) *(0.1038)*	0.0314 (0.0134)** *(0.0183)**	0.0239 (0.0096)** *(0.0126)**	0.0749 (0.1272) *(0.1489)*	0.0344 (0.0128)*** *(0.0180)**	0.0248 (0.0093)*** *(0.0121)***	0.5001 (9.4502) *(10.4434)*	0.0372 (0.0121)*** *(0.0163)***	0.0211 (0.0102)** *(0.0108)**
EMN				-0.3781 (0.9759) *(1.1395)*	0.0577 (0.1948) *(0.2269)*	0.3018 (0.1555)* *(0.1692)**	-0.5019 (1.4055) *(1.6554)*	0.0399 (0.1813) *(0.2137)*	0.2899 (0.1518)* *(0.1641)**	-5.8878 (116.8038) *(129.5313)*	0.0378 (0.1533) *(0.1723)*	0.3042 (0.1474)** *(0.1593)**
EGE1				-0.1317 (0.5449) *(0.5273)*	0.0912 (0.1180) *(0.0819)*	0.1705 (0.1015)* *(0.0861)***	-0.2153 (0.8028) *(0.8323)*	0.0957 (0.1102) *(0.0831)*	0.1660 (0.0997)* *(0.0874)**	-2.4249 (49.2962) *(54.5765)*	0.1218 (0.0920) *(0.0742)*	0.1963 (0.0944)** *(0.0934)***
AGE							0.1000 (0.2066) *(0.2464)*	0.0121 (0.0144) *(0.0169)*	0.0127 (0.0105) *(0.0138)*	0.9354 (18.2349) *(20.2553)*	0.0085 (0.0126) *(0.0150)*	0.0119 (0.0101) *(0.0126)*
LFSTOCK										36.9702 (726.0782) *(800.2190)*	0.0195 (0.6375) *(0.7831)*	-0.1129 (0.7628) *(0.8953)*
ECACRA												
UNEM												
INACTIVE												
ECACRF												
OBS.	400			392			392			392		
	REGRESSION (5)			REGRESSION (6)			REGRESSION (7)					
IMN	0.0151 (0.0124) *(0.0133)*	0.0133 (0.0101) *(0.0099)*	0.0086 (0.0135) *(0.0157)*	0.0144 (0.0187) *(0.0200)*	0.0140 (0.0080)* *(0.0070)***	0.0097 (0.0103) *(0.0103)*	0.0104 (0.0179) *(0.0201)*	0.0173 (0.0126) *(0.0131)*	0.0118 (0.0115) *(0.0124)*			
EMN	-0.1077 (0.1761) *(0.2117)*	-0.1321 (0.1340) *(0.1844)*	-0.2919 (0.2186) *(0.2773)*	-0.1380 (0.2748) *(0.3289)*	-0.0312 (0.1025) *(0.1304)*	-0.0252 (0.1437) *(0.1815)*	-0.1475 (0.2644) *(0.3172)*	-0.1382 (0.1610) *(0.1864)*	-0.2431 (0.1802) *(0.2386)*			
EGE1	-0.0531 (0.1159) *(0.1206)*	0.0199 (0.0831) *(0.0964)*	-0.1783 (0.1534) *(0.1612)*	-0.0581 (0.1769) *(0.1908)*	0.0447 (0.0649) *(0.0750)*	-0.0261 (0.1000) *(0.1073)*	-0.0698 (0.1718) *(0.1833)*	0.0031 (0.0997) *(0.1060)*	-0.1144 (0.1225) *(0.1661)*			
AGE	0.0186 (0.0182) *(0.0238)*	-0.0107 (0.0108) *(0.0132)*	-0.0014 (0.0150) *(0.0200)*	0.0239 (0.0287) *(0.0349)*	-0.0014 (0.0089) *(0.0102)*	0.0147 (0.0121) *(0.0160)*	0.0313 (0.0308) *(0.0355)*	0.0021 (0.0148) *(0.0176)*	0.0165 (0.0151) *(0.0192)*			
LFSTOCK												
ECACRA	-0.0332 (0.0119)*** *(0.0145)***	-0.0223 (0.0071)*** *(0.0085)****	-0.0345 (0.0108)*** *(0.0123)****									
UNEM				-1.7372 (1.8359) *(2.1020)*	0.6224 (0.6127) *(0.7629)*	1.9000 (0.9162)** *(0.8548)***						
INACTIVE							-1.5061 (1.2721) *(1.4377)*	-0.9230 (0.9194) *(1.0003)*	-2.2723 (1.2988)* *(1.7279)*			
ECACRF				-0.0396 (0.0226)* *(0.0285)*	-0.0168 (0.0052)*** *(0.0062)****	-0.0175 (0.0074)** *(0.0072)***	-0.0383 (0.0200)* *(0.0247)*	-0.0230 (0.0088)*** *(0.0101)***	-0.0384 (0.0111)*** *(0.0137)****			
OBS.	325			325			325					

NOTES: (*), (**), and (***) indicates significance at the 10%, 5% and 1% level, respectively. *(*), (**), and (***)* denotes the significance of the White (1980) estimator (robust standard errors) at the 10%, 5% and 1% level, respectively.

Generally speaking, the exogenous, predetermined, and endogenous parameters are similar to one another, denoting the robustness of the dynamic results. First, all of the equations (Appendix A.8) reject that the lagged income inequality coefficient is zero. The coefficient on the lagged dependent variable is higher when the explanatory variables are assumed to be exogenous, except for Regression 1, and lower when the explanatory variables are endogenous, except for Regression 5. Additionally, the coefficients on the lagged dependent variable are statistically significant at the one per cent level in most equations. One expected finding is that income inequality in the current period depends on income inequality in the previous period. The rationale for this result is simple, because income inequality does not change radically over one year and job mobility is rather low. People tend not to change jobs for psychological, technological, and institutional reasons (Gujarati 2003).

Regression 1 indicates that income inequality $(IGE1)$ increases in the long-run as income per capita (IMN) increases, thus leading to a positive correlation between the two variables. The coefficients are also statistically significant in most equations. For instance, if the strictly exogenous income is increased by one per cent, income inequality will rise by 0.0331 per cent in the long-run. This disproves the theory relating to the declining segment of the Kuznets curve, but is likely to fail to reject Lydall's (1979) hypothesis that only a limited number of people can be transferred to higher levels of skills, while the remainder have to wait their turn. This result is consistent with the FEs conditional regressions.

The findings also indicate that income inequality in a region declines over time as the human capital variables (educational attainment (EMN) and educational inequality $(EGE1)$) decline, but only when they are assumed to be endogenous. According to the estimated value and assuming, for example, that human capital variables are endogenous, a one per cent increase in the coefficient on educational attainment would lead in the long-run to a 0.3018 per cent increase in income inequality (Regression 2). The effects of educational attainment and educational inequality obtained after full adjustment of income inequality are positive and statistically significant only when education is endogenous (equations 2c, 3c, and 4c). The combined positive impact of educational attainment and inequality on income inequality implies that, although educational expansion improves the opportinities for individuals, the returns are higher for the rich than for the poor and rich people have more opportunities to engage in higher paid jobs. Additionally, the positive relationship between income and educational inequality highlights the responsiveness of the EU labour market to differences in qualifications and skills. Education is likely to raise the individual's marginal product in the future and therefore his/her future income (Barr 2004: 296).

The long-run effect of the population ageing (AGE) variable on inequality is in most equations positive, which may reflect that with greater longevity, there will be a growing number of elderly people. Since their income is lower than that of younger people, an increasing number of elderly people should lead to a rise in the number of households with a low income (Estudillo 1997: 68), but this variable is not statistically significant. Regression 4 (equations 4a and 4b) shows that the labour force stock $(LFSTOCK)$ has a positive effect on income inequality, but it is not statistically significant either. Nevertheless, the impact of the economic activity rate $(ECACRA)$ has the expected sign (negative) and is statistically significant at the one per cent level (Regression 5). High unemployment $(UNEM)$ is associated with higher inequality in the long-run only when unemployment is endogenous. This outcome is consistent with the outcome of the static regression models, denoting the robustness of the relationship between unemployment and inequality. The dynamic models are likely to allow testing of whether changes in short-term (cyclical) and long-term (structural) unemployment influence changes in income inequality. The short-run and long-run impact of unemployment on inequality has the 'right' sign with respect to the literature and the static regression analysis. Finally, the impact of women's access to work $(ECACRF)$ on income inequality is negative and statistically significant no matter what the explanatory variables are assumed to be.

Equation 6c is the most appropriate, taking into account the specification tests. In this equation, unemployment and female participation in the labour force are the most significant factors in determining income inequality within European regions. More specifically, the higher the unemployment, the higher the income inequality and the higher the female participation, the lower the income inequality.

