SERC DISCUSSION PAPER 11

Education and Income Inequality in the Regions of the European Union

Andres Rodríguez-Pose (SERC, London School of Economics)
Vassilis Tselios (CURDS, Newcastle University)

November 2008

This work was part of the research programme of the independent UK Spatial Economics Research Centre funded by the Economic and Social Research Council (ESRC), Department for Business, Enterprise and Regulatory Reform (BERR), the Department for Communities and Local Government (CLG), and the Welsh Assembly Government. The support of the funders is acknowledged. The views expressed are those of the authors and do not represent the views of the funders.

Education and Income Inequality in the Regions of the European Union

Andrés Rodríguez-Pose* and Vassilis Tselios**

November 2008

* SERC, Department of Geography & Environment, London School of Economics
** Center for Urban and Regional Development Studies (CURDS), Newcastle University

Acknowledgements
The authors are grateful to the European Commission [DYNREG Programme, contract no 028818 (CIT5)].and Eurostat for granting access to the European Community Household Panel (ECHP). The work was also part of the research programme of the independent UK Spatial Economics Research Centre funded by the Economic and Social Research Council (ESRC), Department for Business, Enterprise and Regulatory Reform, Communities and Local Government, and the Welsh Assembly Government. The support of the funders is acknowledged. The views expressed are those of the authors and do not represent the views of the funders or of Eurostat. An early version of the paper was presented at the 46th Congress of the European Regional Science Association in Volos.

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 the relationship between income per capita and income inequality, as well as between a good human capital endowment and income inequality is positive. High levels of inequality in educational attainment are also associated with higher income inequality. The above results are robust to changes in the definition of income distribution and may be interpreted as a sign of the responsiveness of the EU labor market to differences in qualifications and skills. Other results indicate that population ageing, female participation in the labor force, urbanization, agriculture, and industry are negatively associated to income inequality, while unemployment and the presence of a strong financial sector positively affect 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

JEL Classifications: D31, I21, O15, O18

1. INTRODUCTION

It is often claimed that improvements in educational attainment affect income inequality (Berry and Glaeser, 2005; Shapiro, 2006) 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 variation in impact at different levels of education and of the positive correlation between inequality in education and in income for the regions of the EU. It aims 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. We measure human capital distribution in terms of both the percentage of the labor force which has received primary, secondary, or tertiary education and of 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 education policies contribute to a more equal income distribution and whether the EU labor market is responsive to differences in qualifications, knowledge, and skills.

The paper is organized 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 the whole array of sources, but simply to focus on how the impact of income per capita, as well as of primary, secondary, and tertiary education levels and inequality in educational achievement, 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 inequality.

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 (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 wealthier 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 differences are often viewed 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, pp. 14). 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), who posits that income per capita has an inverted U-curve effect on income inequality. Income inequality increases as nations begin to industrialize and, then, declines at later stages of industrialization. This relationship is known as ‘Kuznets curve’. The Kuznets curve shows that in the early stages of industrialization, the labor force is primarily engaged in agriculture. As industrialization 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 industrialization and labor 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 (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; Anand and Kanbur, 1993; 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. 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. Based on the work of Mincer (1958), Schultz (1961) and Becker (1962) income inequality is generally considered to be affected by educational attainment, in a process which is sometimes referred to as ‘skills deepening’ (Williamson, 1991). However, the impact of endowments at different levels of education (i.e. primary, secondary, and tertiary) seems to depend on a country’s level of development (Sianesi and Van Reenen, 2003), with tertiary education being the most important for the variation in income (Berry and Glaeser, 2005; Shapiro, 2006). 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 tertiary education, for example, is likely to increase the earning opportunity of the lowest strata, leading to a reduction in earning inequality (Checchi, 2000). Access to education is likely to provide for upward mobility and thus greater income equality. 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 different types of educational attainment on income inequalities depends on the balance between the ‘composition’ and the ‘wage compression’ effect. Concerning the ‘composition’ effect, an increase in tertiary education 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 tertiary 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, a rise in the educated labor supply is likely to increase competition for positions requiring advanced educational credentials and thereby should reduce the income differential between the more and the less educated (Tinbergen, 1975). 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) signaling 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 labeling 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 (1997), 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, and Wang (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 (i.e. Jacobs, 1985; Chakraborty and Das, 2005). More explicitly, Thorbecke and Charumilind (2002, pp. 1488) have pointed out that, with regard to the supply side of skilled labor 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 labor education, if the demand for unskilled labor is either contracting or growing at a slower rate than the demand for skilled labor, 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 labor force has an equalizing 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 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, and industrial composition. 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 $N \to \infty$ 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 to minimize potential problems of omitted-variable bias. We also use panel data in order to allow for greater degrees of freedom than with cross-regional data and to improve the accuracy of the parameter estimates (i.e. Baltagi, 2005).

This study deals with three methods of panel regression analysis: standard (non-spatial) static models, spatial (static) models, and dynamic models. These models are increasingly popular for panel data analysis among regional scientists. With repeated observations for a maximum 102 regions, panel analysis permits us to study the dynamics of change with short-time series. The basic characteristics of each method are presented below:

(1) The *standard static models* endow regression analysis with both a spatial and temporal dimension. The former dimension pertains to a set of cross-regional units of observation, while the latter to periodic observations of a set of variables characterizing these cross-regional units over a particular time span. As the surveys of the European Community Household Panel (ECHP) dataset – which is our main data source – were conducted regularly at approximately one-year intervals, the error terms of inequality regressions are expected to be correlated with the regional-specific effect. This can be addressed with fixed effects (FEs) panel data analyses. The static model is characterized by one source of persistence over time due to the presence of unobserved regional-specific effects. Based on the specification tests of Hausman's (1978) chi-squared statistic, and Breusch and Pagan's (1980) Lagrange multiplier (LM), FEs correct for unmeasured regional-invariant factors. In addition, as FEs techniques can lead to

misleading results when most of the variation is cross-sectional (Partridge, 2005) – in the case of the income distribution measures over the six-year period considered – the random effects (REs) are also reported. Both FEs and REs estimators are based on a strict exogeneity assumption.

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 disturbance term has a different variance, the robust estimation of the covariance matrix is presented following the White estimator for unspecified heteroskedasticity (White, 1980).

