XXII CONFERENZA

NUOVE FRONTIERE DELL'INTERVENTO PUBBLICO
IN UN MONDO DI INTERDIPENDENZA

Pavia, Università, 20-21 settembre 2010

A STATISTICAL EXPLORATION IN THE AGE GROUP INEQUALITY AROUND THE WORLD

EMANUELE CANEGRATI

società italiana di economia pubblica

dipartimento di economia pubblica e territoriale – università di Pavia

A Statistical Exploration in the Age Group Inequality Around the World

November 24, 2009

Abstract

This paper intends to measure the age group income inequality and the impact of public expenditure and income taxation-related policies on cohort inequality. Age group Gini indexes are calculated from the Luxembourg Income Study. Different hypotheses on standard errors are considered, in order to detect the presence of one-way or two-way fixed effects at different levels of clustering. Results are very robust in demonstrating that fiscal policies do influence age group inequality. Nevertheless, the coefficient signs change according to the underlying hypothesis on the shape of the standard errors and their interpretation is ambiguous.

Keywords: Age group inequality, Multilevel analysis, Public Expenditure, Income Taxation

JEL Classification Codes: E62, E64, H24

Contents

1 **Introduction** **3**

2 **Age group income inequality** **4**
2.1 Distributional and Ranking concepts 4
2.2 Decomposition indices 6

3 **Empirical evidence** **7**
3.1 LIS Dataset 7
3.2 Variables of interest 8
3.2.1 Sources of data 10
3.3 Econometric analysis 10
3.3.1 Non-parametric Analysis 10
3.3.2 Regression analysis 12

4 **Analysis of results** **15**
4.1 Summary statistics 15
4.2 Age group non-parametric distribution 16
4.3 Regression results 17

5 **Conclusions** **17**

List of Tables

1 Age group Gini indexes 24
2 Age group Gini indexes 26
3 Combined One-sample Kolmogorov-Smirnov test 28
4 Kruskal-Wallis equality-of-populations rank test - Country 29
5 Kruskal-Wallis equality-of-populations rank test - Age Group . . 30
6 One-way Error Component Model (I) 31
7 One-way Error Component Model (II) 32
8 One-way Error Component Model (III) 33
9 Two-way Error Component Model 34
10 3-Level Mixed Effects (I) 35
11 3-Level Mixed Effects (II) 36
12 3-Level Mixed Effects (III) 37
13 2-Level Mixed Effects (I) 38
14 2-Level Mixed Effects (II) 39
15 2-Level Mixed Effects (III) 40

1 Introduction

The analysis of the causes of income inequality is one of the most studied topics in the economic science. Nowadays, all modern democracies use fiscal policies in order to achieve redistribution goals and to reduce the level of disparity among social groups. From a normative perspective, a common tenet, taken from the optimal theory of taxation, affirms that a fair income distribution may be achieved through a tax system where income tax, paid as a fraction of before-tax income, increases somewhat with income (Atkinson, 1970).

Nevertheless, despite statutory schedules are revised from time to time by policy-makers, the stylized facts show that in Britain and America, "from the 1970s to the 1990s inequality rose in both countries" and that "redistribution toward the poor tends to happen least in those times and polities where it would seem most justified by the usual goal of welfare policy" (Lindert, 2000). Other evidence, showing an increasing level of inequality within industrialized countries, is found by Gottschalk and Smeeding, 2000. Finally, a comprehensive study made by the United Nations (WIDER, 2000) demonstrates that a recent increase in inequality has taken place in several countries such as Australia, United Kingdom, United States, Chile, Peru, Bangladesh, China, Philippines and Poland. As a result, it seems that redistribution and equity goals are far from being reached even in more industrialized countries and a natural question arises: is the design of fiscal policies effective in reducing income inequality?

A great amount of literature studies the impact of fiscal policies on economic indicators. For example, Fatás and Mihov, 2001, demonstrate the existence of a significant impact of fiscal policies on consumption and employment, while Giavazzi et al., 1999, detects the presence of casual effects on savings. On the contrary, no significant effects are detected between fiscal policies and output (Tsoukalas, 2008).

Oddly enough, the impact of fiscal policies on income distribution is less studied, although there are some authors who have started to fill this gap. For instance, an empirical study by Afonso et al. (2008) adopts a non-parametric approach to assess the efficiency of public spending in promoting more equalization of income in OECD countries and finds that redistributive public spending has a significant effect on income distribution. Over the last years, the literature has been focusing on the government redistribution between ages and between birth cohorts, but only focusing on the analysis of the transfer of resources from the old to the poor, or vice versa (for a good review of the literature see Eschker, 2007 and Altonji et al., 1995)

This paper analyzes the impact of public expenditure and income taxation on age group inequality for seventeen countries. Age group Gini index is calculated by using data taken from the Luxemburg Income Study (LIS). To the best of my knowledge, this is the first study that calculates age group inequality indices and assesses the impact of fiscal policies on the generations' welfare.

The econometric framework is designed in order to allow for different hypotheses on standard errors, in order to detect the presence of idiosyncratic components among clusters of data. The use of mixed effects, obtained by clustering data by country and age group, is certainly not standard in the econometric literature of public policies. Results are very robust in demonstrating that both public expenditure components and income taxation are able to influence the level of age group inequality, but they also show that it is not easy to understand whether fiscal policies reduce or increase the level of inequality.

The paper is organized as follows: section two introduces some basic concepts and tools used in the income inequality measurement and focuses on age group inequality. Section 3 describes the database based on the Luxemburg Income Study and the econometric technique used. Section 4 describes the main results and section 5 concludes.

2 Age group income inequality

The goal of this section is threefold: 1) introducing some useful measuring tools to measure age group income inequality, 2) calculating Gini indices at a cohort level and 3) assessing if this inequality is caused by the structure of fiscal policies, in particular fiscal expenditure components and tax policy.

The main research question addressed in this study is "how much are the age groups afflicted by the different components of public expenditure?" Before answering this question we must remember that inequality measurement is always an attempt to give meaning to comparisons of income distributions in terms of criteria that may be derived from ethical principles, appealing mathematical constructs or simple intuition (Cowell, 2000). Consequently, before measuring the level of inequality in practise, it is necessary to define the concepts, the ranking criteria, and the indices necessary to achieve our goal.

2.1 Distributional and Ranking concepts

I will denote by $\digamma$ the space of all univariate probability distributions with support $\Lambda \subseteq \mathbb{R}; x \in \Lambda$ represents a particular value of income and $\mathcal{F} \in \digamma$ one of the possible income distribution. So $\mathcal{F}(x \leq \widetilde{x})$ represents the proportion of population with income less than $\widetilde{x}$. Furthermore define $\underline{x} := \inf(\Lambda)$ and denote by $\mathcal{F}(\varrho) \subseteq \digamma$ a subset with given mean $\varrho : \digamma \mapsto \mathbb{R}$ given by

$$\varrho(\mathcal{F}) := \int x d\mathcal{F}(x) \tag{1}$$

and $f : \Lambda' \mapsto \mathbb{R}$ as a density function, supposed that $\mathcal{F}$ is continuous over some intervals $\Lambda' \subseteq \Lambda$. Furthermore, in order to compare distributions, I assume the existence of a complete and transitive binary relation $\succcurlyeq_I$on $\digamma$, called *inequality ordering* and represented by $I : \mathcal{F} \mapsto \mathbb{R}$, if the ordering is continuous.[1]

[1] I assume that axioms of *Anonymity*, *Population Principle*, *Principle of Transfers*, *Monotonicity*, *Scale Invariance*, *Decomposability*, *Uniform income growth* and *Translation Invariance* (Cowell, 2000) are satisfied.