The dynamic regression results of income inequality for normally working people $(NGE1)$ are similar to the dynamic regression results of income inequality for the population as a whole.[12] As expected, all the equations reject the lagged income inequality for working people parameter is zero, because few workers change job within one year. Regarding that income persistence is an essential characteristic of rewarding achievement (Lane 1971), the results show that most individuals remain at the same economic status. Analysing the long-run coefficients on the determinants of income variations of normally working people, income per capita, once more, positively affects income inequality, but this impact is sensitive to the model specification in terms of the assumption of the determinants (whether they are exogenous, predetermined, or endogenous). The results also indicate that the long-run impact of human capital distribution on income inequality is not clear. Both educational attainment and educational inequality are not statistically significant, except for educational inequality in equation where explanatory variables are income per capita, educational attainment and inequality, and they are assumed to be predetermined. In this case, the higher the educational inequality, the higher the income inequality. Since both income and human capital inequalities have decreased slightly between 1995 and 2000, more equal education may lead to greater equality in economic opportunities and incomes. Population ageing has an ambiguous effect on income inequality, while the female participation in labour force has a negative and statistically significant effect.

6. Concluding Remarks and Further Research

Different static and dynamic panel data analyses have been conducted in order to examine how microeconomic changes in educational distribution in terms of both educational attainment and educational inequality as well as changes in income per capita affect the evolution of income inequality across regions of the EU over the period 1995-2000. Our methodology incorporates variability both across regions and over time.

[12] The short-run and long-run GMM results of income inequality for normally working people are reported in Appendices A.9 and A.10, respectively.

Taking into account the specification tests applied to the estimated models, the relationship between income per capita and income inequality seems to be positive, no matter what income distribution is considered. Regional economic development seems to increase the occupational choices and the earning opportunities not of the population as a whole, but of rich people. While the impact of educational attainment on income inequality is not clear, educational inequality is associated with higher income inequality. It is human capital inequality that seems to matter. It is worth noting that the coefficients on educational inequality are higher when the dependent variable is income inequality for the population as a whole rather than income inequality for normally working people.

The impact of population ageing within a region on income inequality is sensitive to the definition of income distribution. Unemployment is positively associated to income inequality, while work access negatively. The coefficient on inactivity is negative, but sensitive to the model specification. Taking into account urbanisation, an increasing weight of the urban relative to the rural population means a decrease in income inequality for the population as a whole (OLS and REs results). In contrast, the impact of urbanisation on income inequality for normally working people is not clear. Hence, the impact of urbanisation on income inequality is sensitive to the definition of income distribution. Inequality is also associated to latitude: it is lower in the north than in the south. Finally, considering institutions, the social-democratic welfare states, the mainly Protestant regions and those with Nordic family structures are among the most egalitarian.

The results have policy implications as they shed light on the ambiguous impact of income per capita on income inequality. They show that improving access to education, providing higher quality of education, and generally, increasing educational attainment may not have the desired effect on income inequality. They also indicate that income and educational inequality are connected, highlighting the responsiveness of the EU labour market to differences in qualifications and skills. Since both income and human capital inequalities have decreased slightly between 1995 and 2000, a more equal educational distribution may help to improve the economic opportunities and incomes of the less well-off without challenging the European social systems and without requiring any major redistribution of capital.

Although our methodology seems to address the question of how changes in income per capita, educational attainment, and education inequality affect the observed income inequality, further research is needed. First of all, the fact that data on only a limited time period were available means that the results should be interpreted with some caution. Longer time-series will reinforce the analysis. A potential limitation of the analysis — which is also a limitation in most cross-sectional studies — is the fact that regions are more homogeneous than countries, because the regions are subunits of a single national entity (Nielsen and Alderson 1997). Regions do not encompass as wide a range of variation in income and educational distribution, in economic development and in some unobserved characteristics, such as institutions and socio-cultural conditions, as a cross-national sample. Regional boundaries may not define autonomous and internally integrated socioeconomic systems with respect to the distributional process (Nielsen and Alderson 1997). Thus, the administrative boundaries used to organise the data series do not coincide perfectly with the actual boundaries, introducing nuisance spatial autocorrelation into data (Anselin and Rey 1991). It would be valuable to refine the results on regional economic growth by considering data spanning longer periods. In terms of the quality of data, the fact that people are classified into just three categories with respect to the education level completed is a limitation. Finally, the analysis could be extended to spatial econometrics (i.e. Anselin 1988). Spatial econometric techniques can provide a natural framework to test for the occurrence of interregional externalities, and to estimate their magnitude (Vaya, Lopez-Bazo et al. 2004).

Bibliography

Aghion, P. and P. Bolton (1992). "Distribution and Growth in Models of Imperfect Capital Markets." European Economic Review **36**(2-3): 603-611.

Ahluwalia, M. S. (1976). "Inequality, Poverty and Development." Journal of Development Economics **3**(4): 307-342.

Ahn, S. C. and P. Schmidt (1995). "Efficient Estimation of Models for Dynamic Panel Data." Journal of Econometrics **68**(1): 5-27.

Anand, S. and S. M. R. Kanbur (1993). "The Kuznets Process and the Inequality Development Relationship." Journal of Development Economics **40**(1): 25-52.

Anderson, T. W. and C. Hsiao (1981). "Estimation of Dynamic Models with Error Components." Journal of the American Statistical Association **76**(375): 598-606.

Anderson, T. W. and C. Hsiao (1982). "Formulation and Estimation of Dynamic Models Using Panel Data." Journal of Econometrics **18**(1): 47-82.

Anselin, L. (1988). Spatial Econometrics: Methods and Models. Dordrecht; Boston, Kluwer Academic Publishers.

Anselin, L. and S. Rey (1991). "Properties of Tests for Spatial Dependence in Linear Regression Models." Geographical Analysis **23**(2): 112-131.

Arellano, M. and S. Bond (1991). "Some Tests of Specification for Panel Data - Monte-Carlo Evidence and an Application to Employment Equations." Review of Economic Studies **58**(2): 277-297.

Ayala, L., R. Martinez, et al. (2002). "Institutional Determinants of the Unemployment Earnings Inequality Trade-Off." Applied Economics **34**(2): 179-195.

Baltagi, B. H. (2005). Econometric Analysis of Panel Data. Chichester, John Wiley.

Barr, N. A. (2004). The Economics of the Welfare State. Oxford, Oxford University Press.

Becker, G. S. (1962). "Investment in Human Capital - a Theoretical Analysis." Journal of Political Economy **70**(5): 9-49.

Becker, G. S. (1964). Human Capital: a Theoretical and Empirical Analysis with Special Reference to Education. New York, Columbia University Press for National Bureau of Economic Research.

Becker, G. S. and B. R. Chiswick (1966). "Economics of Education and Distribution of Earnings." American Economic Review **56**(2): 358-369.

Berthoud, R. and M. Iacovou (2004). Social Europe: Living Standards and Welfare States. Cheltenham, UK, Northampton, MA, Edward Elgar.

Blinder, A. S. (1974). Toward an Economic Theory of Income Distribution. Cambridge, Mass. ; London, M.I.T. Press.

Bourguignon, F. and C. Morrisson (1990). "Income Distribution, Development and Foreign Trade - a Cross-Sectional Analysis." European Economic Review **34**(6): 1113-1132.

Bourguignon, F. and C. Morrisson (1998). "Inequality and Development: the Role of Dualism." Journal of Development Economics **57**(2): 233-257.

Breusch, T. S. and A. R. Pagan (1980). "The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics." Review of Economic Studies **47**(1): 239-253.

Brown, E. H. P. (1977). The Inequality of Pay. Oxford, Oxford University Press.

Chakraborty, S. and M. Das (2005). "Mortality, Human Capital and Persistent Inequality." Journal of Economic Growth **10**(2): 159-192.

Champernowne, D. G. and F. A. Cowell (1998). Economic Inequality and Income Distribution. Cambridge [England]; New York, Cambridge University Press.

Checchi, D. (2000). Does Educational Achievement Help to Explain Income Inequality?, Departmental Working Papers 2000-11, Department of Economics, University of Milan, Italy, revised Jan 2000.

Combes, P. P., G. Duranton, et al. (2005). "Agglomeration and the Adjustment of the Spatial Economy." Papers in Regional Science **84**(3): 311-349.

Cooper, S. J. (1998). Redistribution and the Persistence of Income Inequality, John F. Kennedy School of Government, Harvard University.

Cooper, S. J., S. N. Durlauf, et al. (1994). "On the Evolution of Economic Status Across Generations." American Statistical Association Papers and Proceedings: 50-58.

Cornia, G. A., T. Addison, et al. (2001). Income Distribution Changes and their Impact in the Post-World War II Period, WIDER Discussion Paper No. 2001/89, World Institute for Development Economics Research, United Nations University.

De Gregorio, J. and J. W. Lee (2002). "Education and Income Inequality: New Evidence from Cross-Country Data." Review of Income and Wealth(3): 395-416.

Durlauf, S. N. (1996). "A Theory of Persistent Income Inequality." Journal of Economic Growth **1**(1): 75-93.

Esping-Andersen, G. (1990). The Three Worlds of Welfare Capitalism. Cambridge, Polity.

Estudillo, J. P. (1997). "Income Inequality in the Philippines, 1961-91." Developing Economies **35**(1): 68-95.

Ezcurra, R., Pascual, P. and Rapún, M. (2008) "Is income inequality harmful for regional growth? Evidence from the European Union." Urban Studies (forthcoming).

Ferrera, M. (1996). "The “Southern Model” of Welfare in Social Europe." Journal of European Social Policy **6**(1): 17-37.