(2) The *spatial models* deal with substantive and nuisance spatial dependence induced by an exonesously determined weights matrix and provide a framework to test for the occurrence of interregional externalities (Rey and Montouri, 1999). We use two panel data models: the spatial autoregressive (SAR) and the spatial error (SE) model (Anselin, 1988). The SAR model indicates how income inequality in a region is affected by those of neighboring regions and is a substantive type. Income is likely to spill over across regions through interregional trade, transfer payments, network and social capital and pecuniary, technological, and information externalities. In the SE model, spatial dependence works through omitted variables and is a nuisance type. Both models are estimated by maximum likelihood (ML) (Elhorst, 2003). In the SAR model, the spatial autoregressive parameter indicates the extent of interregional interactions, and in the SE model, the spatial error parameter expresses the intensity of spatial correlation between

regression residuals. Since the question of the correct spatial specification is a very important one and there are no spatial diagnostic tests for panel data models, the selection of one of the two models is based on the significance of the coefficients, the value of the log-likelihood function, and the diagnostic tests for the spatial cross-sectional models, such as the Moran's I test adapted to estimated residuals (Cliff and Ord, 1981), the Lagrange multiplier test for residual spatial autocorrelation, as well as the Lagrange multiplier test for an additional residual spatial autocorrelation in the spatial autoregressive model (Anselin, 1988).

(3) The *dynamic models* test for the existence of autocorrelation. In these models, we can obtain both short-run and long-run parameters. However, the equilibrium, for instance, may be constrained in the short-run because of supply rigidities or factor immobilities that are removed in the longer-run (Combes, Duranton, and Overman, 2005). 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. Long-run standard errors are calculated using the Delta method. The dynamic panel structure of our data is exploited by a generalized method of moments (GMM) estimation (Arellano and Bond, 1991). 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 (Baltagi, 2005).

The dynamic model is characterized 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). FEs and REs estimators are likely to be biased and inconsistent, because the dynamic econometric model contains a lagged endogenous variable (Baltagi, 2005). The correlation between the explanatory variables and the error is handled by instrument variables. In GMM-DIFF estimations (Arellano and Bond, 1991), the 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. This procedure is more efficient than the Anderson and Hsiao (1981) two stage least squares estimator which does not make use of all of the available moment conditions (Ahn and Schmidt, 1995). The GMM-DIFF estimator may also be improved using the GMM-SYS estimator (Arellano and Bover, 1995; Blundell and Bond, 1998) which uses not only lagged levels of the instruments for equations in first differences, but also lagged differences as instruments for equations in levels. The GMM methodology is based on a set of diagnostics. The tests of overidentifying restrictions are associated with Sargan (1958) and Hansen (1982) statistics. They should not indicate correlation between the instruments and the error term. Additionally, the tests regarding serial correlation should reject the absence of first and second order serial correlation. Both the homoskedastic one-step and the robust one-step GMM-DIFF and GMM-SYS estimators are presented.

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. Broadly speaking, the advantage of dynamic over static models is that the former correct the inconsistency introduced by lagged endogenous variables and, also, permit a certain degree of endogeneity in the regressors. However, dynamic models do not deal with spatial dependence.

4. DATA AND VARIABLES

As in other recent studies dealing with human capital variables across European regions (Rodríguez-Pose and Vilalta-Bufi, 2005; Ezcurra, 2007), 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 from 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 13 countries in the EU.[1] 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 regionalized. Income is collected not only for each individual in the household, so as to measure income per capita and income inequality for the population as a whole, but also for each normally working (15+ hours/week) individual[2] in the household in order to measure income per capita and income inequality for normally working people. Income per capita is transformed for the same level of prices using the harmonized indices for consumer prices and then is divided by 1,000. 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

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

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

other social benefits or grants, and private income. Wages are the main source of personal income, and they constitute the 45 percent of the personal income of the whole of the population and the 78 percent of the personal income of normally working people (Rodríguez-Pose and Tselios, 2007).

Income inequality is calculated using the generalized Theil entropy index. 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. Income inequality within a region is defined as $\textit{Income Inequality} = \sum_{i=1}^{N} y_i \log(Ny_i)$, where y_i is the income share that is individual i's total income as a proportion of the total income for the entire regional population. This index varies from 0 for perfect equality to $\log N$ for perfect inequality.

The education variables 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: recognized third level education completed, second stage of secondary level education completed, and less than second stage of secondary level education completed. These categories, which are mutually exclusive, allow for international comparisons, because they are defined by the International Standard Classification of Education. We describe the educational attainment within a region in terms of the percentage of the population who have successfully achieved the above three levels of formal education in order to find which educational category is the critical factor in income inequality variations. For instance, the work of Berry and Glaeser (2005) and Shapiro (2006) indicate that tertiary education is critical in terms of spatial variations in earnings.

Following the work of Thomas, Wang, and Fan (2001), we also calculate the inequalities in educational attainment using an education Theil index. This is defined as $Educational\ Inequality = \sum_{i=1}^{N} z_i \log(Nz_i)$, where z_i is the human capital share, that is, individual i's higher education level completed as a proportion of the 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 effect of additional quantitative time-variant variables on income inequality: the average age of individuals, the percentage of normally working (15+ hours/week) respondents, the percentage of unemployed respondents, and the percentage of inactive respondents within a region. The source of these variables is again the ECHP dataset. Other controls include the economic activity rate of the population, female activity, and the added value per capita of agriculture, industry, and services from the Eurostat's Regio dataset. These are also time-variant indicators. The urbanization ratio of a region 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 (ECHP data source), and not for all countries. We assume that the urbanization ratio from 1995 to 2001 remains constant. This variable, therefore, introduces observed time-invariant effects.

The transformed dataset with means, standard deviation, and minimum and maximum value for each of the variables is reported in Table 1.[3] The descriptive statistics show that the dataset is unbalanced, which is amenable to estimation methods that manage potential

[3] The descriptive statistics of the ECHP quantitative and qualitative variables can be provided upon request.

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, while the percentage of respondents with tertiary education has increased. Mapping income and educational inequalities in 1995 and 2000 shows that (1) inequalities are not randomly distributed in space, highlighting the spatial autocorrelation in inequalities, and that (2) the spatial distribution of inequalities has remained relatively stable between 1995 and 2000, underscoring the persistence of inequalities (Appendix A.1).