In order to compare distributions we also need to define some ranking criteria over $\digamma$. I use the notation $\succcurlyeq_{\mathcal{T}}$to indicate the *ranking* induced by a comparison principle $\mathcal{T}$. Three possible situations arise:

Definition 1 *For all $\mathcal{F}, \mathcal{G} \in \digamma$:*

(a) (strict dominance) $\mathcal{G} \succ_{\mathcal{T}} \mathcal{F} \Leftrightarrow \mathcal{G} \succcurlyeq_{\mathcal{T}} \mathcal{F} \quad \wedge \quad \mathcal{F} / \succcurlyeq_{\mathcal{T}} \mathcal{G}$.
(b) (equivalence) $\mathcal{G} \sim_{\mathcal{T}} \mathcal{F} \Leftrightarrow \mathcal{G} \succcurlyeq_{\mathcal{T}} \mathcal{F} \quad \wedge \quad \mathcal{F} \succcurlyeq_{\mathcal{T}} \mathcal{G}$.
(c) (non-comparability) $\mathcal{G} \perp_{\mathcal{T}} \quad \mathcal{F} \Leftrightarrow \mathcal{G} / \succcurlyeq_{\mathcal{T}} \mathcal{F} \quad \wedge \quad \mathcal{F} / \succcurlyeq_{\mathcal{T}} \mathcal{G}$.

Suppose now to focus on the concept of Social-Welfare Function (SWF), expressed in the following additively separable form:

$$\mathcal{W}(\mathcal{F}) = \int \mathcal{U}(x)\, d\mathcal{F}(x) \tag{2}$$

where $\mathcal{U} : \digamma \mapsto \mathbb{R}$ is an evaluation function. Denote by $\widehat{\mathcal{W}_1}$ the subclass of SWFs where $\mathcal{U}$ is increasing and by $\widehat{\mathcal{W}_2}$ the subclass of $\widehat{\mathcal{W}_1}$ where $\mathcal{U}$ is also concave. Furthermore, define the set of age years $\mathcal{A}$ where $\mathfrak{a}$ is a given age in $\mathcal{A}$. Finally, introduce the following

Definition 2 *For all $\mathcal{F} \in \digamma$, $\mathfrak{a} \in \mathcal{A}$ and for all $0 \leq \mathfrak{q} \leq 1$, the quantile functional for a given age year is defined by*

$$\mathcal{Q}(\mathcal{F}; (\mathfrak{q}, \mathfrak{a})) = \inf \{x | \mathcal{F}(x) \geq \mathfrak{q}, \mathfrak{a}\} = x_{\mathfrak{q}\mathfrak{a}} \tag{3}$$

This definition enables us to state the theorem of the *first-order distributional dominance*

Theorem 3 $\mathcal{G} \succcurlyeq_{\mathcal{Q}} \mathcal{F} \Leftrightarrow \mathcal{W}(\mathcal{G}) \geq \mathcal{W}(\mathcal{F}) \; \forall \left(\mathcal{W} \in \widehat{\mathcal{W}_1}\right)$

Otherwise, if we consider this other

Definition 4 *For all $\mathcal{F} \in \digamma$, $\mathfrak{a} \in \mathcal{A}$ and for all $0 \leq \mathfrak{q} \leq 1$, the cumulative income functional for a given age year is defined by*

$$\mathcal{C}(\mathcal{F}; (\mathfrak{q}, \mathfrak{a})) := \int_{x}^{Q(\mathcal{F}; (\mathfrak{q}, \mathfrak{a}))} x d\mathcal{F}(x) \tag{4}$$

2

we can write the theorem of the *second-order distributional dominance*

Theorem 5 $\forall \mathcal{F}, \mathcal{G} \in \digamma \; (\varrho) : \mathcal{G} \succcurlyeq_{\mathcal{C}} \mathcal{F} \Leftrightarrow \mathcal{W}(\mathcal{G}) \geq \mathcal{W}(\mathcal{F}) \; \forall \left(\mathcal{W} \in \widehat{\mathcal{W}_2}\right)$

[2] The graph $C(F; q)$ against q describes the *generalised Lorenz curve.*

Suppose now that a distribution depends on the effects of a policy $\mathfrak{p} \in \mathcal{P}$, where $\mathcal{P}$ is the space of all the possible policies. Without loss of generality, I suppose that $\mathcal{P} = \left\{\mathfrak{p}^1, \mathfrak{p}^2\right\}$. Suppose also that distribution $\mathcal{F}$ is generated under policy $\mathfrak{p}^1$ and distribution $\mathcal{G}$ is generated under policy $\mathfrak{p}^2$. We denote by $\mathcal{F} = F\left(\mathfrak{p}^1, \mathfrak{a}\right)$ and $\mathcal{G} = G\left(\mathfrak{p}^2, \mathfrak{a}\right)$ the distributions obtained under the two policies for a given age group $\mathfrak{a}$.

We want to define a comparison criterion for judging policies and their effects on the distribution of age groups.

Theorem 6 *(First-order distributional dominance) For all* $\mathfrak{p}^1, \mathfrak{p}^2 \in \mathcal{P}$, $\mathfrak{a} \in \mathcal{A}$: $\mathfrak{p}^1 \succcurlyeq_{\mathcal{Q}} \mathfrak{p}^2 \Leftrightarrow \mathcal{W}\left(\mathcal{F}\left(\mathfrak{p}^1, \mathfrak{a}\right)\right) \geq \mathcal{W}\left(\mathcal{G}\left(\mathfrak{p}^2, \mathfrak{a}\right)\right) \; \forall \left(\mathcal{W} \in \widehat{\mathcal{W}_1}\right)$

Theorem 7 *(Second-order distributional dominance) For all* $\mathfrak{p}^1, \mathfrak{p}^2 \in \mathcal{P}$, $\mathfrak{a} \in \mathcal{A}$, $\mathcal{F}, \mathcal{G} \in F\;(\varrho) : \mathfrak{p}^1 \succcurlyeq_{\mathcal{C}} \mathfrak{p}^2 \Leftrightarrow \mathcal{W}\left(\mathcal{F}\left(\mathfrak{p}^1, \mathfrak{a}\right)\right) \geq \mathcal{W}\left(\mathcal{G}\left(\mathfrak{p}^2, \mathfrak{a}\right)\right) \; \forall \left(\mathcal{W} \in \widehat{\mathcal{W}_2}\right)$

These two theorems simply state that a policy $\mathfrak{q}^1$ is preferred to policy $\mathfrak{q}^2$ if and only if the welfare obtained under the distribution it generates is higher than the welfare obtained under the distribution generated by the other policy for every age group. Notice that this condition must hold *for every age group*; that means that we should see an improvement in welfare of all cohorts.

2.2 Decomposition indices

The *Generalized Entropy* measure is the more suitable index to analyze inequality within and between groups because of its decomposability. It may be written as

$$GE(\alpha) = \overbrace{\int_h f^h \left(\frac{x_h}{x}\right)^{\alpha} \mathcal{I}_h(\alpha)}^{within-group\ inequality} + \overbrace{\mathcal{I}_{bet}(\alpha)}^{between-group\ inequality} \tag{5}$$

where

$$\mathcal{I}_{bet}(\alpha) = \frac{1}{\alpha(\alpha-1)} \left[\int_h f^h \left(\frac{x_h}{x}\right)^{\alpha} - 1\right] \tag{6}$$

The α in 5 is a parameter that characterizes different members of the GE class: a high positive value of α yields an index that is very sensitive to income transfers at the top of the distribution. In particular, $GE(0)$ represents the mean logarithmic deviation, $GE(1)$ the Theil index, and $GE(2)$ the half of square of the coefficient of variation.