Fields, G. S. (1979). "Welfare Economic Approach to Growth and Distribution in the Dual Economy." Quarterly Journal of Economics **93**(3): 325-353.

Firebaugh, G. (2003). The New Geography of Global Income Inequality. Cambridge, MA, Harvard University Press.

Galor, O. and D. Tsiddon (1997a). "The Distribution of Human Capital and Economic Growth." Journal of Economic Growth **2**(1): 93-124.

Galor, O. and D. Tsiddon (1997b). "Technological Progress, Mobility, and Economic Growth." American Economic Review **87**(3): 363-382.

Galor, O. and J. Zeira (1993). "Income Distribution and Macroeconomics." Review of Economic Studies **60**(1): 35-52.

Glomm, G. and B. Ravikumar (1992). "Public versus Private Investment in Human Capital - Endogenous Growth and Income Inequality." Journal of Political Economy **100**(4): 818-834.

Gottschalk, P. and T. M. Smeeding (1997). "Cross-National Comparisons of Earnings and Income Inequality." Journal of Economic Literature **35**(2): 633-687.

Graham, C. (2002). "Mobility, Opportunity, and Vulnerability: the Dynamics of Poverty and Inequality in a Global Economy." Journal of Human Development **3**(1): 57-94.

Greene, W. H. (2003). Econometric Analysis. Upper Saddle River, N.J., [Great Britain], Prentice Hall.

Griliches, Z. (1997). "Education, Human Capital, and Growth: a Personal Perspective." Journal of Labor Economics **15**(1): S330-S344.

Gujarati, D. N. (2003). Basic Econometrics. Boston, London, McGraw-Hill.

Hannum, E. and C. Buchmann (2005). "Global Educational Expansion and Socio-Economic Development: an Assessment of Findings from the Social Sciences." World Development **33**(3): 333-354.

Hassler, J. and J. V. R. Mora (2000). "Intelligence, Social Mobility, and Growth." American Economic Review **90**(4): 888-908.

Hausman, J. A. (1978). "Specification Tests in Econometrics." Econometrica **46**(6): 1251-1271.

Higgins, M. and J. Williamson (1999). Explaining Inequality the World Round: Cohort Size, Kuznets Curves, and Openness, NBER Working Paper No. 7224, National Bureau of Economic Research.

Hsiao, C. (2003). Analysis of Panel Data. Cambridge, Cambridge University Press.

Jacobs, D. (1985). "Unequal Organizations or Unequal Attainments - an Empirical Comparison of Sectoral and Individualistic Explanations for Aggregate Inequality." American Sociological Review **50**(2): 166-180.

Knight, J. B. (1976a). "Devaluation and Income Distribution in Less Developed Economies." Oxford Economic Papers-New Series **28**(2): 208-227.

Knight, J. B. (1976b). "Explaining Income Distribution in Less Developed Countries - Framework and an Agenda." Oxford Bulletin of Economics and Statistics **38**(3): 161-177.

Knight, J. B. and R. H. Sabot (1983). "Educational Expansion and the Kuznets Effect." American Economic Review **73**(5): 1132-1136.

Kuznets, S. (1955). "Economic Growth and Income Inequality." American Economic Review **45**(1): 1-28.

Lane, D. S. (1971). The End of Inequality? Stratification under State Socialism. Harmondsworth, Penguin.

Lecaillon, J. (1984). Income Distribution and Economic Development: an Analytical Survey. Geneva, International Labour Office.

Lopez, R. E., V. Thomas, et al. (1998). Addressing the Education Puzzle: the Distribution of Education and Economic Reforms. Washington, DC, World Bank, Economic Development Institute, Office of the Director and Macroeconomic Management and Policy Division.

Lydall, H. (1979). A Theory of Income Distribution. Oxford, Clarendon Press.

Mincer, J. (1958). "Investment in Human Capital and Personal Income Distribution." Journal of Political Economy **66**(4): 281-302.

Mincer, J. (1962). "On-the-Job Training - Costs, Returns, and Some Implications." Journal of Political Economy **70**(5): 50-79.

Mincer, J. (1974). Schooling, Experience and Earnings. New York, National Bureau of Economic Research.

Mocan, H. N. (1999). "Structural Unemployment, Cyclical Unemployment, and Income Inequality." Review of Economics and Statistics **81**(1): 122-134.

Motonishi, T. (2000). On the Effects of the Development of International Financial Markets When the World Economy Is Stratified, Discussion Paper Series No. 2000-02, Faculty of Economics, Nagasaki University.

Motonishi, T. (2006). "Why Has Income Inequality in Thailand Increased? an Analysis Using Surveys from 1975 to 1998." Japan and the World Economy **18**(4): 464-487.

Nielsen, F. and A. S. Alderson (1997). "The Kuznets Curve and the Great U-Turn: Income Inequality in US Countries, 1970 to 1990." American Sociological Review **62**(1): 12-33.

Papanek, G. F. and O. Kyn (1986). "The Effect on Income Distribution of Development, the Growth Rate and Economic Strategy." Journal of Development Economics **23**(1): 55-65.

Park, K. H. (1996). "Educational Expansion and Educational Inequality on Income Distribution." Economics of Education Review **15**(1): 51-58.

Psacharopoulos, G. and A.-M. Arriagada (1986). The Educational Attainment of the Labor Force: an International Comparison, World Bank Discussion Paper, Washington, D.C.

Ram, R. (1990). "Educational Expansion and Schooling Inequality - International Evidence and Some Implications." Review of Economics and Statistics **72**(2): 266-273.

Robinson, S. (1976). "U-Hypothesis Relating Income Inequality and Economic Development." American Economic Review **66**(3): 437-440.

Rodríguez-Pose, A. and Vilalta-Bufi. M. (2005) "Education, migration, and job satisfaction: the regional returns of human capital in the EU." Journal of Economic Geography **5**(5): 545-566.

Ryscavage, P., G. Green, et al. (1992). The Impact of Demographic, Social, and Economic Change on the Distribution of Income, U.S. Bureau of the Census, Ser. P60, no. 183. Washington, D.C., U.S. Bureau of the Census.

Saint-Paul, G. and T. Verdier (1993). "Education, Democracy and Growth." Journal of Development Economics **42**(2): 399-407.

Sargan, J. D. (1958). "The Estimation of Economic Relationships Using Instrumental Variables." Econometrica **26**(3): 393-415.

Schultz, T. W. (1961). "Investment in Human Capital." American Economic Review **51**(1-2): 1-17.

Schultz, T. W. (1962). "Reflections on Investment in Man." Journal of Political Economy **70**(5): 1-8.

Schultz, T. W. (1963). The Economic Value of Education. New York, Columbia University Press.

Spence, A. M. (1973). "Job Market Signaling." Quarterly Journal of Economics **87**(3): 355-374.

Spence, A. M. (1974). Market Signalling: Informational Transfer in Hiring and Related Screening Processes. Cambridge [Mass], Harvard University Press.

Spence, A. M. (1976). "Informational Aspects of Market Structure - Introduction." Quarterly Journal of Economics **90**(4): 591-597.

Theil, H. (1967). Economics and Information Theory. Amsterdam, North-Holland.

Thomas, V., Y. Wang, et al. (2001). Measuring Education Inequality: GINI Coefficients of Education. Washington, D.C., World Bank.
Thorbecke, E. and C. Charumilind (2002). "Economic Inequality and its Socioeconomic Impact." World Development **30**(9): 1477-1495.
Tinbergen, J. (1975). Income Distribution: Analysis and Policies. Amsterdam; Oxford, North Holland Publishing Co. [etc.].
Vaya, E., E. Lopez-Bazo, et al. (2004). Growth and Externalities across Economies: an Empirical Analysis Using Spatial Econometrics. Advances in Spatial Econometrics: Methodology, Tools and Applications. L. Anselin, R. J. G. M. Florax and S. J. Rey. Berlin, Springer**:** 433-455.
White, H. (1980). "A Heteroskedasticity Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." Econometrica **48**(4): 817-838.
Williamson, J. G. (1991). Inequality, Poverty, and History: the Kuznets Memorial Lectures of the Economic Growth Center, Yale University. Oxford, Basil Blackwell.
Wolf, A. (2004). "Education and Economic Performance: Simplistic Theories and their Policy Consequences." Oxford Review of Economic Policy **20**(2): 315-333.
Wooldridge, J. M. (2003). Introductory Econometrics: a Modern Approach. Mason, Ohio, United Kingdom, South-Western College Publishing.
World Bank (2002). Achieving Education for All by 2015: Simulation Results for 47 Low-Income Countries, Washington, DC: Human Development Network, Africa Region and Education Department, World Bank: 76.