TABLE 1: Summary Statistics

Variable	Definition	Year	Source	Obs	Mean or percent	Std. Dev.	Min	Max
Income inequality	Income inequality for the whole of the population (Theil index)	1995	ECHP	94	0.42	0.16	0.18	0.83
		2000		102	0.36	0.14	0.11	0.74
Income per capita	Income per capita for the whole of the population (/1000)	1995	ECHP	94	9.76	3.54	3.40	18.93
		2000		102	12.81	4.55	4.05	21.14
Income inequality	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
Income per capita	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
Primary	Percentage of respondents with less than second stage of secondary level education completed	1995	ECHP	94	53.60	17.34	14.44	90.26
		2000		102	45.54	17.59	11.51	85.95
Secondary	Percentage of respondents with second stage of secondary level education completed	1995	ECHP	94	27.29	16.58	7.25	63.34
		2000		102	28.44	18.35	7.98	68.23
Tertiary	Percentage of respondents with third level education completed	1995	ECHP	94	19.11	10.66	1.80	40.94
		2000		102	26.03	15.02	3.58	55.56
Educational inequality	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
Population ageing	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
Work access	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
Work access	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
Unemployment	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
Inactivity	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
Women's work access	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
Agriculture	Added value per capita of agriculture, hunting, forestry and fishing	1995	Eurostat	101	0.44	0.32	0.01	1.42
		2000		97	0.44	0.33	0.01	1.44
Industry	Added value per capita of mining and quarrying, manufacturing, electricity, gas and water supply, construction	1995	Eurostat	101	4.33	1.77	0.84	9.28
		2000		97	5.62	1.93	1.33	10.48
Services	Added value per capita of services (excluding extra-territorial organizations and bodies)	1995	Eurostat	101	10.05	5.06	3.64	33.77
		2000		97	14.41	5.92	5.12	38.71
Wholesale and retail trade	Added value per capita of wholesale and retail trade, repair of motor vehicles, motorcycles and	1995	Eurostat	85	3.06	1.28	1.20	9.13

	personal and household goods, hotels and restaurants, transport, storage and communication	2000		97	4.53	1.66	1.76	10.03
Finance	Added value per capita of financial intermediation, real estate, renting and business activities	1995	Eurostat	85	3.15	2.30	0.99	14.94
		2000		97	5.20	3.29	1.20	19.68
Public administration	Added value per capita of public administration and defense, compulsory social security; education; health and social work; other community, social and personal service activities; private households with employed persons	1995	Eurostat	85	3.09	1.35	1.17	9.70
		2000		97	4.68	1.42	1.93	11.09

Source: ECHP dataset and Eurostat's Regio dataset

The qualitative explanatory variables (time-invariant) classify regions into categories that are hypothesized 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, pp. 28).

- Religion: The European regions' religious affiliation is classified into four categories[4]: mainly Protestant (Sweden, Denmark, northern Germany, Scotland), mainly Catholic (France, Ireland, Luxembourg, Portugal, Spain, Italy, Austria,

[4] 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 parts of southern Germany, Belgium), mainly Anglican (England and Wales) and mainly Orthodox (Greece). It is hypothesized 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.

- 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 FEs and REs estimations in the standard static models, ML in the spatial models, and GMM estimation (both GMM-DIFF and GMM-SYS) in the dynamic models taking into account the unobserved regional-specific effects. We first report the static regression models, followed by the dynamic ones.[5]

Estimations of the Static Models

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

---

[5] In our study we have considered the differences between the sample and the population (Gelman, 2007). Our results for the inequality measures are, however, robust by region with and without weights. We, therefore, only report the regression results without weights. This may be the result of Eurostat's role in leading both the elaboration process of the survey design of the ECHP data set and of the Eurostat's region database, making comparisons reliable.

OLS models, and the p-values of Hausman's test reject the GLS estimator as an appropriate alternative to the FEs estimator. Therefore, 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. The determinants of income inequality are thus not sensitive to the model specification of the error term. Table 2 displays the FEs regression results, complemented by REs for those equations where time-invariant indicators are considered.

TABLE 2: FEs and REs Regression Results

	FEs					REs			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Income per capita	-0.0001 (0.0011) *(0.0013)*	0.0020 (0.0014) *(0.0016)*	0.0048 (0.0016)*** *(0.0017)****	0.0158 (0.0022)*** *(0.0025)****	0.0208 (0.0024)*** *(0.0027)****	-0.0007 (0.0016) *(0.0017)*	0.0026 (0.0013)* *(0.0014)**	0.0000 (0.0013) *(0.0015)*	0.0028 (0.0013)** *(0.0014)***
Secondary		0.3652 (0.0833)*** *(0.1222)****	0.2405 (0.0785)*** *(0.1027)***	0.2854 (0.0781)*** *(0.1023)****	0.2678 (0.0779)*** *(0.0928)****	0.3402 (0.1161)*** *(0.1904)**	0.1500 (0.0767)* *(0.1031)*	0.1789 (0.0796)** *(0.1281)*	0.2015 (0.0711)*** *(0.1014)***
Tertiary		0.2661 (0.0747)*** *(0.1127)***	0.1564 (0.0720)** *(0.0941)**	0.2497 (0.0710)*** *(0.0952)****	0.2492 (0.0705)*** *(0.0844)****	0.2751 (0.0851)*** *(0.1482)**	0.0905 (0.0688) *(0.0950)*	0.1813 (0.0720)** *(0.1135)*	0.1156 (0.0682)* *(0.0979)*
Educational inequality		0.1661 (0.0318)*** *(0.0506)****	0.1021 (0.0309)*** *(0.0422)***	0.1249 (0.0300)*** *(0.0423)****	0.1064 (0.0300)*** *(0.0378)****	0.1563 (0.0343)*** *(0.0650)***	0.0880 (0.0290)*** *(0.0430)***	0.1222 (0.0303)*** *(0.0521)***	0.1015 (0.0284)*** *(0.0437)***
Population ageing			-0.0056 (0.0022)*** *(0.0024)***						
Unemployment			0.5325 (0.1391)*** *(0.1482)****						
Women's work access			-0.0063 (0.0012)*** *(0.0013)****						
Agriculture				-0.0941 (0.0336)*** *(0.0274)****	-0.0773 (0.0338)** *(0.0262)****				
Industry				-0.0262 (0.0048)*** *(0.0056)****	-0.0231 (0.0050)*** *(0.0058)****				
Services				-0.0068 (0.0019)*** *(0.0019)****					
Wholesale and retail trade					-0.0267 (0.0100)*** *(0.0104)***				
Finance					0.0082 (0.0043)* *(0.0044)**				
Public administration					-0.0270 (0.0088)*** *(0.0105)***				
Urbanisation						-0.2530 (0.0500)*** *(0.0512)****			
Welfare regime							X1		
Religion								X2	
Family structure									X3
Observations	604	596	513	586	566	378	596	596	596
R-within	0.0000	0.0648	0.1933	0.1922	0.2350				
LM test (p-value)	916.46 (0.0000)	662.92 (0.0000)	654.90 (0.0000)	478.98 (0.0000)	388.13 (0.0000)	431.91 (0.0000)	847.77 (0.0000)	819.77 (0.0000)	925.07 (0.0000)