Another useful indicator to measure the inequality between groups is represented by Gini:

$$G = 1 + \frac{1}{\mathcal{N}} - \left[\frac{2}{\mathcal{N}^2 x}\right] \left[\int_h (\mathcal{N} - h + 1)\, x_h\right] \tag{7}$$

where $\mathcal{N} = \int w_h$, $w_h = f^h \mathcal{N}$. When data are unweighted, $w_h = 1$ and $\mathcal{N} = \mathcal{H}$. Individuals are ranked in ascending order of h.

3 Empirical evidence

3.1 LIS Dataset

The Luxembourg Income Study (LIS) is a panel database including 30 countries and made by 5 *waves* of data from 1979 – 2002[3]. The source of data is represented by country specific household income surveys. For example, individual data from the United States are taken from the *Current Population Survey*. Datasets are identified by a code made by two letters denoting a country and two numbers which identify the wave of data. For instance, US00 identifies the wave 2000 for the United States. In the analysis I used a reduced panel of 15 countries (letters in brackets represent the LIS codes): Austria(AT), Belgium(BE), Canada(CA), Switzerland(CH), Germany(DE), Denmark(DK), Spain(ES), Greece(GR), Ireland(IE), Italy(IT), Luxembourg(LU), Mexico(MX), Norway(NO), Sweden(SE) ,and United States(US).

The dataset includes data at both an individual and household level on demographics, expenditure, income, labor market outcomes and tax variables. Inequality indexes were calculated using the definition of *disposable income*, calculated as follows:

disposable income = compensation of employees
+ gross self
−employment income
+ realised property income
+ occupational pensions[4]
+ other cash income[5]
+ social insurance cash transfers[6]
+ universal cash transfers[7]
+ social assistance[8]
− direct taxes
− social security contributions.

[3] A new Wave VI is being released at the moment this paper is written.

[4] Occupational pensions include all pensions paid from non-social retirement schemes including employer-based pensions for private sector workers and public employees.

[5] Other cash income includes regular private transfers, alimony and child support benefits, other sources of regular cash income, not classified above.

[6] Social insurance transfers include: accident or short-term disability pay, long-term disability pay, social retirement benefits (old age and survivors), unemployment pay, maternity allowances, military or veteran's benefis, other social insurance.

[7] Universal cash transfers include child and/or family allowances if paid directly by governments. Universal cash transfers paid as refundable income tax credits are counted as negative amounts in the income tax of some countries.

[8] Social assistance includes all income-tested and means-tested benefits, both cash and near-cash.

This choice is natural because the disposable income allows us to assess the impact of taxation on individuals' welfare and thus to evaluate the degree of inequality as a result of the candidates' choice.

3.2 Variables of interest

In order to evaluate if and how the cohort-specific inequality depends upon the structure of fiscal policies chosen by the government, I build an econometric model where Gini indexes, calculated for every age group, represent the dependent variable. The regressors are represented by some variables which capture the two sides of fiscal policies, namely the tax structure and the public expenditure components (expressed in percentage of GDP), plus two macroeconomic control variables, the GDP growth rate and the Consumer Price Index (CPI).

Since it is not easy to evaluate how long the effects of a fiscal policy take to affect individuals' wealth (due to the so-called *transmission lag effect*), I take the values of public expenditure components measured in two different years, 1990 and 1995. Otherwise, we may reasonably assume that taxation affects directly and instantaneously the welfare of households, which means that the transmission lag is particularly low.

I use different proxies to capture the two sides of fiscal policy. The variables, with relative components, are reported below:

1. Public Expenditure Components

- *Old-age*[9]

  Cash benefits
  - Pensions
  - Early retirement pensions
  - Other cash benefits

  Benefits in kind
  - Residential care / Home-help services
  - Other benefits in kind

- *Health*[10]

[9] more in details: comprises all cash expenditures (including lump-sum payments) on old-age pensions. Old-age cash benefits provide an income for persons retired from the labour market or guarantee incomes when a person has reached a 'standard' pensionable age or fulfilled the necessary contributory requirements. This category also includes early retirement pensions: pensions paid before the beneficiary has reached the 'standard' pensionable age relevant to the programme. Excluded are programmes concerning early retirement for labour market reasons which are classified under unemployment. Old-age includes supplements for dependants paid to old-age pensioners with dependants under old-age cash benefits. Old age also includes social expenditure on services for the elderly people, services such as day care and rehabilitation services, home-help services and other benefits in kind. It also includes expenditure on the provision of residential care in an institution (for example, the cost of operating homes for the elderly).

[10] Social expenditure data in the health policy area is taken from the *OECD Health Data* (OECD, 2006). All public expenditure on health is included (not total health expenditure): current expenditure on health (personal and collective services and investment). Expenditure

Spending on in- and out-patient care
Medical goods
Prevention

- *Family*[11]

Cash benefits
Family allowances
Maternity and parental leave
Other cash benefits
Benefits in kind
Day-care / Home-care services
Other benefits in kind

- *Active labour programmes*[12]

Employment service and administration
Labour market training
Youth measures
Subsided employment
Employment measures for disabled

- *Housing*[13]

Employment measures for disabled
Housing assistance
Other benefits in kind

2. Taxation

in this category encompasses, among other things, expenditure on in-patient care, ambulatory medical services and pharmaceutical goods. Individual health expenditure, insofar as it is not reimbursed by a public institution, is not included. Cash benefits related to sickness are recorded under sickness benefits. Voluntary private social health expenditure are estimates on the benefits to recipients that derive from private health plans which contain an element of redistribution (such private health insurance plan are often employment-based and/or tax-advantaged).

[11] Includes expenditure which supports families (i.e. excluding one-person households). This expenditure is often related to the costs associated with raising children or with the support of other dependants. Expenditure related to maternity and parental leave is grouped under the family cash benefits sub-category

[12] Contains all social expenditure (other than education) which is aimed at the improvement of the beneficiaries' prospect of finding gainful employment or to otherwise increase their earnings capacity. This category includes spending on public employment services and administration, labour market training, special programmes for youth when in transition from school to work, labour market programmes to provide or promote employment for unemployed and other persons (excluding young and disabled persons) and special programmes for the disabled.

[13] Spending items recorded under this heading include rent subsidies and other benefits to the individual to help with housing costs. This includes direct public subsidies to tenants (in some countries, e.g. Norway, homeowners living in their house) 'earmarked' for support with the cost of housing.

- *attw67* - total tax wedge as a 67% of average wage; average personal income tax and social security contribution rates on gross labour income;
- *attw100* - total tax wedge as a 100% of average wage; average personal income tax and social security contribution rates on gross labour income;
- *attw133* - total tax wedge as a 133% of average wage; average personal income tax and social security contribution rates on gross labour income;
- *attw167* - total tax wedge as a 167% of average wage; average personal income tax and social security contribution rates on gross labour income;

The average tax rates 'all-in' for employees include personal income tax and employee social security contributions, less cash benefits, for a single individual without children at different income levels. They measure how much total net income after tax changes if an individual decides to join (or exit from) the labour market (OECD, 2004).I also consider two macroeconomic variables:

- *GDP Growth Rate*;
- *Consumer price index (CPI)*

calculated for years 1997, 1998 and1999.

3.2.1 Sources of data

For all OECD countries data on public expenditure on health and public expenditure on active labour market policies (ALMPs) are taken from the *OECD Health Data* and the *OECD database on Labour Market Programmes*, respectively (OECD, 2006a, and 2006b, Statistical Annex). Data on unemployment compensation (cash transfers) are taken from the *LMP database* for OECD countries that do not belong to the EU and from *ESSPROS* for EU countries. For some Non-European OECD countries, data delivered by through the services of the delegates to the Working party on Social Policy of the Employment Labour and Social Affairs committee. For some European countries data on social expenditure is provided by EUROSTAT as based on the information in their ESSPROS database (EUROSTAT, 2006). Data for tax rates are derived from the *OECD Taxing Wages* framework. GDP Growth rate and CPI are taken from EUROSTAT database.