Appendix A.1: Descriptive statistics of ECHP dataset

Year	Statistic	Quantitative variables			Qualitative variables			
					Main activity status			
		Income	Educational attainment	Age	Unemployed	Inactive	Normally working	Urbanisation
1995	Obs	120413	119463	125395	7915	55169	61406	26863
	Mean	9744.58	0.60	44.96				
	Percentage				6.36	44.32	49.33	46.68
	Std. Dev.	11782.83	0.73	18.23				
	Variance	1.39E+08	0.53	332.35				
	Skewness	8.39	0.78	0.34				
	Kurtosis	311.52	2.27	2.12				
1996	Obs	124663	114529	120413	7685	58933	53214	26863
	Mean	10163.60	0.60	45.05				
	Percentage				6.41	44.41	49.18	46.68
	Std. Dev.	11234.33	0.73	18.28				
	Variance	1.26E+08	0.53	334.28				
	Skewness	6.45	0.79	0.35				
	Kurtosis	205.83	2.27	2.12				
1997	Obs	117886	118402	124756	7760	54183	62221	26863
	Mean	10472.71	0.62	45.22				
	Percentage				6.25	43.64	50.11	46.68
	Std. Dev.	11529.87	0.74	18.32				
	Variance	1.33E+08	0.55	335.47				
	Skewness	6.87	0.73	0.34				
	Kurtosis	213.47	2.17	2.13				
1998	Obs	113455	115953	117980	6775	50646	59978	26863
	Mean	10617.48	0.68	45.54				
	Percentage				5.77	43.14	51.09	46.68
	Std. Dev.	12648.77	0.76	18.32				
	Variance	1.60E+08	0.57	335.66				
	Skewness	16.09	0.60	0.34				
	Kurtosis	1049.18	1.97	2.13				
1999	Obs	108731	112406	113536	5908	48802	58342	26863
	Mean	11037.64	0.68	45.78				
	Percentage				5.23	43.17	51.61	46.68
	Std. Dev.	13552.43	0.77	18.33				
	Variance	1.84E+08	0.59	336.04				
	Skewness	30.58	0.63	0.33				
	Kurtosis	3616.64	1.96	2.13				
2000	Obs	104953	107751	108848	5165	46890	56384	26863
	Mean	11368.55	0.69	46.07				
	Percentage				4.76	43.24	52	46.68
	Std. Dev.	12884.93	0.77	18.45				
	Variance	1.66E+08	0.59	340.32				
	Skewness	10.55	0.59	0.32				
	Kurtosis	442.83	1.92	2.12				

Appendix A.2: REs: Dependent variable is income inequality for the population as a whole (IGE1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
IMN	-0.0036 (0.0011)*** (0.0013)***	-0.0012 (0.0014) (0.0015)	-0.0009 (0.0015) (0.0016)	0.0008 (0.0015) (0.0015)	-0.0001 (0.0015) (0.0015)	0.0020 (0.0015) (0.0015)***	0.0014 (0.0015) (0.0015)	0.0020 (0.0017) (0.0017)	0.0042 (0.0014)*** (0.0014)***	0.0053 (0.0015)*** (0.0015)***	0.0030 (0.0015)** (0.0014)**	0.0054 (0.0015)*** (0.0015)***
EMN		0.0371 (0.0304) (0.0339)	0.0370 (0.0305) (0.0340)	0.0658 (0.0298)*** (0.0310)***	0.0175 (0.0286) (0.0293)	0.0359 (0.0275) (0.0270)***	0.0386 (0.0278) (0.0278)	0.0697 (0.0318)** (0.0342)**	0.0217 (0.0257) (0.0249)	0.0189 (0.0272) (0.0266)	0.0496 (0.0276)* (0.0290)*	0.0230 (0.0260) (0.0259)
EGE1		0.0847 (0.0222)*** (0.0267)***	0.0879 (0.0223)*** (0.0268)***	0.0901 (0.0213)*** (0.0244)***	0.0519 (0.0202)** (0.0205)**	0.0600 (0.0193)*** (0.0182)***	0.0591 (0.0194)*** (0.0181)***	0.0802 (0.0255)*** (0.0282)***	0.0422 (0.0180)** (0.0170)**	0.0446 (0.0192)** (0.0173)**	0.0684 (0.0194)*** (0.0208)***	0.0477 (0.0182)*** (0.0170)***
AGE			-0.0042 (0.0022)* (0.0025)*	-0.0056 (0.0021)*** (0.0027)***	-0.0078 (0.0020)*** (0.0021)***	-0.0044 (0.0020)** (0.0020)***	-0.0069 (0.0020)*** (0.0022)***	-0.0061 (0.0026)** (0.0025)**	-0.0057 (0.0018)*** (0.0019)***	-0.0061 (0.0019)*** (0.0020)***	-0.0058 (0.0020)*** (0.0020)***	-0.0061 (0.0019)*** (0.0020)***
LFSTOCK				-0.6963 (0.0788)*** (0.0895)***								
ECACRA					-0.0131 (0.0010)*** (0.0011)***							
UNEM						0.3933 (0.1301)*** (0.1402)***		0.5955 (0.2030)*** (0.2215)***	0.4711 (0.1215)*** (0.1327)***	0.5059 (0.1272)*** (0.1374)***	0.4550 (0.1300)*** (0.1436)***	0.5122 (0.1248)*** (0.1374)***
INACTIVE							0.1725 (0.0882)* (0.0894)*					
ECACRF						-0.0111 (0.0008)*** (0.0008)***	-0.0110 (0.0008)*** (0.0009)***	-0.0083 (0.0011)*** (0.0012)***	-0.0073 (0.0008)*** (0.0009)***	-0.0073 (0.0009)*** (0.0009)***	-0.0089 (0.0008)*** (0.0010)***	-0.0072 (0.0009)*** (0.0009)***
URBANDPAV (fixed)								-0.1538 (0.0467)*** (0.0446)***				
LAT (fixed)									-0.0120 (0.0013)*** (0.0012)***			
DWSLIB										0.0621 (0.0284)** (0.0241)**		
DWSCORP										0.0594 (0.0291)** (0.0249)**		
DWSRES										0.2259 (0.0357)*** (0.0301)***		
DRLCATH											0.0955 (0.0221)*** (0.0248)***	
DRLORTH											0.2243 (0.0411)*** (0.0373)***	
DRLANGL											0.0262 (0.0219) (0.0248)	
DFNORD												-0.0599 (0.0265)** (0.0222)***
DFSC												0.1680 (0.0200)*** (0.0193)***
OBS.	604	596	596	596	513	513	513	299	513	513	513	513

Appendix A.3: Dummy variables definition

Variable	**Definition**
Welfare state	
DWSSOC	Socialism (social democratic)
DWSLIB	Liberal
DWSCORP	Corporatist (conservatism)
DWSRES	Residual (‘Southern’)
Religion	
DRLPROT	Mainly Protestant
DRLCATH	Mainly Catholic
DRLORTH	Mainly Orthodox
DRLANGL	Mainly Anglicans
Family structure	
DFNORD	Nordic (Scandinavian)
DFNC	North/Central
DFSC	Southern/Catholic

Appendix A.4: Standardised coefficients

Dependent variable is income inequality for the population as a whole (IGE1)

DEPENDENT VARIABLE: IGE1									
	REGR. 1	REGR. 2	REGR. 3	REGR. 4	REGR. 5	REGR. 6	REGR. 7	REGR. 8	REGR. 9
IMN	-0.6514	-0.3659	-0.3360	-0.0449	-0.1675	-0.0845	-0.1105	-0.2136	0.0526
EMN		-0.5168	-0.5331	-0.1467	0.0171	0.0877	0.1149	0.1418	0.0624
EGE1		-0.1598	-0.1185	0.2067	0.2553	0.2854	0.2460	0.1985	0.1545
AGE			-0.1662	-0.2178	-0.1712	-0.0964	-0.1661	-0.0537	-0.0945
LFSTOCK				-0.5644					
ECACRA					-0.5712				
UNEM						0.0531		0.1887	0.0501
INACTIVE							0.1974		
ECACRF						-0.6773	-0.5612	-0.5035	-0.4929
URBANDPAV (fixed)								-0.1148	
LAT (fixed)									-0.4330

Dependent variable is income inequality for normally working people (NGE1)

DEPENDENT VARIABLE: NGE1						
	REGR. 1	REGR. 2	REGR. 3	REGR. 4	REGR. 5	REGR. 6
NMN	-0.3975	-0.0187	-0.0196	-0.0309	-0.1803	0.1063
EMN		0.0020	0.0023	0.3836	0.1752	0.3665
EGE1		0.5368	0.5340	0.6557	0.3556	0.4877
AGE			0.0118	0.0515	0.1522	0.0567
ECACRF				-0.3757	-0.1102	-0.0985
URBANDPAV (fixed)					-0.0883	
LAT (fixed)						-0.5556

Appendix A.5: FEs: Dependent variable is income inequality for normally working people (NGE1)