Hausman test (p-value)	71.46 (0.0000)	66.30 (0.0000)	53.98 (0.0000)	101.31 (0.0000)	202.09 (0.0000)				

Notes: (*), (**), and (***) indicates significance at the 10 percent, 5 percent and 1 percent 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. The models contain constant. X1, X2, and X3 indicate the presence of welfare regime, religion and family structure dummies, respectively. The full set of results can be provided upon request.

Regression 1 analyses the impact of income per capita on income inequality. This equation is unconditioned by any other effects. The relationship between income per capita and inequality is negative, but statistically insignificant. 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, this suggests a poor unconditioned model. In the FEs conditional regressions (Regressions 3–5) income per capita becomes 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 5 shows that an increase of one per cent in income per capita is associated with, on average, about 0.0208 per cent more income inequality, as measured by the Theil index.

The next step in the analysis sees the introduction of human capital distribution. Considering primary education level completed as our base category, we include the percentage of respondents with secondary and tertiary education, as well as the within-region educational inequality. The regression coefficients indicate that both secondary and tertiary education influence the resulting income distribution. The relationship is positive, robust, and statistically significant. The higher the secondary and the tertiary educational attainment, the higher the income inequality, with secondary education normally having a greater sway on the variation in income inequality, as its coefficient is higher than the coefficient on tertiary education. The empirical results also show that a highly unequal distribution of education level completed is associated with higher income inequality. This relationship is robust and statistically significant.

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. Our results are in line with Dickey's (2007) view that income inequality increases with the level of education, but clash with many of those reported earlier that point to education as a powerful instrument in reducing inequality (i.e. Checchi, 2000; World Bank, 2002). Another potential explanation is that the demand for unskilled labor grows at a slower rate than the demand for skilled labor. This positive relationship may also be a sign of the responsiveness of the EU labor market to differences in qualifications and skills.

The remaining regressions include the control variables described earlier. 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 labor force and young people in work earn less in a European labor market that traditionally rewards seniority, increasing the inequality levels within a society (Higgins and Williamson, 1999). In order to capture the economic activity characteristics of the regions, unemployment and women's participation in employment are also included in Regression 3. 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). The coefficient on the female economic activity rate is negative and significant. The impact of the increase in women's access to work has been to reduce income inequality.[6]

Regression 4 controls for sectoral composition. An increase in the added value per capita of agriculture, industry, and services is associated with a decrease in inequality. However, decomposing the service sector into wholesale and retail trade, finance, and public administration (Regression 5), highlights how different sub-sectors have a different association with income inequality. Whereas a greater emphasis on wholesale and retail trade and on public administration is negatively associated with inequality, a specialization in finance leads to greater income polarization.

The FEs estimator is not provided for the time-invariant controls as there is no within-group variation in these variables. Hence for Regressions 6-9, we display the REs results of the impact of urbanization and institutional variables on income inequality. Regression 6 reports the negative correlation between urbanization and inequality. Considering Kuznets' assumption that urbanization 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 rejects Estudillo's (1997) and Sassen's (2001) hypothesis that the heterogeneity of urban areas enhances, rather than lowers, inequality. Highly-urbanized regions seem not only to be more prosperous — the

---

[6] The work access variables measured by the percentage of normally working respondents (source: ECHP) and the economic activity rate of the total population (source: Eurostat) are negatively associated with income inequality and are statistically significant, while the coefficient on inactivity is not statistically significant. These results can be provided upon request.

correlation between income per capita and urbanization is positive (0.46) — but also less unequal.

Regression 7 checks for the influence of welfare regimes. 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 lead to 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 minimizes 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 'residual' ones.

Regression 8 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 9). Regions with a Nordic family structure are the most egalitarian societies and Southern/Catholic regions have the highest inequality.

The regression results of income inequality for normally working people are similar to the regression results of income inequality for the population as a whole, apart from the coefficients on population ageing, agriculture, and services which are not statistically significant.[7] More specifically, income per capita is positively associated with income inequality. Once more, the impact of secondary and tertiary educational achievement, as well as of educational inequality on income inequality is positive, robust, and statistically significant. Finally, income inequality for normally working people is, once again, lower in social-democratic welfare states, in mainly Protestant areas, and in regions with Nordic family structures.

Table 3 displays the estimation results by ML for the SAR and SE models for panel data. Results were obtained for the 3- and 5-nearest neighbors. First of all, the Moran's I test (Cliff and Ord, 1981) adapted to estimated residuals suggests spatial dependence. While the value of the log-likelihood function is slightly higher for the SE than the SAR models, the significance of the coefficients is higher for the SAR. In addition, the robust version of the Lagrange multiplier test for spatially lagged endogenous variable rejects the null hypothesis of no spatial dependence, but the robust version of this test for residual spatial autocorrelation rejects it (Anselin and Florax, 1995; Anselin et al., 1996).[8] Hence, the SAR is the most favored specification. This specification shows positive and statistically significant coefficients on income per capita, secondary education, tertiary education, and educational inequality, which are consistent with the non-spatial regression models of Table 2. Moreover, a spatial spillover effect is found, as the average income inequality within a given region is influenced by those of neighboring regions. The above results are robust to the choice of the spatial weights matrix.

---

[7] These results can be provided upon request.

[8] These results can be provided upon request.