3.3 Econometric analysis

3.3.1 Non-parametric Analysis

Kernel Density Estimation Kernel Density Estimation (Silverman, 1986 and Simonoff, 1996) intends to give a shape to the distribution of the age group Gini index for each country. Kernel estimators smooth out the contribution of each observed data point over a local neighborhood of that data point . Data point x_i contributes to the estimate at point $\widehat{x}$, depending on how apart x_i and $\widehat{x}$ are. The extent of this contribution depends on two factors: the shape of the Kernel function chosen and its bandwidth. The estimated density may be written as:

$$\widehat{\beta} = \frac{1}{n}\sum_{i=1}^{n} \mathcal{K}\left(\frac{\widehat{x} - x_i}{j}\right) \tag{8}$$

where $\mathcal{K}$ is a Kernel function, j the bandwidth and $\widehat{x}$ the point where the density is evaluated. The parabolic-shaped Epanechnikov Kernel

$$\mathcal{K}\left[\mathfrak{z}\right] = \begin{cases} \frac{3}{4}\left(1 - \frac{1}{5}\mathfrak{z}^2\right)/5 & if \ \ |\mathfrak{z}| < \sqrt{5} \\ \quad 0 & otherwise \end{cases} \tag{9}$$

is the Kernel function I use, since it is the most efficient in minimizing the asymptotic mean integrated squared error (Wand and Jones, 1995). Notice that the choice of j will decide how many values are included in estimating the density at each point and in this model is determined as

$$\mathfrak{m} = \min\left(\sqrt{variance_{\overline{x}}}, \frac{interquartile\ range_{\overline{x}}}{1.349}\right) \tag{10}$$

so that the bandwidth dimension is equal to

$$j = \frac{0.9\mathfrak{m}}{n^{\frac{1}{5}}} \tag{11}$$

where $\overline{x}$ is the variable for which the Kernel is estimated and n the number of observations.

One-sample Kolmogorov-Smirnov test The one-sample Kolmogorov-Smirnov (K-S) tests the equality of two distributions (Chakravarti et al., 1967). In our case we test, for each country, if the empirical cumulative density function of the sample is equal to a (assumed) normal cumulative density function.

To illustrate the test, I denote by X_i the random variable representing the Gini index of the sample and by $X_{\mathcal{N}}$ the values obtainable if the observations of the sample were distributed as normal. Furthermore, I denote by $\mathcal{F}(X_i)$ the *empirical distribution function* of the sample, defined as

$$\mathcal{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n}\mathcal{I}\left(X_i \leq x\right) \tag{12}$$

and by $\mathcal{F}(X_{\mathcal{N}})$ the *normal cumulative density function.* The K-S statistic for $\mathcal{F}(X_{\mathcal{N}})$ is

$$\mathcal{S}_n = \sup_x \left|\mathcal{F}(X_n) - \mathcal{F}(X_{\mathcal{N}})\right| \tag{13}$$

Under the null hypothesis that the sample comes from the normal distribution $\mathcal{F}(X_{\mathcal{N}})$, $\sqrt{n}\mathcal{S}_n$ converges at the limit to the Kolmogorov distribution, that is

$$\sqrt{n}\mathcal{S}_n \overset{n \longrightarrow \infty}{\longrightarrow} \sup |\mathcal{BF}(t)| \tag{14}$$

where $\mathcal{B}(t)$ is a given Brownian bridge.

Kruskall-Wallis Rank Test The Kruskall-Wallis Rank Test (Kruskal and Wallis, 1952) is a non-parametric econometric technique for testing equality of population medians among groups. The test assumes the existence of an identically-shaped and scaled distribution for each group, except for any difference in medians. Suppose a sample size divided in k groups and to rank the sample. Then compute $\mathcal{R}_i$ as the sum of the ranks for group i. Then the Kruskal Wallis test statistic is:

$$KW = \frac{12}{n(n+1)} \times \sum_{i=1}^{k} \frac{\mathcal{R}_i^2}{n_i} - 3(n+1) \tag{15}$$

This statistic approximates a *chi-square distribution* with $k-1$ degrees of freedom if the null hypothesis of equal populations is true. We reject the null hypothesis of equal population means if the test statistic KW is greater than $CHIPPF(\alpha, K-1)$, where $CHIPPF$ is the Chi-square Percent Point Function.

3.3.2 Regression analysis

One-way Error Component Model (OECM) We start to write the basic econometric regression model as

$$GINI_{ij} = \alpha + X_{ij}\beta + \varepsilon_{ij} \tag{16}$$

where $GINI_{ij}$ denotes the Gini index for the i-th age group and for the j-th country, calculated for year 2000. OLS regressions were performed by using different hypotheses on the shape of error components. Following Baltagi, 2008, I assume that observations could have unobserved fixed effects. First, we assume that residuals consist of an age group specific component, μ_i, and an idiosyncratic component, unique to each observation, v_{ij}, independent and identically distributed $IID\left(0, \sigma_v^2\right)$. That is, the error components may be written as

$$\varepsilon_{ij} = \mu_i + v_{ij} \qquad i = 1, ..., N; \quad j = 1, ..., M \tag{17}$$

Otherwise, following the same reasoning, we assume that residuals consist of a country specific component, η_j

$$\varepsilon_{ij} = \eta_j + v_{ij} \qquad i = 1, ..., N; \quad j = 1, ..., M \tag{18}$$

This structure of residuals produces *White standard errors*, which are robust to *within cluster correlation* (*Clustered* or *Rogers* standard errors). These residuals are correlated across observations of the same cohort, but are independent across countries:

$$corr\left(\varepsilon_{ij},\varepsilon_{ts}\right)=\begin{cases}1 & for\ i=t\ and\ j=s\\ \rho_{\varepsilon}=\frac{\sigma_{\mu}^{2}}{\sigma_{\varepsilon}^{2}} & for\ i=t\ and\ j\neq s\\ 0 & for\ i\neq t\end{cases} \tag{19}$$

or, alternatively, correlated across observations of the same country, but independent across cohorts:

$$corr\left(\varepsilon_{ij},\varepsilon_{ts}\right)=\begin{cases}1 & for\ i=t\ and\ j=s\\ \rho_{\varepsilon}=\frac{\sigma_{\eta}^{2}}{\sigma_{\varepsilon}^{2}} & for\ i\neq t\ and\ j=s\\ 0 & for\ j\neq s\end{cases} \tag{20}$$

Note that if $\rho_{\varepsilon}>0$ the OLS standard errors underestimate the true standard errors. It can be demonstrated (Petersen, 2006) that clustered standard errors are designed to correct the correlation of the residuals within cluster.

Two-way Error Component Model (TECM) Unlike the OECM, the TECM assumes the existence of both a cohort and a country specific component, μ_i and η_j respectively, so that the standard errors can be written as

$$\varepsilon_{ij}=\mu_i+\eta_j+v_{ij} \qquad i=1,...,N; \qquad j=1,...,M \tag{21}$$

This approach allows for correlations among different age groups in the same country and different countries in the same age group. Therefore, we have the following structure:

$$corr\left(\varepsilon_{ij},\varepsilon_{ts}\right)=\begin{cases}1 & for\ i=t\ and\ j=s\\ \rho_{\varepsilon}=\begin{cases}\frac{\sigma_{\mu}^{2}}{\sigma_{\varepsilon}^{2}} & for\ i=t\ and\ \forall\ j\neq s\\ \frac{\sigma_{\eta}^{2}}{\sigma_{\varepsilon}^{2}} & for\ i\neq t\ and\ \forall\ j=s\end{cases} & \\ 0 & for\ \forall\ j\neq s\ and\ \forall\ j\neq s\end{cases} \tag{22}$$

The Linear Multilevel Model (LMM) Multilevel models suppose the existence of a hierarchy of data in a sample, consisting of observations grouped at different levels (Goldstein, 1995 and Hox, 1995). This technique is very used in Biometrics and Medical science, while it is almost unknown in Economics.