	(1)	(2)	(3)	(4)	(5)	(6)
NMN	0.0014 (0.0008)* *(0.0013)*	0.0022 (0.0011)** *(0.0015)*	0.0023 (0.0011)** *(0.0016)*	0.0020 (0.0012) *(0.0014)*	0.0074 (0.0019)*** *(0.0021)****	0.0046 (0.0014)*** *(0.0016)****
EMN		0.0347 (0.0304) *(0.0292)*	0.0349 (0.0304) *(0.0293)*	0.0322 (0.0295) *(0.0254)*	-0.0055 (0.0419) *(0.0330)*	0.0250 (0.0325) *(0.0268)*
EGE1		0.0545 (0.0233)** *(0.0169)****	0.0546 (0.0233)** *(0.0169)****	0.0326 (0.0220) *(0.0147)***	0.0596 (0.0319)* *(0.0219)****	0.0377 (0.0221)* *(0.0146)***
AGE			-0.0006 (0.0022) *(0.0020)*	-0.0017 (0.0021) *(0.0019)*	-0.0011 (0.0028) *(0.0024)*	0.0000 (0.0023) *(0.0019)*
ECACRF				-0.0035 (0.0012)*** *(0.0011)****	-0.0012 (0.0018) *(0.0016)*	-0.0020 (0.0013) *(0.0013)*
YR96*UR BANDPAV					-0.0101 (0.0155) *(0.0134)*	
YR97*UR BANDPAV					-0.0316 (0.0156)** *(0.0145)***	
YR98*UR BANDPAV					0.0126 (0.0171) *(0.0157)*	
YR99*UR BANDPAV					-0.0129 (0.0180) *(0.0168)*	
YR00*UR BANDPAV					-0.0570 (0.0188)*** *(0.0167)****	
YR96*LAT						0.0000 (0.0001) *(0.0001)*
YR97*LAT						-0.0002 (0.0001) *(0.0001)**
YR98*LAT						0.0000 (0.0001) *(0.0001)*
YR99*LAT						-0.0002 (0.0001) *(0.0001)**
YR00*LAT						-0.0004 (0.0001)*** *(0.0001)****
CONSTANT	0.2019 (0.0127)*** *(0.0186)****	0.1231 (0.0390)*** *(0.0328)****	0.1486 (0.1035) *(0.0878)**	0.3855 (0.1096)*** *(0.0841)****	0.1991 (0.1658) *(0.1255)*	0.2071 (0.1320) *(0.1040)***
ADJ R-SQ	0.0057	0.0207	0.0209	0.0337	0.1556	0.0682
OBS.	604	596	596	513	299	513
LM TEST (p-value)	676.24 (0.0000)	555.86 (0.0000)	555.66 (0.0000)	557.12 (0.0000)	259.68 (0.0000)	538.47 (0.0000)
HAUSMAN TEST (p-value)	38.07 (0.0000)	34.03 (0.0000)	34.36 (0.0000)	14.72 (0.0116)		

NOTES: (*), (**), and (***) indicates significance at the 10%, 5% and 1% level, respectively. *(*), (**), and (***)* denotes the significance of the White (1980) estimator (robust standard errors). LM TEST is the Lagrange multiplier test for the random effects model based on the OLS residuals (Breusch and Pagan 1980). HAUSMAN TEST is the Hausman (1978) test for fixed or random effects.

Appendix A.6: OLS: Dependent variable is income inequality for normally working people (NGE1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NMN	-0.0068 (0.0006)*** (0.0008)***	-0.0003 (0.0008) (0.0009)	-0.0003 (0.0008) (0.0009)	-0.0005 (0.0008) (0.0009)	-0.0027 (0.0016)* (0.0019)	0.0018 (0.0008)** (0.0009)*	0.0034 (0.0010)*** (0.0011)***	0.0017 (0.0008)** (0.0010)*	0.0039 (0.0009)*** (0.0011)***
EMN		0.0006 (0.0198) (0.0180)	0.0006 (0.0198) (0.0180)	0.1061 (0.0241)*** (0.0262)***	0.0404 (0.0313) (0.0298)	0.1013 (0.0224)*** (0.0232)***	0.0435 (0.0259)* (0.0287)	0.1008 (0.0263)*** (0.0297)***	0.0626 (0.0226)*** (0.0248)**
EGE1		0.0949 (0.0129)*** (0.0134)***	0.0944 (0.0130)*** (0.0140)***	0.1203 (0.0138)*** (0.0155)***	0.0578 (0.0184)*** (0.0177)***	0.0895 (0.0133)*** (0.0139)***	0.0710 (0.0152)*** (0.0168)***	0.1235 (0.0151)*** (0.0179)***	0.0791 (0.0131)*** (0.0142)***
AGE			0.0005 (0.0014) (0.0013)	0.0020 (0.0014) (0.0014)	0.0053 (0.0018)*** (0.0016)***	0.0022 (0.0013)* (0.0013)*	0.0026 (0.0013)** (0.0013)*	-0.0010 (0.0014) (0.0013)	0.0026 (0.0013)** (0.0013)**
ECACRF				-0.0031 (0.0004)*** (0.0005)***	-0.0008 (0.0006) (0.0006)	-0.0008 (0.0005)* (0.0005)	0.0007 (0.0006) (0.0006)	-0.0019 (0.0004)*** (0.0005)***	0.0011 (0.0005)** (0.0005)**
URBANDPAV (fixed)					-0.0261 (0.0181) (0.0172)				
LAT (fixed)						-0.0064 (0.0007)*** (0.0008)***			
DWSLIB							0.1068 (0.0134)*** (0.0102)***		
DWSCORP							0.0995 (0.0133)*** (0.0099)***		
DWSRES							0.1945 (0.0201)*** (0.0187)***		
DRLCATH								0.0352 (0.0086)*** (0.0086)***	
DRLORTH								0.1528 (0.0152)*** (0.0155)***	
DRLANGL								0.0212 (0.0088)** (0.0093)**	
DFNORD									-0.1054 (0.0124)*** (0.0087)***
DFSC									0.1061 (0.0114)*** (0.0114)***
ADJ R-SQ	0.1566	0.2974	0.2963	0.3557	0.2191	0.4358	0.4512	0.4556	0.4763
OBS.	604	596	596	513	299	513	513	513	513

Appendix A.7: REs: Dependent variable is income inequality for normally working people (NGE1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NMN	-0.0009 (0.0008) *(0.0010)*	0.0011 (0.0010) *(0.0012)*	0.0011 (0.0010) *(0.0012)*	0.0008 (0.0010) *(0.0011)*	0.0002 (0.0013) *(0.0013)*	0.0016 (0.0010) *(0.0011)*	0.0019 (0.0010)* *(0.0011)**	0.0015 (0.0010) *(0.0011)*	0.0019 (0.0010)* *(0.0011)**
EMN		0.0564 (0.0248)** *(0.0249)***	0.0556 (0.0249)** *(0.0251)***	0.0762 (0.0251)*** *(0.0251)****	0.0783 (0.0305)** *(0.0279)****	0.0704 (0.0245)*** *(0.0249)****	0.0523 (0.0260)** *(0.0259)***	0.0705 (0.0259)*** *(0.0264)****	0.0636 (0.0245)*** *(0.0251)***
EGE1		0.0963 (0.0178)*** *(0.0168)****	0.0952 (0.0179)*** *(0.0170)****	0.0762 (0.0177)*** *(0.0158)****	0.0828 (0.0239)*** *(0.0194)****	0.0657 (0.0172)*** *(0.0163)****	0.0573 (0.0183)*** *(0.0158)****	0.0735 (0.0181)*** *(0.0168)****	0.0654 (0.0171)*** *(0.0155)****
AGE			0.0010 (0.0018) *(0.0015)*	-0.0002 (0.0018) *(0.0015)*	0.0003 (0.0024) *(0.0020)*	-0.0007 (0.0017) *(0.0015)*	-0.0003 (0.0017) *(0.0016)*	-0.0012 (0.0017) *(0.0015)*	-0.0003 (0.0017) *(0.0015)*
ECACRF				-0.0037 (0.0006)*** *(0.0006)****	-0.0019 (0.0009)** *(0.0009)***	-0.0014 (0.0007)** *(0.0007)***	-0.0011 (0.0008) *(0.0008)*	-0.0025 (0.0007)*** *(0.0007)****	-0.0006 (0.0008) *(0.0007)*
URBANDPAV (fixed)					-0.0308 (0.0377) *(0.0334)*				
LAT (fixed)						-0.0059 (0.0012)*** *(0.0011)****			
DWSLIB							0.0888 (0.0223)*** *(0.0181)****		
DWSCORP							0.0721 (0.0234)*** *(0.0174)****		
DWSRES							0.1482 (0.0298)*** *(0.0226)****		
DRLCATH								0.0474 (0.0171)*** *(0.0194)***	
DRLORTH								0.1645 (0.0315)*** *(0.0331)****	
DRLANGL								0.0412 (0.0164)** *(0.0193)***	
DFNORD									-0.0840 (0.0209)*** *(0.0158)****
DFSC									0.0773 (0.0166)*** *(0.0149)****
OBS.	604	596	596	513	299	513	513	513	513

Appendix A.8: Short Run GMM: Dependent variable is income inequality for the population as a whole (IGE1)