TABLE 3: ML Regression Results

	SAR model				SE model			
	3 nearest neighbours		5 nearest neighbours		3 nearest neighbours		5 nearest neighbours	
	Spatial fixed effects	Spatial and time period fixed effects	Spatial fixed effects	Spatial and time period fixed effects	Spatial fixed effects	Spatial and time period fixed effects	Spatial fixed effects	Spatial and time period fixed effects
Income per capita	0.0025 (2.0694)**	0.0110 (7.4369)***	0.0027 (2.2678)**	0.0105 (6.9906)***	0.0065 (3.8820)***	0.0139 (8.5526)***	0.0107 (5.5390)***	0.0153 (8.6921)***
Secondary	0.2632 (3.5967)***	0.2032 (2.9591)***	0.2337 (3.2796)***	0.1863 (2.7393)***	0.1889 (2.3261)**	0.1600 (2.2027)**	0.0990 (1.2260)	0.1114 (1.5086)
Tertiary	0.1722 (2.6310)***	0.1347 (2.1540)**	0.1359 (2.1342)**	0.1117 (1.8041)*	0.0371 (0.4836)	0.0732 (1.0750)	-0.1042 (-1.3062)	0.0039 (0.0554)
Educational inequality	0.1221 (4.3626)***	0.1059 (4.0328)***	0.1065 (3.9048)***	0.0966 (3.7085)***	0.0951 (3.0769)***	0.0911 (3.2907)***	0.0601 (1.9539)*	0.0719 (2.5538)**
Spatial error parameter					0.3490 (8.6628)***	0.1960 (4.3499)***	0.5260 (13.1743)***	0.3190 (6.3295)***
Spatial autoregressive parameter	0.3050 (6.9716)***	0.1240 (2.5822)***	0.4120 (8.6282)***	0.2150 (3.9069)***				
R-squared	0.9566	0.9627	0.9589	0.9635	0.9572	0.9634	0.9609	0.9646
Log-Likelihood	1162.6891	1212.342	1175.4238	1216.5509	1163.6524	1215.5082	1181.4173	1222.1353
Observations	564	564	564	564	564	564	564	564

Notes: (*), (**), and (***) indicates significance at the 10 percent, 5 percent and 1 percent level, respectively.

Estimations of the Dynamic Models

Table 4 presents the short-run and long-run results of the dynamic models of income inequality for the population as a whole. The first column of each model specification assumes that the explanatory variables are strictly exogenous, while in the second column the explanatory variables are endogenous. This table also reports the tests statistics of serial correlation and overidentifying restrictions.

TABLE 4: GMM Regression Results

	GMM-DIFF				GMM-SYS			
	1		2		3		4	
	exogenous	endogenous	exogenous	endogenous	exogenous	endogenous	exogenous	endogenous
Lagged income inequality	0.7531 (0.1234)*** _(0.1199)_***	0.6965 (0.1451)*** _(0.1525)_***	0.8834 (0.1439)*** _(0.1517)_***	0.5191 (0.1361)*** _(0.1801)_***	0.7640 (0.1128)*** _(0.0939)_***	0.6680 (0.0898)*** _(0.0946)_***	0.8703 (0.1370)*** _(0.1349)_***	0.5049 (0.0801)*** _(0.1037)_***
Income per capita	0.0139 (0.0026)*** _(0.0027)_***	0.0132 (0.0042)*** _(0.0050)_***	0.0173 (0.0032)*** _(0.0033)_***	0.0258 (0.0057)*** _(0.0061)_***	0.0126 (0.0024)*** _(0.0026)_***	0.0116 (0.0027)*** _(0.0031)_***	0.0153 (0.0029)*** _(0.0032)_***	0.0166 (0.0027)*** _(0.0034)_***
Lagged income per capita	-0.0057 (0.0031)* _(0.0032)_*	-0.0017 (0.0065) _(0.0045)_	-0.0106 (0.0045)** _(0.0047)_**	-0.0138 (0.0065)** _(0.0072)_*	-0.0058 (0.0031)* _(0.0031)_*	-0.0065 (0.0025)** _(0.0027)_**	-0.0096 (0.0044)** _(0.0047)_**	-0.0117 (0.0037)*** _(0.0046)_**
Secondary			0.1831 (0.1230) _(0.1180)_	0.2948 (0.1944) _(0.1603)_*			0.1572 (0.1209) _(0.1189)_	0.1734 (0.1025)* _(0.1025)_*
Lagged secondary			-0.1389 (0.1332) _(0.0787)_*	0.1262 (0.2037) _(0.1609)_			-0.1207 (0.1303) _(0.0781)_	-0.0097 (0.1006) _(0.0701)_
Tertiary			0.2288 (0.1159)** _(0.1110)_**	0.4721 (0.1847)** _(0.1844)_**			0.1963 (0.1139)* _(0.1119)_*	0.3433 (0.0952)*** _(0.1111)_***
Lagged tertiary			-0.2530 (0.1165)** _(0.0862)_***	-0.1325 (0.1539) _(0.1193)_			-0.2196 (0.1142)* _(0.0864)_**	-0.0944 (0.0842) _(0.0628)_
Educational inequality			0.0850 (0.0454)* _(0.0399)_**	0.0907 (0.0686) _(0.0538)_*			0.0740 (0.0446)* _(0.0403)_*	0.0968 (0.0329)*** _(0.0367)_***

Lagged educational inequality			-0.0684 (0.0480) *(0.0248)****	0.0148 (0.0668) *(0.0461)*			-0.0639 (0.0471) *(0.0246)****	-0.0114 (0.0349) *(0.0216)*
Observations	400	400	392	392	400	400	392	392
Sargan test (p-value) *Hansen test* *(p-value)*	12.26 (0.1989)	18.09 (0.1541)	11.41 (0.2485)	30.35 (0.2116)	10.54 (0.309) *8.50* *(0.485)*	57.96 (0.000) *37.36* *(0.015)*	9.68 (0.377) *7.42* *(0.593)*	102.53 (0.000) *63.90* *(0.247)*
AR(1) test (p-value)	-5.85 (0.0000) *-4.42* *(0.0000)*	-4.82 (0.0000) *-4.09* *(0.0000)*	-5.59 (0.0000) *-3.85* *(0.0001)*	-4.88 (0.0000) *-3.31* *(0.0009)*	-5.96 (0.000) *-4.42* *(0.000)*	-6.60 (0.000) *-4.72* *(0.000)*	-5.68 (0.000) *-3.78* *(0.000)*	-6.83 (0.000) *-4.26* *(0.000)*
AR(2) test (p-value)	-1.19 (0.2339) *-0.68* *(0.4977)*	-1.14 (0.2562) *-0.65* *(0.5188)*	-1.50 (0.1332) *-0.88* *(0.3774)*	-1.32 (0.1853) *-0.87* *(0.3865)*	-1.22 (0.224) *-0.68* *(0.494)*	-1.27 (0.204) *-0.71* *(0.480)*	-1.51 (0.131) *-0.86* *(0.391)*	-1.28 (0.202) *-0.72* *(0.471)*
Long-run parameters								
Income per capita	0.0331 (0.0137)** *(0.0143)***	0.0377 (0.0136)*** *(0.0151)***	0.0577 (0.0681) *(0.0784)*	0.0251 (0.0107)** *(0.0137)**	0.0285 (0.0131)** *(0.0140)***	0.0154 (0.0073)** *(0.0068)***	0.0436 (0.0478) *(0.0561)*	0.0099 (0.0062) *(0.0069)*
Secondary			0.3786 (1.3684) *(1.1513)*	0.8754 (0.4303)** *(0.4506)**			0.2810 (1.2010) *(0.9881)*	0.3306 (0.2391) *(0.1972)**
Tertiary			-0.2079 (1.4621) *(1.4002)*	0.7062 (0.3506)** *(0.3872)**			-0.1797 (1.2731) *(1.2060)*	0.5029 (0.2258)** *(0.2372)***
Educational inequality			0.1420 (0.5254) *(0.4217)*	0.2194 (0.1225)* *(0.1016)***			0.0785 (0.4723) *(0.3737)*	0.1726 (0.0810)** *(0.0779)***