The Linear Multilevel Model analyses the structure of data at *different levels of variation*. In this econometric framework, we can observe a level-1 variation, that is the variation existing among individuals. Then, we can also observe level-2 variation, that is the variation existing, for example, among generations *or* countries, and the level-3 variation, that is the variation among generations *and* countries.

The 2-Level model The specification of the 2-Level model is the following:

$$gini_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + \ldots + \beta_{kj}x_{ij} + \varepsilon_{ij} \tag{23}$$

where subscript i refers to the level-1 unit, the age group, and j to the level-2 unit, the country; the regression coefficients are supposed to have a random component

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \varepsilon_{0j} \\ \beta_{1j} &= \gamma_{10} + \varepsilon_{1j} \\ &\vdots \\ \beta_{kj} &= \gamma_{k0} + \varepsilon_{kj} \end{aligned} \tag{24}$$

Subscript ij indicates that an item varies from age group to age group within a country. The subscript j denotes that it varies from country to country but it does not vary among the taxpayers of the same country. Finally, when an item has neither an ij subscript nor a j subscript, it means that it is constant across all age groups and countries. Linear random intercept model allows random variation in intercept for countries; in other words, slope coefficients can also be random.

The 3-level model I analyze a model where the age group Gini indexes are considered at the age groups and countries within age groups. That is, countries are the second level of the analysis and age groups are the third level. Therefore, the specification of the 3-Level model is the following:

$$gini_{ijw} = \beta_0 + \beta_1 x_{ijw} + \ldots + \beta_{kj}x_{ijw} + (v_w + u_{jw} + \varepsilon_{ijw}) \tag{25}$$

where $w$denotes the new level of clustering. The model assumes the variance-covariance structure of the within-equation random effects as unstructured, that is, covariances allow all variances and covariances to be distinct[14].

These techniques of analysis have a serious drawback, which is the presence of multicollinearity among variables from the random-effects equation. The matrix singularity caused by the collinearity usually causes the estimation to fail. This is why it is important to use a small number of regressors and it is preferable to substitute regressions from time to time. For instance, in the analysis I decide to use one control variable at a time[15]. The search of the minimum point of the function is made through the *gradient-based optimization* (Snyman, 2005). Furthermore the *Conservative Log restricted-likelihood test* (Gutierrez et al.,

[14] According to this structure, if an equation consists of p random effects, the unstructured covariance matrix will have $p(p+1)/2$ parameters to be estimated.

[15] An attempt to perform a regression with all the three control variables taken together caused a failure in the estimation.

2001, McLachlan et al., 1988, Self and Yang, 1987 and Stram and Lee, 1994) is performed. The test is a comparison of the fitted mixed model to standard regression with no group-level random effects. This LR test assesses whether all random-effects parameters of the mixed model are simultaneously zero

To explain the test, consider the following Linear Mixed Model (LMM) with one variance component, written in matrix notation

$$\mathbf{GINI} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \boldsymbol{\epsilon} \tag{26}$$

where $\boldsymbol{\beta}$ is a p-dimensional vector of fixed-effects parameters and

$$cov\left(\mathbf{GINI}\right) = \sigma_\epsilon^2 \mathbf{V}_\lambda \tag{27}$$

$\mathbf{V}_\lambda = \mathbf{I}_n + \lambda \mathbf{Z}\boldsymbol{\Sigma}\mathbf{Z}^\top$, $\lambda = \frac{\sigma_b^2}{\sigma_\epsilon^2}$ and $\boldsymbol{\Sigma}$ is a known symmetric positive definite $K \times K$ matrix. Test the null hypothesis

$$H_0 : \beta_{p+1-q} = \beta^0_{p+1-q}, \ldots, \beta_p = \beta^0_p \ \sigma^2_p = 0 \ (equivalently, \lambda = 0 \ q > 0) \tag{28}$$

against the alternative hypothesis

$$H_A : \beta_{p+1-q} \neq \beta^0_{p+1-q} \ or, \ldots, \ or \beta_p \neq \beta^0_p \ or \ \sigma^2_p \neq 0 \ (equivalently, \lambda > 0) \tag{29}$$

The restricted profile *log-likelihood function* is

$$2\ell^{K,n}\left(\lambda\right) = -\log\left|\mathbf{V}_\lambda\right| - \log\left|\mathbf{X}^\intercal \mathbf{V}_\lambda^{-1}\mathbf{X}\right| - (n-p)\left(\mathbf{GINI}^\intercal \mathbf{P}_\lambda^\intercal \mathbf{V}_\lambda^{-1}\mathbf{P}_\lambda \mathbf{GINI}\right) \tag{30}$$

Then, under the null hypothesis

$$RLRT_n \underset{\lambda \geq 0}{\overset{\mathcal{D}}{=}} \left[(n-p)\log\left\{1 + \frac{N_n\left(\lambda\right)}{D_n\left(\lambda\right)}\right\} - \sum_{s=1}^{K}\log\left(1 + \lambda\mu_{s,n}\right)\right] \tag{31}$$

and the probability of having a local maximum at $\lambda = 0$ is

$$\Pr\left(\frac{\sum_{s=1}^{K}\mu_{s,n}w_s^2}{\sum_{s=1}^{n-p}w_s^2} \leq \frac{1}{n-p}\sum_{s=1}^{K}\mu_{s,n}\right) \tag{32}$$

where $w_1, \ldots, w_K$ are independent $\mathcal{N}(0,1)$ random variables and $\mu_{s,n}$ is the eigenvalue of the matrix $\Sigma^{1/2} Z^\intercal P_0 Z \Sigma^{1/2}$, satisfying $\lim_{n \longrightarrow \infty} \mu_{s,n} = \mu_s$.

4 Analysis of results

4.1 Summary statistics

Tables 1-2 show the values of Gini index for each cohort and for each country, from the cohort of nineteen-year-old to the cohort of eightyfive-year old.

Furthermore, the average Gini index and the variance is calculated for each country.

It can be seen that the inequality trend is very variegate and changeable from country to country. Tendentiously, Gini index values are below the 0.4 threshold. Nevertheless, in many cases, we can observe how this value is broadly exceeded. This especially happens in American countries such as Mexico and the United States, where the average of age group Gini indexes is higher than that of other countries (0.487 and 0.389, respectively). It is emblematic, for instance, the Mexican case, where only five cohorts have a Gini index value below the 0.4. On the contrary, other countries such as Austria, Germany, Luxembourg, Norway and Sweden are characterized both by an average level particularly low and by the absence (or almost) of cohorts having values above the 0.4.

Finally, some countries with intermediate situations, such as Ireland, Greece, and Italy, are characterized by average level of inequality at a country level and by many cohorts that exceed the 0.4 threshold.

An interesting aspect of the analysis is noticeable when we observe the values of the variance. Countries characterized by a higher variability of intergenerational inequality are Belgium, Ireland, and Mexico, while Germany, Sweden, and the United States have an almost nil variance. These latest seem to have minimized the intergenerational inequality, regardless of other generating causes of inequality (e.g. social class, ethnicity, etc.).