	REGRESSION 1			REGRESSION 2			REGRESSION 3			REGRESSION 4		
	(a) x_{it} strictly exogenous	(b) x_{it} predetermined	(c) x_{it} endogenous	(a) x_{it} strictly exogenous	(b) x_{it} predetermined	(c) x_{it} endogenous	(a) x_{it} strictly exogenous	(b) x_{it} predetermined	(c) x_{it} endogenous	(a) x_{it} strictly exogenous	(b) x_{it} predetermined	(c) x_{it} endogenous
$IGE1_{i,t-1}$	0.7531 (0.1234)*** *(0.1199)****	0.8135 (0.1230)*** *(0.1445)****	0.6965 (0.1451)*** *(0.1525)****	0.8993 (0.1441)*** *(0.1563)****	0.6388 (0.1232)*** *(0.1743)****	0.4526 (0.1574)*** *(0.2283)***	0.9188 (0.1469)*** *(0.1662)****	0.6125 (0.1212)*** *(0.1717)****	0.4405 (0.1543)*** *(0.2289)**	0.9913 (0.1688)*** *(0.1864)****	0.5709 (0.1219)*** *(0.1857)****	0.4193 (0.1539)*** *(0.2203)**
IMN_{it} $IMN_{i,t-1}$	0.0139 (0.0026)*** *(0.0027)**** -0.0057 (0.0031)* *(0.0032)**	0.0063 (0.0038)* *(0.0044)* -0.0014 (0.0050) *(0.0042)*	0.0132 (0.0042)*** *(0.0050)**** -0.0017 (0.0065) *(0.0045)*	0.0175 (0.0032)*** *(0.0033)**** -0.0109 (0.0045)** *(0.0048)***	0.0202 (0.0055)*** *(0.0061)**** -0.0089 (0.0068) *(0.0081)*	0.0239 (0.0058)*** *(0.0064)**** -0.0108 (0.0075) *(0.0085)*	0.0184 (0.0033)*** *(0.0035)**** -0.0124 (0.0047)*** *(0.0054)***	0.0204 (0.0055)*** *(0.0056)**** -0.0071 (0.0068) *(0.0074)*	0.0241 (0.0058)*** *(0.0061)**** -0.0103 (0.0073) *(0.0081)*	0.0181 (0.0034)*** *(0.0036)**** -0.0137 (0.0050)*** *(0.0061)***	0.0195 (0.0051)*** *(0.0052)**** -0.0035 (0.0066) *(0.0069)*	0.0231 (0.0055)*** *(0.0053)**** -0.0108 (0.0076) *(0.0067)*
EMN_{it} $EMN_{i,t-1}$				0.0901 (0.0518)* *(0.0493)** -0.1282 (0.0504)** *(0.0494)****	0.1584 (0.0775)** *(0.0913)** -0.1375 (0.0503)*** *(0.0448)****	0.2503 (0.0846)*** *(0.1029)*** -0.0850 (0.0701) *(0.0687)*	0.1004 (0.0521)* *(0.0517)** -0.1412 (0.0513)*** *(0.0520)****	0.1577 (0.0763)** *(0.0873)** -0.1423 (0.0498)*** *(0.0439)****	0.2517 (0.0842)*** *(0.0995)*** -0.0895 (0.0694) *(0.0694)*	0.0950 (0.0540)* *(0.0530)** -0.1465 (0.0531)*** *(0.0543)****	0.1478 (0.0703)** *(0.0755)** -0.1316 (0.0492)*** *(0.0416)****	0.2666 (0.0829)*** *(0.0843)**** -0.0900 (0.0688) *(0.0711)*
$EGE1_{it}$ $EGE1_{i,t-1}$				0.0587 (0.0346)* *(0.0256)*** -0.0720 (0.0357)** *(0.0249)****	0.1006 (0.0479)** *(0.0419)*** -0.0677 (0.0370)* *(0.0264)***	0.1275 (0.0572)** *(0.0551)*** -0.0342 (0.0506) *(0.0465)*	0.0560 (0.0352) *(0.0258)*** -0.0735 (0.0361)** *(0.0265)****	0.1029 (0.0478)** *(0.0433)*** -0.0658 (0.0366)* *(0.0259)***	0.1293 (0.0567)** *(0.0559)*** -0.0364 (0.0502) *(0.0468)*	0.0560 (0.0363) *(0.0266)*** -0.0772 (0.0374)** *(0.0280)****	0.1124 (0.0437)** *(0.0398)**** -0.0601 (0.0350)* *(0.0240)***	0.1524 (0.0550)*** *(0.0522)**** -0.0384 (0.0483) *(0.0472)*
AGE_{it} $AGE_{i,t-1}$							0.0092 (0.0049)* *(0.0054)** -0.0011 (0.0033) *(0.0036)*	0.0082 (0.0045)* *(0.0050)** -0.0035 (0.0027) *(0.0030)*	0.0081 (0.0044)* *(0.0051)* -0.0010 (0.0028) *(0.0030)*	0.0100 (0.0051)* *(0.0057)** -0.0018 (0.0034) *(0.0038)*	0.0077 (0.0044)* *(0.0052)* -0.0041 (0.0028) *(0.0030)*	0.0073 (0.0045) *(0.0051)* -0.0004 (0.0030) *(0.0030)*
$LFSTOCK_{it}$ $LFSTOCK_{i,t}$										0.2505 (0.1565) *(0.1739)* 0.0726 (0.1291) *(0.1161)*	0.1588 (0.2936) *(0.3475)* -0.1505 (0.1747) *(0.1589)*	-0.2972 (0.3870) *(0.4391)* 0.2316 (0.3129) *(0.3589)*
$ECACRA_{it}$ $ECACRA_{i,t-1}$												
$UNEM_{it}$ $UNEM_{i,t-1}$												
$INACTIVE_{it}$ $INACTIVE_{i,t}$												
$ECACRF_{it}$ $ECACRF_{i,t-1}$												

OBS.	400			392			392			392		
SARGAN TEST (p-value)	12.26 (0.1989)	26.20 (0.0709)	18.09 (0.1541)	10.67 (0.2988)	49.79 (0.0306)	32.29 (0.0547)	9.54 (0.3888)	48.36 (0.0412)	31.29 (0.0690)	9.29 (0.4107)	59.13 (0.0331)	35.24 (0.0840)
AR(1) TEST (p-value)	-5.85 (0.0000) *-4.42* *(0.0000)*	-6.11 (0.0000) *-4.29* *(0.0000)*	-4.82 (0.0000) *-4.09* *(0.0000)*	-5.64 (0.0000) *-3.82* *(0.0001)*	-5.39 (0.0000) *-3.58* *(0.0003)*	-3.44 (0.0006) *-2.32* *(0.0202)*	-5.72 (0.0000) *-3.77* *(0.0002)*	-5.35 (0.0000) *-3.47* *(0.0005)*	-3.40 (0.0007) *-2.24* *(0.0254)*	-5.57 (0.0000) *-3.72* *(0.0002)*	-5.33 (0.0000) *-3.37* *(0.0008)*	-3.61 (0.0003) *-2.51* *(0.0120)*
AR(2) TEST (p-value)	-1.19 (0.2339) *-0.68* *(0.4977)*	-1.38 (0.1671) *-0.79* *(0.4289)*	-1.14 (0.2562) *-0.65* *(0.5188)*	-1.45 (0.1480) *-0.85* *(0.3941)*	-1.35 (0.1783) *-0.83* *(0.4078)*	-0.89 (0.3725) *-0.60* *(0.5470)*	-1.28 (0.2018) *-0.74* *(0.4573)*	-1.23 (0.2193) *-0.73* *(0.4679)*	-0.78 (0.4356) *-0.51* *(0.6100)*	-1.17 (0.2428) *-0.68* *(0.4996)*	-1.11 (0.2680) *-0.63* *(0.5274)*	-0.96 (0.3361) *-0.69* *(0.4912)*