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

Overall, the specification tests are satisfactory. The Sargan tests do not indicate correlation between the instruments and the error term of the first differenced equation, because they do not reject the overidentifying restrictions, except for the GMM-SYS estimators which assume that the explanatory variables are endogenous. The Hansen tests also do not reject the overidentifying restrictions, apart from Regression 3 and assuming that the explanatory variables are endogenous. The tests for serial correlation reject the absence of first order, but not second order serial correlation.[9]

All the equations reject that the lagged income inequality coefficient is zero. In both GMM-DIFF and GMM-SYS estimators, the coefficient on the lagged dependent variable

[9] We have also contemplated the possibility of weak instruments in the GMM estimation. Weak instruments correspond to a weak identification of some or all of the unknown parameters which may result in GMM statistics with nonnormal distributions, leading to the possibility of misleading conventional GMM inferences (Stock, Wright, and Yogo, 2002). Our results are robust to experimentation with different lag lengths, allowing us in all likelihood to discard the possibility of weak instruments.

is positive and statistically significant at the one per cent level, and it is higher when the explanatory variables are assumed to be exogenous than endogenous. Hence, 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: 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).

The short-run coefficient on income per capita is positive and statistically significant, regardless of the explanatory variables considered. In addition, the coefficients on secondary education, tertiary education, and educational inequality are positive, as in the case of the FEs regression results which also capture the short-run effects (Mairesse, 1990). The reason why some lagged educational variables are not significant may be that the time series variation in these variables is limited.

Considering the long-run parameters, the results indicate that income inequality increases in the long-run as income per capita increases, thus leading to a positive correlation between the two variables. For instance, if the endogenous income is increased by one per cent, income inequality will rise by 0.0377 per cent in the long-run for the GMM-DIFF estimator and 0.0154 per cent in the long-run for the GMM-SYS estimator (Regressions 1 and 3, respectively). This goes against the assumption of the presence of a declining segment of the Kuznets curve, but also fails 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 the higher the secondary education, the tertiary education, and the educational inequality, the higher the income inequality in the long-run, but only when the explanatory variables are assumed to be endogenous. According to the estimated value and assuming, for example, that income and human capital variables are endogenous, a one per cent increase in the coefficient on tertiary education would lead in the long-run to a 0.7062 per cent increase in income inequality for the GMM-DIFF estimator and a 0.5029 per cent increase for the GMM-SYS (Regressions 2 and 4, respectively). Once more, secondary education has the strongest association with the variation in income inequality. The combined positive impact of educational attainment and inequality on income inequality implies that, although educational expansion improves the opportunities for individuals, the returns tend to be 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 further indicates a responsiveness of the EU labor 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, pp. 296).

6. CONCLUDING REMARKS

Different static and dynamic panel data analyses have been conducted in order to examine how microeconomic changes in educational distribution in terms of both the percentage of the labor force that has received primary, secondary, or tertiary education and inequality in educational achievement, 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.

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 more the occupational choices and the earning opportunities of the rich, rather than of the population as a whole. The short-run and long-run impact of secondary and tertiary education on income inequality is positive with secondary education having the strongest association with inequality. There is also a positive and robust relationship between educational inequality and income inequality. Other results indicate that population ageing, female participation in the labor force, urbanization, agriculture and industry are negatively associated to income inequality, while unemployment and a specialization in the financial sector positively affect inequality. Finally, income inequality is lower in social-democratic welfare states, in Protestant areas, and in regions with Nordic family structures.

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 secondary and tertiary education relative to primary education and providing higher skills 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 labor 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 addresses the question of how changes in income per capita, educational attainment, and educational inequality affect the observed income inequality, further research is needed. 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. The classification of individuals into just three educational attainment categories represents a further simplification and limitation.

Despite these caveats, the results of the paper have contributed to shed light on the complex relationship between education and inequality, with results that, in some cases, tend to challenge the dominant views. From this perspective, it raises interesting questions that future research will need to address.

REFERENCES

Ahluwalia, Montek S. 1976. "Inequality, Poverty and Development," *Journal of Development Economics*, 3, 307-342.

Ahn, Seung C. and Peter Schmidt. 1995. "Efficient Estimation of Models for Dynamic Panel Data," *Journal of Econometrics*, 68, 5-27.

Anand, Sudhir and S. M. Ravi Kanbur. 1993. "The Kuznets Process and the Inequality Development Relationship," *Journal of Development Economics*, 40, 25-52.

Anderson, Theodore W. and Cheng Hsiao. 1981. "Estimation of Dynamic Models with Error Components," *Journal of the American Statistical Association*, 76, 598-606.

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht; Boston: Kluwer Academic Publishers.

Anselin, Luc, Anil K. Bera, Raymond J.G.M. Florax and Mann J. Yoon. 1996. "Simple Diagnostic Tests for Spatial Dependence," *Regional Science and Urban Economics*, 26, 77-104.

Anselin, Luc and Raymond J.G.M. Florax. 1995. "Small Sample Properties of Tests for Spatial Dependence in Regression Models," in L. Anselin and R. J. G. M. Florax (eds),*New Directions in Spatial Econometrics* Berlin; New York: Springer.