4.2 Age group non-parametric distribution

[Graphs 1.a - 1.n HERE]

Graphs 1.a - 1.n show the result of the Kernel density estimation for each country, depicted by the blue solid line. On every graph I added the Normal density plot (dashed line) and the Student's t density plot (dash-dotted line), for sake of comparison. The inspection of the graphs provides an interesting variability in the distribution. It is perfectly visible how the shape of the distribution changes from country to country, as for skewness and kurtosis. Furthermore, some countries, such as Belgium, Canada, Denmark, Greece, Sweden, and Switzerland exhibit a double-peaked function, with the second peak located on the right tail. Table 3 shows the results of the K-S test, where the distribution obtained under the Kernel density estimation is tested against the Normal density. The null hypothesis of equality in the two distributions is accepted for Austria, Switzerland, Germany, Spain, Greece, Ireland, Italy, Luxembourg, Mexico, Norway, Sweden, and the United States. Otherwise, it is rejected at the 10 per cent of the confidence interval for Belgium and Denmark and at the 5 per cent of the confidence interval for Canada.

Furthermore, the Kruskal-Wallis equality-of-populations rank test (Tables 4 - 5) strongly rejects the hypothesis that observations of different age groups or countries come from the same population.

Summarizing, we may affirm that the non-parametric analysis depicts a high degree of variability of cohort inequality in different countries.

4.3 Regression results

[Tables 6-15 HERE]

Tables 6-15 show regression results, obtained by clustering standard errors using different techniques. Overall, we notice a high degree of variability of results, both for coefficients values and for the levels of statistical significance.

Table 6 compares the results of regressions obtained by using the One-way error component model. As we notice, clustering by country makes all the variables significant at the 1% of the confidence interval, while clustering by age group, variables housing00, labourprogramme00 and attw67 loose statistical significance and labourprogramme95 and attw100 notably reduce it. The family policy-related variables remain particularly significant, in particular old age and health, while for tax variables the significance is maintained only for the highest average tax rates. The interpretation of the coefficient signs is not immediate because we do not have a definite sign capturing the relation between expenditure variables or tax variables and inequality. For example, expenditure policies carried out in 1995 have a positive effect on the Gini index, but this effect is capsized if we take the same policies carried out in 2000 into account. The same consideration holds also for health and old age policies, as long as for the average tax rates. Instead, housing and labour-related policies maintain the sign. Nevertheless, the effect is negative, since an increase in these two expenditure chapters generates an increase in inequality levels.

The substitution of tax-related variables with control variables determines a change in signs and significance levels for many expenditure variables. Table 7 shows the results of the regression performed by using the inflation rate as control variable, while table 8 shows the results obtained with the use of the GDP growth rate. For example, consider the family expenditure case. With respect to the model that uses tax-related variables, the variant of the model using the inflation rate produces a lower significance if standard errors are clustered by country, or, statistical significance being equal, the sign of coefficient is inverted.

The same thing happens also if we analyze the results of regressions performer by using the mixed effect technique. Both the coefficient signs and the statistical significance change according to the clustering level and the variables specified by the model. Instead, it is confirmed the significance at the 1% of the confidence interval of the fixed components related to clustering variables, country and age group, confirming the existence of idiosyncratic elements for the two dimensions.

5 Conclusions

This paper intended to measure the age group income inequality and analyze the role played by fiscal policies on inequality among cohorts. It also represents one of the first attempts to consider the problem of inequality from a microeconomic perspective, which aims to detect differences in income distribution between social groups. Results are not always easy to interpret; while we observe a robust statistical significance of expenditure and tax-related variables on age group Gini index, less clarity does exist as for the interpretation of coefficient signs. In particular, the adoption of different clustering techniques, namely one-way and two-way error component models and mixed effects models, causes the non-robustness of the results.

This study, of course, could be improved in many ways. For instance, it is difficult to separate the effects generated by fiscal policies from those generated by monetary policy, since we have to assume that the two instruments produce effects on households. Secondly, it would be important to measure the age group inequality by clustering the standard errors according to other variables, such as location, social status, and ethnicity. I hope that future researches could welcome these suggestions.

References

[1] Chakravarti, Laha, and Roy, (1967). *Handbook of Methods of Applied Statistics*, Volume I, John Wiley and Sons, pp. 392-394

[2] Cowell, F. (2000) *Measurement of Inequality*, in Handbook of Income Distribution, Amsterdam, Elsevier

[3] Afonso, A. and al. (2008) *Income Distribution Determinants and Public Spending Efficiency*, mimeo

[4] Afonso, A. and al. (2005) *Public Sector Efficiency: An International Comparison*, Public Choice, 123 (3-4), 321-347.

[5] Afonso, A. and al. (2006) *Public Sector Efficiency: Evidence for New EU Member States and Emerging Markets*, European Central

Bank, Working Paper n. 581, January

[6] Altonji, J. G., Hayashi, F. and Kotlikoff, L. J. (1995) *Parental Altruism and Inter Vivos Transfers: Theory and Evidence*, National Bureau of Economic Research Working Paper, no 5378

[7] Atkinson AB (1970) *On the measurement of inequality*. Journal of Economic Theory 2: 244-263

[8] Baltagi, B. (2008) Econometric Analysis of Paneld Data, Fourth Edition, John Wiley & Sons, London

[9] Eschker, E. (2007) *Fiscal Redistribution by Age and Generational Inequality in the Twentieth Century*, unpublished paper

[10] EUROSTAT (2006), *European Social Statistics, Social ¨Protection, Expenditure and Receipts*, Data 1995-2003, Luxembourg.

[11] Fatás, A. and Mihov, I. (2001) *The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence*, CEPR Discussion Paper No. 2760. Available at SSRN: http://ssrn.com/abstract=267281

[12] Giavazzi F. et al. (1999), *Searching for Non-Keynesian Effects of Fiscal Policy*, CSEF Working Paper, No. 16

[13] Gottschalk, T. & Smeeding, T. (2000) *Empirical Evidence on Income Inequality in Industrialized Countries*, in Handbook of Income Distribution, Amsterdam, Elsevier

[14] Gutierrez, R. G., S. L. Carter, and D. M. Drukker (2001) *On boundary-value likelihood-ratio tests*. Stata Technical Bulletin 60:15-18. Reprinted in Stata Technical Bulletin Reprints, vol. 8, pp. 233-236.

[15] Hox, J. (1995) *Applied Multilevel Analysis*, Amsterdam, TT-Publikaties

[16] Huggett and al. (2007), *Source of Lifetime Inequality*, mimeo

[17] Kruskal, W. and Wallis, W. (1952). *Use of ranks in one-criterion variance analysis*. Journal of the American Statistical Association 47 (260): 583–621

[18] Lindert, P. (2000) *Three Centuries of Inequality in Britain and America*, in Handbook of Income Distribution, Amsterdam, Elsevier

[19] McLachlan, G. J., and K. E. Basford. (1988) *Mixture Models*. New York: Marcel Dekker

[20] OECD (2004) *Recent Tax Policy Trends and Reforms in OECD Countries*, Tax Policy Studies No.9

[21] OECD (2006a), *OECD Health Data 2006*, OECD, Paris.

[22] OECD (2006b), *OECD Employment Outlook*, OECD, Paris.

[23] Petersen, M. (2006) *Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches*, Kellogg School of Management, Northwestern University, mimeo

[24] Petersen, M. (2006) *Programming Advice - Finance Panel Data Set*s, Kellogg School of Management, mimeo

[25] Shorak, G.R., Wellner, J.A. (1986) *Empirical Processes with Applications to Statistics*, Wiley

[26] Self, S. G., and K.-Y. Liang (1987) *Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions*, Journal of the American Statistical Association 82: 605-610.

[27] Silverman, B.W. (1986) *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.

[28] Simonoff, J.S. (1996) *Smoothing Methods in Statistics*. New York: Springer

[29] Snyman, J. (2005) *Practical mathematical optimization: an introduction to basic optimization theory and classical and new gradient-based algorithms*, Lavoisier

[30] Stock, J. and Watson, M. (2006) *Heteroskedasticity-Robust Standard Errors for Fixed Effects Panel Data Regression*, mimeo

[31] Stram, D. O., and J. W. Lee (1994) *Variance components testing in the longitudinal mixed effects model*. Biometrics 50: 1171-1177.