	REGRESSION 5			REGRESSION 6			REGRESSION 7					
	(a) x_{it} strictly exogenous	(b) x_{it} predetermined	(c) x_{it} endogenous	(a) x_{it} strictly exogenous	(b) x_{it} predetermined	(c) x_{it} endogenous	(a) x_{it} strictly exogenous	(b) x_{it} predetermined	(c) x_{it} endogenous			
$IGE1_{i,t-1}$	0.6263 (0.1278)*** *(0.1423)****	0.4689 (0.1113)*** *(0.1382)****	0.5554 (0.1392)*** *(0.1788)****	0.7371 (0.1434)*** *(0.1626)****	0.3899 (0.0977)*** *(0.1225)****	0.4300 (0.1255)*** *(0.1537)****	0.7274 (0.1365)*** *(0.1499)****	0.5741 (0.1072)*** *(0.1369)****	0.4963 (0.1341)*** *(0.1656)****			
IMN_{it} $IMN_{i,t-1}$	0.0163 (0.0040)*** *(0.0047)**** -0.0106 (0.0045)** *(0.0056)**	0.0054 (0.0062) *(0.0074)* 0.0016 (0.0062) *(0.0081)*	0.0075 (0.0077) *(0.0096)* -0.0037 (0.0076) *(0.0108)*	0.0168 (0.0043)*** *(0.0049)**** -0.0130 (0.0048)*** *(0.0060)***	0.0127 (0.0056)** *(0.0060)*** -0.0042 (0.0054) *(0.0059)*	0.0138 (0.0071)* *(0.0076)** -0.0083 (0.0070) *(0.0080)*	0.0157 (0.0042)*** *(0.0048)**** -0.0128 (0.0047)*** *(0.0055)***	0.0095 (0.0058) *(0.0063)* -0.0021 (0.0055) *(0.0062)*	0.0109 (0.0071) *(0.0081)* -0.0050 (0.0069) *(0.0076)*			
EMN_{it} $EMN_{i,t-1}$	0.0780 (0.0520) *(0.0563)* -0.1182 (0.0473)** *(0.0534)***	0.0277 (0.0751) *(0.0979)* -0.0978 (0.0513)* *(0.0503)**	0.0391 (0.0899) *(0.1158)* -0.1689 (0.0679)** *(0.0810)***	0.0851 (0.0548) *(0.0541)* -0.1214 (0.0504)** *(0.0560)***	0.0866 (0.0654) *(0.0697)* -0.1057 (0.0486)** *(0.0474)***	0.1129 (0.0841) *(0.0960)* -0.1273 (0.0628)** *(0.0676)**	0.0865 (0.0539) *(0.0533)* -0.1267 (0.0498)** *(0.0588)***	0.0312 (0.0669) *(0.0618)* -0.0900 (0.0506)* *(0.0508)**	-0.0036 (0.0846) *(0.0849)* -0.1188 (0.0635)* *(0.0739)*			
$EGE1_{it}$ $EGE1_{i,t-1}$	0.0456 (0.0318) *(0.0269)** -0.0655 (0.0317)** *(0.0263)***	0.0765 (0.0448)* *(0.0527)* -0.0659 (0.0351)* *(0.0282)***	0.0504 (0.0618) *(0.0590)* -0.1297 (0.0537)** *(0.0520)***	0.0511 (0.0337) *(0.0287)** -0.0664 (0.0336)** *(0.0282)***	0.0702 (0.0404)* *(0.0406)** -0.0429 (0.0319) *(0.0205)***	0.0439 (0.0559) *(0.0526)* -0.0587 (0.0464) *(0.0388)*	0.0525 (0.0331) *(0.0272)** -0.0715 (0.0332)** *(0.0300)***	0.0524 (0.0424) *(0.0369)* -0.0511 (0.0342) *(0.0252)***	0.0016 (0.0578) *(0.0601)* -0.0592 (0.0470) *(0.0480)*			
AGE_{it} $AGE_{i,t-1}$	0.0080 (0.0049)* *(0.0057)* -0.0011 (0.0030) *(0.0036)*	0.0013 (0.0050) *(0.0061)* -0.0070 (0.0027)** *(0.0032)***	0.0027 (0.0055) *(0.0070)* -0.0033 (0.0031) *(0.0035)*	0.0083 (0.0051) *(0.0055)* -0.0021 (0.0032) *(0.0036)*	0.0050 (0.0046) *(0.0053)* -0.0059 (0.0026)** *(0.0031)**	0.0088 (0.0054) *(0.0068)* -0.0005 (0.0031) *(0.0035)*	0.0108 (0.0053)** *(0.0056)** -0.0022 (0.0032) *(0.0037)*	0.0080 (0.0055) *(0.0062)* -0.0071 (0.0029)** *(0.0035)***	0.0113 (0.0063)* *(0.0075)* -0.0030 (0.0032) *(0.0035)*			
$LFSTOCK_{it}$ $LFSTOCK_{i,t}$												
$ECACRA_{it}$ $ECACRA_{i,t-1}$	-0.0078 (0.0022)*** *(0.0021)**** -0.0046 (0.0023)** *(0.0021)***	-0.0051 (0.0035) *(0.0036)* -0.0067 (0.0032)** *(0.0032)***	-0.0072 (0.0042)* *(0.0039)** -0.0082 (0.0046)* *(0.0050)*									
$UNEM_{it}$ $UNEM_{i,t-1}$				-0.0865 (0.2213) *(0.1836)* -0.3702 (0.2206)*	0.1723 (0.3225) *(0.3195)* 0.2074 (0.2431)	0.2386 (0.3890) *(0.3674)* 0.8445 (0.3645)**						

				_(0.2556)_	_(0.2703)_	_(0.2979)_***						
$INACTIVE_{it}$ $INACTIVE_{i,t}$							-0.4672 (0.1766)*** _(0.2104)_** 0.0567 (0.1394) _(0.1236)_	-0.6287 (0.3249)* _(0.3580)_* 0.2356 (0.1733) _(0.1577)_	-0.8120 (0.4393)* _(0.5851)_ -0.3325 (0.3420) _(0.3591)_			
$ECACRF_{it}$ $ECACRF_{i,t-1}$				-0.0048 (0.0020)** _(0.0020)_** -0.0056 (0.0021)*** _(0.0020)_***	-0.0043 (0.0026) _(0.0025)_* -0.0059 (0.0026)** _(0.0030)_**	-0.0066 (0.0034)** _(0.0032)_** -0.0033 (0.0040) _(0.0043)_	-0.0053 (0.0019)*** _(0.0021)_** -0.0052 (0.0020)** _(0.0019)_***	-0.0062 (0.0033)* _(0.0029)_** -0.0036 (0.0028) _(0.0030)_	-0.0132 (0.0047)*** _(0.0051)_** -0.0062 (0.0041) _(0.0044)_			
OBS.	325			325			325					
SARGAN TEST (p-value)	9.12 (0.4264)	58.44 (0.0378)	27.06 (0.3527)	8.71 (0.4644)	86.75 (0.0007)	36.89 (0.1491)	7.32 (0.6041)	64.35 (0.0696)	32.70 (0.2899)			
AR(1) TEST (p-value)	-4.93 (0.0000) _-3.51_ _(0.0005)_	-4.79 (0.0000) _-3.36_ _(0.0008)_	-4.09 (0.0000) _-2.92_ _(0.0035)_	-5.03 (0.0000) _-3.56_ _(0.0004)_	-4.93 (0.0000) _-3.22_ _(0.0013)_	-4.02 (0.0001) _-3.01_ _(0.0026)_	-5.20 (0.0000) _-3.79_ _(0.0002)_	-5.28 (0.0000) _-3.44_ _(0.0006)_	-2.99 (0.0028) _-2.31_ _(0.0210)_			
AR(2) TEST (p-value)	-0.87 (0.3866) _-0.50_ _(0.6168)_	-1.46 (0.1441) _-0.77_ _(0.4422)_	-1.36 (0.1723) _-0.76_ _(0.4443)_	-0.67 (0.5056) _-0.40_ _(0.6876)_	-1.66 (0.0960) _-0.92_ _(0.3583)_	-1.82 (0.0692) _-1.15_ _(0.2493)_	-0.65 (0.5181) _-0.39_ _(0.6996)_	-0.75 (0.4558) _-0.43_ _(0.6705)_	-1.36 (0.1752) _-0.95_ _(0.3415)_			

NOTES: (*), (**), and (***) indicates significance at the 10%, 5% and 1% level, respectively. _(*), (**)_, and _(***)_ denotes the significance of the White (1980) estimator (robust standard errors) at the 10%, 5% and 1% level, respectively. SARGAN TEST is the Sargan test for overidentifying restrictions (Sargan 1958). AR(1) TEST and AR(2) TEST are the Arellano-Bond test for the first and the second-order autocorrelation in the first differenced residuals, respectively. Time dummies and a constant are included.

Appendix A.9: Short Run GMM: Dependent variable is income inequality for normally working people (NGE1)