Arellano, Manuel and Stephen Bond. 1991. "Some Tests of Specification for Panel Data - Monte-Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58, 277-297.

Arellano, Manuel and Olympia Bover. 1995. "Another Look at the Instrumental Variable Estimation of Error-Components Models," *Journal of Econometrics*, 68, 29-51.

Baltagi, Badi H. 2005. *Econometric Analysis of Panel Data*. Chichester: John Wiley.

Barr, Nicholas A. 2004. *The Economics of the Welfare State*. Oxford: Oxford University Press.

Becker, Gary S. 1962. "Investment in Human Capital - a Theoretical Analysis," *Journal of Political Economy*, 70, 9-49.

Becker, Gary S. and Barry R. Chiswick. 1966. "Economics of Education and Distribution of Earnings," *American Economic Review*, 56, 358-369.

Berry, Christopher R. and Edward L. Glaeser. 2005. "The Divergence of Human Capital Levels across Cities," *Papers in Regional Science*, 84, 407-444.

Berthoud, Richard and Maria Iacovou. 2004. *Social Europe: Living Standards and Welfare States*. Cheltenham, UK, Northampton, MA: Edward Elgar.

Blundell, Richard and Stephen Bond. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, 115-143.

Breusch, Trevor S. and Adrian R. Pagan. 1980. "The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics," *Review of Economic Studies*, 47, 239-253.

Chakraborty, Shankha and Mausumi Das. 2005. "Mortality, Human Capital and Persistent Inequality," *Journal of Economic Growth*, 10, 159-192.

Champernowne, David G. and Frank A. Cowell. 1998. *Economic Inequality and Income Distribution*. Cambridge [England]; New York: Cambridge University Press.

Checchi, Daniele. 2000. "Does Educational Achievement Help to Explain Income Inequality?," Departmental Working Papers 2000-11, Department of Economics, University of Milan, Italy.

Cliff, Andrew D. and J. Keith Ord. 1981. *Spatial Processes: Models & Applications*. London: Pion.

Combes, Pierre-Philippe, Gilles Duranton and Henry G. Overman. 2005. "Agglomeration and the Adjustment of the Spatial Economy," *Papers in Regional Science*, 84, 311-349.

Cooper, Suzanne J. 1998. "Redistribution and the Persistence of Income Inequality." Working Paper, John F. Kennedy School of Government, Harvard University.

De Gregorio, José and Jong-Wha Lee. 2002. "Education and Income Inequality: New Evidence from Cross-Country Data," *Review of Income and Wealth*, 395-416.

Dickey, Heather. 2007. "Regional Earnings Inequality in Great Britain: Evidence from Quantile Regressions," *Journal of Regional Science*, 47, 775-806.

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

Elhorst, J. Paul. 2003. "Specification and Estimation of Spatial Panel Data Models," *International Regional Science Review*, 26, 244-268.

Esping-Andersen, Gøsta. 1990. *The Three Worlds of Welfare Capitalism*. Cambridge: Polity.

Estudillo, Jonna P. 1997. "Income Inequality in the Philippines, 1961-91," *Developing Economies*, 35, 68-95.

Ezcurra, Roberto. 2007. "Is Income Inequality Harmful for Regional Growth? Evidence from the European Union," *Urban Studies*, 44, 1953-1971.

Ferrera, Maurizio. 1996. "The “Southern Model” of Welfare in Social Europe," *Journal of European Social Policy*, 6, 17-37.

Firebaugh, Glenn. 2003. *The New Geography of Global Income Inequality*. Cambridge, MA: Harvard University Press.

Galor, Oded and Daniel Tsiddon. 1997. "Technological Progress, Mobility, and Economic Growth," *American Economic Review*, 87, 363-382.

Gelman, Andrew. 2007. "Struggles with Survey Weighting and Regression Modeling," *Statistical Science*, 22, 153-164.

Graham, Carol. 2002. "Mobility, Opportunity, and Vulnerability: The Dynamics of Poverty and Inequality in a Global Economy," *Journal of Human Development*, 3, 57-94.

Gujarati, Damodar N. 2003. *Basic Econometrics*. Boston, London: McGraw-Hill.

Hansen, Lars Peter. 1982. "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029-1054.

Hausman, Jerry A. 1978. "Specification Tests in Econometrics," *Econometrica*, 46, 1251-1271.

Higgins, Matthew and Jeffrey Williamson. 1999. "Explaining Inequality the World Round: Cohort Size, Kuznets Curves, and Openness," NBER Working Paper No. 7224, National Bureau of Economic Research.

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

Knight, John B. and Richard H. Sabot. 1983. "Educational Expansion and the Kuznets Effect," *American Economic Review*, 73, 1132-1136.

Kuznets, Simon. 1955. "Economic Growth and Income Inequality," *American Economic Review*, 45, 1-28.

López, Ramon E., Vinod Thomas and Yan Wang. 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, Harold. 1979. *A Theory of Income Distribution*. Oxford: Clarendon Press.

Mairesse, Jacques. 1990. "Time-Series and Cross-Sectional Estimates on Panel Data: Why Are They Different and Why They Should Be Equal," in J. Hartog, G. Ridder and J. Theeuwes (eds.),*Panel Data and Labor Market Studies* New York: North Holland. pp. 81-95.

Mincer, Jacob. 1958. "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy*, 66, 281-302.

Mocan, H. Naci. 1999. "Structural Unemployment, Cyclical Unemployment, and Income Inequality," *Review of Economics and Statistics*, 81, 122-134.

Motonishi, Taizo. 2006. "Why Has Income Inequality in Thailand Increased? An Analysis Using Surveys from 1975 to 1998," *Japan and the World Economy*, 18, 464-487.

Nielsen, François and Arthur S. Alderson. 1997. "The Kuznets Curve and the Great U-Turn: Income Inequality in U.S Countries, 1970 to 1990," *American Sociological Review*, 62, 12-33.

Park, Kang H. 1996. "Educational Expansion and Educational Inequality on Income Distribution," *Economics of Education Review*, 15, 51-58.

Partridge, Mark D. 2005. "Does Income Distribution Affect U.S State Economic Growth?," *Journal of Regional Science*, 45, 363-394.

Rey, Sergio J. and Brett D. Montouri. 1999. "U.S Regional Income Convergence: A Spatial Econometric Perspective," *Regional Studies*, 33, 143-156.