[32] Tsoukalas, D. (2008), *Comparing Monetary and Fiscal Policy Trasmission Across the Americas and the Emu-Area*, IABR & TLC Conference Proceedings

[33] Wand, M. and Jones, M. (1995) *Kernel Smoothing*, Monographs on Statistics and Applied Probability, Chapman & Hall

[34] WIDER (World Institute for Development Economics Research), (2000) *World Income Inequality Database: User Guide and Data Sources*. Helsinki: United Nations University

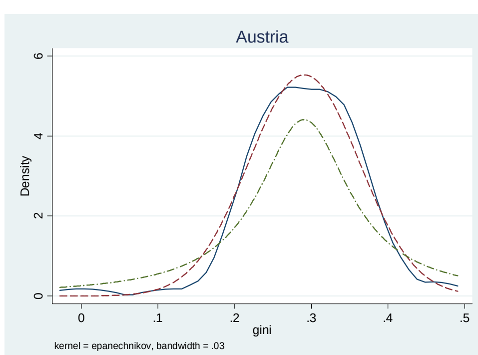


Figure 1.a

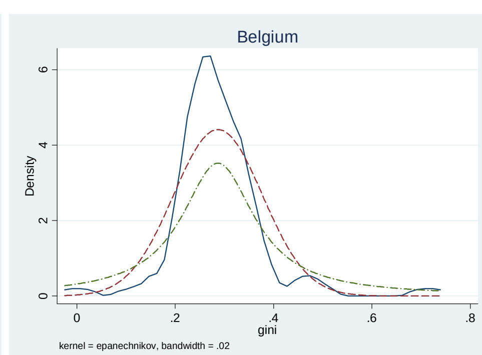
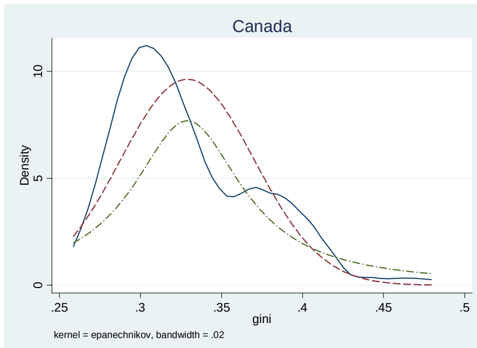


Figure 1.b



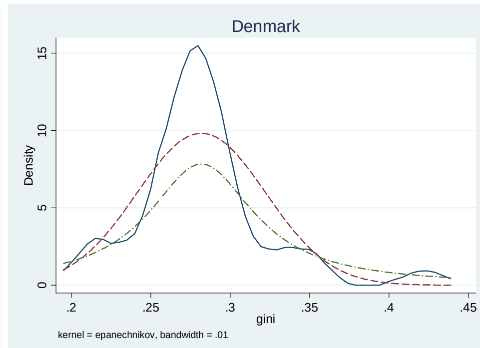
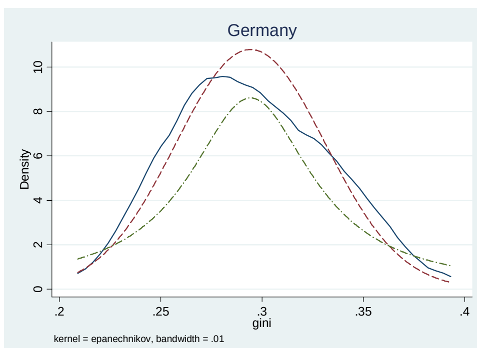


Figure 1.c

Figure 1.d



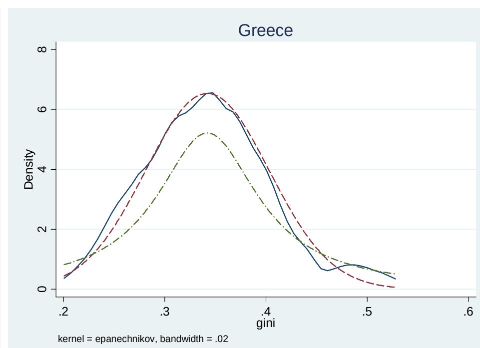
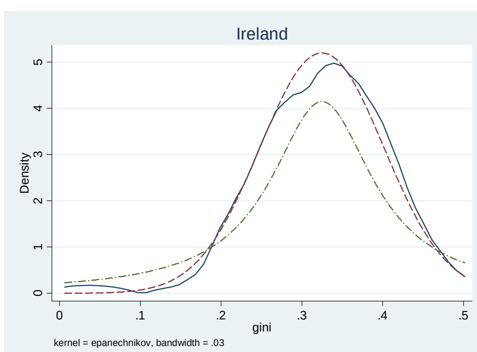
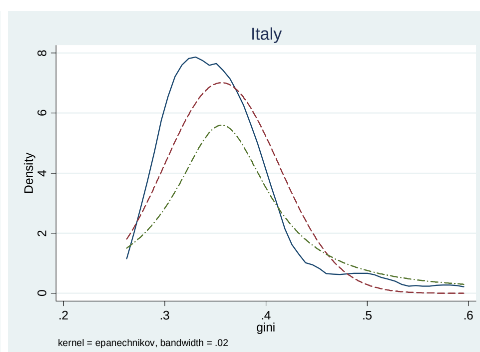