	REGRESSION 1			REGRESSION 2			REGRESSION 3			REGRESSION 4		
	(a) x_{it} strictly exogenous	(b) x_{it} predetermined	(c) x_{it} endogenous	(a) x_{it} strictly exogenous	(b) x_{it} predetermined	(c) x_{it} endogenous	(a) x_{it} strictly exogenous	(b) x_{it} predetermined	(c) x_{it} endogenous	(a) x_{it} strictly exogenous	(b) x_{it} predetermined	(c) x_{it} endogenous
$NGE1_{i,t-1}$	0.7220 (0.1354)*** _(0.1248)***_	0.8090 (0.1496)*** _(0.1637)***_	0.8863 (0.1751)*** _(0.1895)***_	0.8326 (0.1602)*** _(0.1553)***_	0.4428 (0.1134)*** _(0.1089)***_	0.4213 (0.1616)*** _(0.2213)*_	0.8360 (0.1600)*** _(0.1553)***_	0.4375 (0.1134)*** _(0.1064)***_	0.4019 (0.1615)** _(0.2165)*_	0.5717 (0.1233)*** _(0.1644)***_	0.3222 (0.0958)*** _(0.1240)***_	0.4142 (0.1248)*** _(0.1246)***_
NMN_{it} $NMN_{i,t-1}$	0.0061 (0.0022)*** _(0.0031)*_ -0.0009 (0.0025) _(0.0029)_	-0.0023 (0.0038) _(0.0051)_ -0.0055 (0.0061) _(0.0051)_	0.0038 (0.0061) _(0.0087)_ -0.0197 (0.0110)* _(0.0087)**_	0.0091 (0.0026)*** _(0.0037)**_ -0.0045 (0.0034) _(0.0045)_	0.0127 (0.0050)** _(0.0052)**_ -0.0057 (0.0058) _(0.0065)_	0.0147 (0.0070)** _(0.0064)**_ -0.0076 (0.0069) _(0.0078)_	0.0090 (0.0027)*** _(0.0038)**_ -0.0048 (0.0034) _(0.0046)_	0.0143 (0.0049)*** _(0.0048)***_ -0.0072 (0.0056) _(0.0059)_	0.0169 (0.0068)** _(0.0058)***_ -0.0087 (0.0066) _(0.0069)_	0.0094 (0.0030)*** _(0.0045)**_ -0.0066 (0.0030)** _(0.0043)_	0.0190 (0.0048)*** _(0.0063)***_ -0.0134 (0.0043)*** _(0.0054)**_	0.0171 (0.0056)*** _(0.0066)**_ -0.0140 (0.0054)** _(0.0061)**_
EMN_{it} $EMN_{i,t-1}$				0.0999 (0.0579)* _(0.0618)_ -0.1644 (0.0605)*** _(0.0555)***_	0.1921 (0.0788)** _(0.0943)**_ -0.1322 (0.0620)** _(0.0577)**_	0.2511 (0.1074)** _(0.1097)**_ -0.1241 (0.0894) _(0.1015)_	0.0974 (0.0579)* _(0.0608)_ -0.1670 (0.0611)*** _(0.0570)***_	0.2049 (0.0784)*** _(0.0931)**_ -0.1469 (0.0618)** _(0.0559)***_	0.2688 (0.1084)** _(0.1089)**_ -0.1400 (0.0888) _(0.0929)_	0.1371 (0.0523)*** _(0.0531)**_ -0.1621 (0.0525)*** _(0.0595)***_	0.2248 (0.0682)*** _(0.0766)***_ -0.2169 (0.0568)*** _(0.0627)***_	0.2318 (0.0840)*** _(0.0850)***_ -0.2059 (0.0738)*** _(0.0684)***_
$EGE1_{it}$ $EGE1_{i,t-1}$				0.0388 (0.0405) _(0.0299)_ -0.0855 (0.0422)** _(0.0302)***_	0.1078 (0.0516)** _(0.0414)***_ -0.0444 (0.0437) _(0.0295)_	0.1162 (0.0723) _(0.0635)*_ -0.0515 (0.0614) _(0.0640)_	0.0293 (0.0413) _(0.0286)_ -0.0863 (0.0423)** _(0.0301)***_	0.1045 (0.0528)** _(0.0413)**_ -0.0467 (0.0438) _(0.0288)_	0.1131 (0.0731) _(0.0627)*_ -0.0528 (0.0617) _(0.0627)_	0.0485 (0.0335) _(0.0253)*_ -0.0851 (0.0343)** _(0.0286)***_	0.0493 (0.0456) _(0.0374)_ -0.0675 (0.0372)* _(0.0288)**_	0.0602 (0.0602) _(0.0573)_ -0.0754 (0.0569) _(0.0456)*_
AGE_{it} $AGE_{i,t-1}$							0.0047 (0.0057) _(0.0058)_ 0.0033 (0.0040) _(0.0031)_	0.0056 (0.0052) _(0.0058)_ -0.0003 (0.0032) _(0.0027)_	0.0063 (0.0057) _(0.0063)_ 0.0005 (0.0034) _(0.0029)_	0.0092 (0.0052)* _(0.0058)_ 0.0026 (0.0033) _(0.0027)_	0.0115 (0.0052)** _(0.0059)*_ -0.0013 (0.0028) _(0.0027)_	0.0130 (0.0059)** _(0.0064)**_ 0.0004 (0.0031) _(0.0028)_
$ECACRF_{it}$ $ECACRF_{i,t-1}$										-0.0026 (0.0020) _(0.0020)_ -0.0073 (0.0021)*** _(0.0020)***_	-0.0025 (0.0030) _(0.0026)_ -0.0082 (0.0029)*** _(0.0032)**_	-0.0049 (0.0036) _(0.0034)_ -0.0036 (0.0046) _(0.0040)_
OBS.	400			392			392			325		
SARGAN TEST (p-value)	10.84 (0.2871)	16.09 (0.5175)	9.96 (0.6974)	8.88 (0.4484)	43.72 (0.1005)	38.10 (0.0126)	8.68 (0.4674)	42.85 (0.1170)	37.38 (0.0152)	4.75 (0.8557)	49.94 (0.1597)	26.57 (0.3776)
AR(1) TEST (p-value)	-5.57 (0.0000) _-4.78_ _(0.0000)_	-5.32 (0.0000) _-4.46_ _(0.0000)_	-5.16 (0.0000) _-4.48_ _(0.0000)_	-5.28 (0.0000) _-4.60_ _(0.0000)_	-5.07 (0.0000) _-4.46_ _(0.0000)_	-3.40 (0.0007) _-2.56_ _(0.0105)_	-5.32 (0.0000) _-4.58_ _(0.0000)_	-5.10 (0.0000) _-4.37_ _(0.0000)_	-3.30 (0.0010) _-2.59_ _(0.0095)_	-5.12 (0.0000) _-3.50_ _(0.0005)_	-5.24 (0.0000) _-3.35_ _(0.0004)_	-4.40 (0.0010) _-3.79_ _(0.0002)_
AR(2) TEST (p-value)	-1.79 (0.0739) _-1.07_ _(0.2851)_	-1.72 (0.0848) _-1.07_ _(0.2836)_	-1.44 (0.1500) _-0.99_ _(0.3234)_	-2.10 (0.0355) _-1.31_ _(0.1895)_	-2.95 (0.0032) _-1.65_ _(0.0988)_	-2.53 (0.0113) _-1.54_ _(0.1244)_	-2.04 (0.0411) _-1.26_ _(0.2077)_	-2.91 (0.0036) _-1.60_ _(0.1087)_	-2.46 (0.0140) _-1.47_ _(0.1429)_	-1.19 (0.2356) _-0.73_ _(0.4633)_	-0.76 (0.4468) _-0.57_ _(0.5656)_	-0.49 (0.6217) _-0.37_ _(0.7088)_

NOTES: (*), (**), and (***) indicates significance at the 10%, 5% and 1% level, respectively. _(*), (**), and (***)_ denotes the significance of the White (1980) estimator (robust standard errors) at the 10%, 5% and 1% level, respectively. SARGAN TEST is the Sargan test for overidentifying restrictions (Sargan 1958). AR(1) TEST and AR(2) TEST are the Arellano-Bond test for the first and the second-order autocorrelation in the first differenced residuals, respectively. Time dummies and a constant are included.

Appendix A.10: Long Run GMM: Dependent variable is income inequality for normally working people (NGE1)

	REGRESSION (1)			REGRESSION (2)			REGRESSION (3)			REGRESSION (4)		
	(a) X_{it} strictly exogenous	(b) X_{it} predetermined	(c) X_{it} endogenous	(a) X_{it} strictly exogenous	(b) X_{it} predetermined	(c) X_{it} endogenous	(a) X_{it} strictly exogenous	(b) X_{it} predetermined	(c) X_{it} endogenous	(a) X_{it} strictly exogenous	(b) X_{it} predetermined	(c) X_{it} endogenous
NMN	0.0186 (0.0107)* *(0.0118)*	-0.0408 (0.0530) *(0.0650)*	-0.1397 (0.2707) *(0.3017)*	0.0277 (0.0301) *(0.0338)*	0.0125 (0.0064)* *(0.0077)*	0.0123 (0.0088) *(0.0098)*	0.0256 (0.0293) *(0.0336)*	0.0126 (0.0065)* *(0.0078)*	0.0136 (0.0086) *(0.0096)*	0.0066 (0.0080) *(0.0079)*	0.0083 (0.0058) *(0.0057)*	0.0052 (0.0076) *(0.0073)*
EMN				-0.3854 (0.6791) *(0.7199)*	0.1074 (0.1253) *(0.1346)*	0.2195 (0.1791) *(0.1865)*	-0.4239 (0.7223) *(0.7517)*	0.1031 (0.1249) *(0.1355)*	0.2153 (0.1745) *(0.1786)*	-0.0583 (0.1520) *(0.1689)*	0.0116 (0.0913) *(0.1077)*	0.0443 (0.1355) *(0.1522)*
EGE1				-0.2789 (0.4984) *(0.4951)*	0.1138 (0.0823) *(0.0671)**	0.1118 (0.1202) *(0.1136)*	-0.3477 (0.5684) *(0.5574)*	0.1028 (0.0839) *(0.0673)*	0.1007 (0.1184) *(0.1074)*	-0.0854 (0.1114) *(0.1153)*	-0.0269 (0.0687) *(0.0699)*	-0.0259 (0.1066) *(0.1087)*
AGE							0.0487 (0.0651) *(0.0649)*	0.0095 (0.0106) *(0.0106)*	0.0113 (0.0111) *(0.0131)*	0.0274 (0.0171) *(0.0203)*	0.0151 (0.0093) *(0.0096)*	0.0229 (0.0135)* *(0.0131)**
ECACRF										-0.0232 (0.0091)** *(0.0127)**	-0.0159 (0.0052)*** *(0.0062)***	-0.0145 (0.0082)* *(0.0094)*
OBS.	400			392			392			325		

NOTES: (*), (**), and (***) indicates significance at the 10%, 5% and 1% level, respectively. *(*), (**), and (***)* denotes the significance of the White (1980) estimator (robust standard errors) at the 10%, 5% and 1% level, respectively.