Rodríguez-Pose, Andrés and Vassilis Tselios. 2007. "Mapping Regional Personal Income Distribution in Western Europe: Income Per Capita and Inequality," Working Paper, Dynamic Regions in a Knowledge-Driven Global Economy: Lessons and Policy Implications for the EU.

Rodríguez-Pose, Andrés and Montserrat Vilalta-Bufi. 2005. "Education, Migration, and Job Satisfaction: The Regional Returns of Human Capital in the E.U," *Journal of Economic Geography*, 5, 545-566.

Sargan, John D. 1958. "The Estimation of Economic Relationships Using Instrumental Variables," *Econometrica*, 26, 393-415.

Sassen, Saskia. 2001. *The Global City: New York, London, Tokyo*. Princeton: Princeton University Press.

Schultz, Theodore William. 1961. "Investment in Human Capital," *American Economic Review*, 51, 1-17.

———. 1962. "Reflections on Investment in Man," *Journal of Political Economy*, 70, 1-8.

Shapiro, Jesse M. 2006. "Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital," *Review of Economics and Statistics*, 88, 324-335.

Sianesi, Barbara and John Van Reenen. 2003. "The Returns to Education: Macroeconomics," *Journal of Economic Surveys*, 17, 157-200.

Spence, A. Michael. 1973. "Job Market Signaling," *Quarterly Journal of Economics*, 87, 355-374.

Stock, James H., Jonathan H. Wright and Motohiro Yogo. 2002. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business & Economic Statistics*, 20, 518-529.

Thomas, Vinod, Yan Wang and Xibo Fan. 2001. *Measuring Education Inequality: Gini Coefficients of Education*. Washington, D.C.: World Bank.

Thorbecke, Erik and Chutatong Charumilind. 2002. "Economic Inequality and Its Socioeconomic Impact," *World Development*, 30, 1477-1495.

Tinbergen, Jan. 1975. *Income Distribution: Analysis and Policies*. Amsterdam; Oxford: North Holland Publishing Co. [etc.].

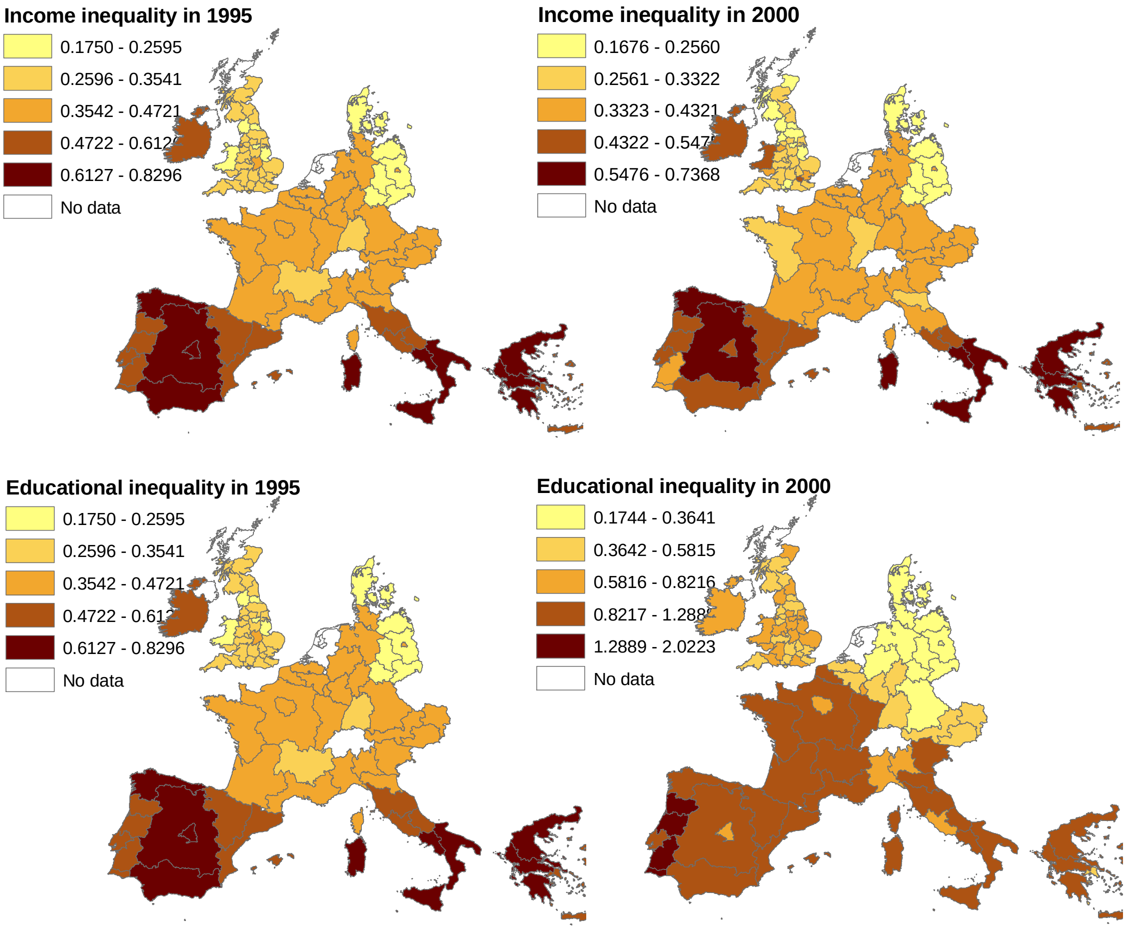
White, Halbert. 1980. "A Heteroskedasticity Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48, 817-838.

Williamson, Jeffrey G. 1991. *Inequality, Poverty, and History: The Kuznets Memorial Lectures of the Economic Growth Center, Yale University*. Oxford: Basil Blackwell.

Wolf, Alison. 2004. "Education and Economic Performance: Simplistic Theories and Their Policy Consequences," *Oxford Review of Economic Policy*, 20, 315-333.

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: The Spatial Distribution of Income and Educational Inequalities (Theil Index)



Spatial Economics Research Centre (SERC)
London School of Economics
Houghton Street
London WC2A 2AE

Tel: 020 7852 3565
Fax: 020 7955 6848
Web: www.spatialeconomics.ac.uk

SERC is an independent research centre funded by the Economic and Social Research Council (ESRC), Department for Business, Enterprise and Regulatory Reform (BERR), the Department for Communities and Local Government (CLG) and the Welsh Assembly Government.