Figure 1.e

Figure 1.f





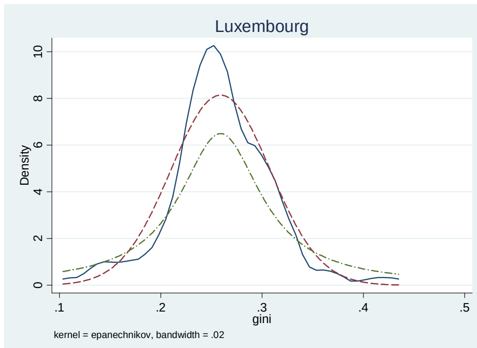


Figure 1.g

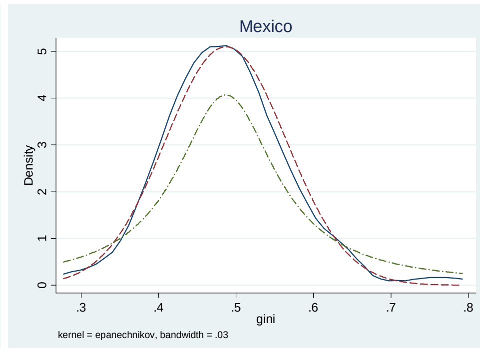


Figure 1.h

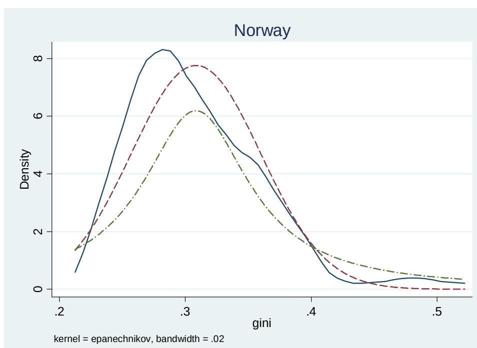


Figure 1.k

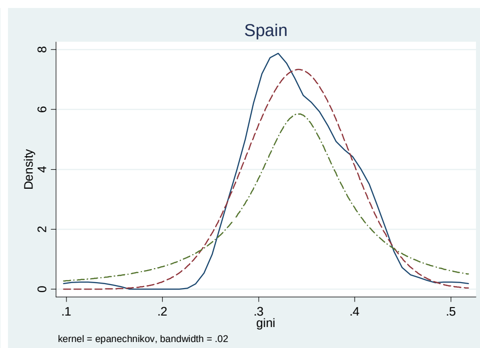


Figure 1.i

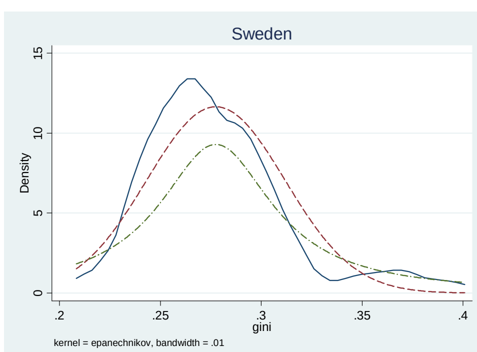


Figure 1.l

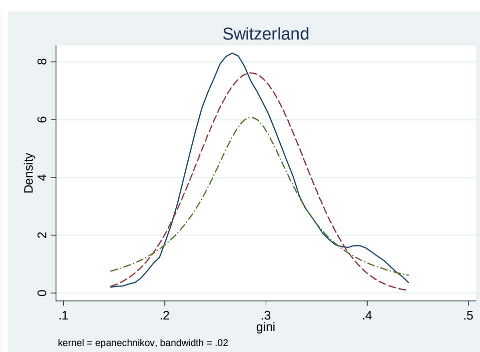


Figure 1.m

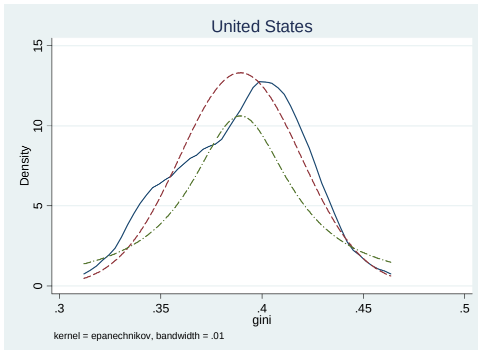

United States
Density
0
5
10
15
.3
.35
.4
.45
.5
gini
kernel = epanechnikov, bandwidth = .01

Figure 1.n

Table 1: Age group Gini indexes

Age group	Austria	Belgium	Canada	Switzerland	Germany	Denmark	Spain
19	-	-	0.463	0.222	0.371	0.429	0.118
20	0.125	0.164	0.394	0.166	0.309	0.415	0.270
21	0.265	0.300	0.393	0.265	0.344	0.352	0.340
22	0.385	0.323	0.330	0.272	0.303	0.355	0.298
23	0.319	0.322	0.342	0.288	0.351	0.341	0.292
24	0.209	0.227	0.318	0.241	0.317	0.347	0.377
25	0.228	0.239	0.302	0.275	0.296	0.320	0.311
26	0.295	0.234	0.310	0.258	0.286	0.305	0.280
27	0.249	0.230	0.315	0.267	0.286	0.282	0.288
28	0.231	0.230	0.304	0.256	0.224	0.295	0.280
29	0.237	0.359	0.300	0.241	0.264	0.272	0.343
30	0.231	0.269	0.298	0.262	0.295	0.257	0.306
31	0.266	0.256	0.295	0.238	0.250	0.260	0.323
32	0.259	0.219	0.289	0.239	0.259	0.252	0.311
33	0.326	0.295	0.280	0.278	0.256	0.272	0.294
34	0.244	0.343	0.294	0.243	0.259	0.248	0.285
35	0.291	0.253	0.324	0.400	0.239	0.261	0.298
36	0.224	0.261	0.291	0.290	0.246	0.265	0.278
37	0.215	0.225	0.325	0.203	0.235	0.267	0.297
38	0.327	0.364	0.293	0.215	0.252	0.271	0.317
39	0.240	0.271	0.307	0.250	0.254	0.270	0.305
40	0.194	0.260	0.288	0.252	0.248	0.270	0.332
41	0.194	0.283	0.305	0.247	0.268	0.274	0.342
42	0.268	0.328	0.318	0.267	0.274	0.261	0.321
43	0.293	0.293	0.308	0.228	0.261	0.290	0.295
44	0.248	0.264	0.296	0.254	0.309	0.267	0.299
45	0.235	0.265	0.313	0.379	0.287	0.269	0.330
46	0.231	0.185	0.344	0.266	0.258	0.282	0.333
47	0.265	0.317	0.312	0.262	0.302	0.289	0.327
48	0.271	0.340	0.317	0.299	0.287	0.281	0.332
49	0.260	0.314	0.323	0.327	0.282	0.291	0.387
50	0.346	0.249	0.354	0.256	0.277	0.282	0.332

Age group	Austria	Belgium	Canada	Switzerland	Germany	Denmark	Spain
51	0.219	0.297	0.346	0.315	0.282	0.294	0.320
52	0.283	0.288	0.383	0.272	0.325	0.281	0.335
53	0.296	0.373	0.378	0.291	0.309	0.323	0.324
54	0.346	0.714	0.363	0.313	0.272	0.296	0.296
55	0.220	0.269	0.388	0.327	0.336	0.281	0.337
56	0.312	0.473	0.374	0.291	0.321	0.285	0.381
57	0.328	0.294	0.396	0.325	0.320	0.289	0.359
58	0.280	0.356	0.391	0.335	0.310	0.280	0.350
59	0.311	0.323	0.391	0.339	0.314	0.280	0.433
60	0.333	0.362	0.362	0.247	0.340	0.289	0.434
61	0.381	0.265	0.360	0.307	0.292	0.299	0.399
62	0.346	0.346	0.376	0.385	0.334	0.286	0.398
63	0.315	0.308	0.414	0.275	0.325	0.319	0.497
64	0.269	0.246	0.400	0.336	0.350	0.300	0.407
65	0.371	0.265	0.342	0.281	0.271	0.338	0.380
66	0.360	0.248	0.305	0.293	0.277	0.308	0.416
67	0.331	0.226	0.340	0.254	0.336	0.294	0.406
68	0.320	0.231	0.317	0.402	0.290	0.288	0.358
69	0.451	0.226	0.303	0.288	0.273	0.288	0.333
70	0.364	0.317	0.287	0.305	0.268	0.254	0.389
71	0.375	0.334	0.275	0.404	0.286	0.279	0.350
72	0.345	0.357	0.313	0.294	0.298	0.264	0.36
73	0.265	0.255	0.279	0.324	0.301	0.268	0.393
74	0.340	0.297	0.309	0.239	0.379	0.264	0.364
75	0.300	0.189	0.307	0.240	0.322	0.257	0.391
76	0.346	0.312	0.276	0.300	0.351	0.281	0.325
77	0.298	0.249	0.287	0.238	0.339	0.246	0.411
78	0.352	0.236	0.306	0.319	0.290	0.243	0.344
79	0.463	0.488	0.278	0.220	0.367	0.214	0.398
80	0.379	0.445	0.300	0.290	0.343	0.220	0.330
81	0.317	0.274	-	0.290	0.266	0.227	0.379
82	0.375	0.230	-	0.421	0.291	0.228	0.297
83	0.239	0.118	-	0.215	0.245	0.205	0.394
84	0.306	0.294	-	0.360	0.327	0.205	0.403
85	0.360	0.236	-	0.366	0.223	0.219	-
Avg	**0.295**	**0.291**	**0.329**	**0.285**	**0.294**	**0.282**	**0.341**
Var	0.004	0.007	0.002	0.003	0.001	0.002	0.003

Table 2: Age group Gini indexes

Age group	Greece	Ireland	Italy	Luxembourg	Mexico	Norway	Sweden	United States
19	0.498	0.036	0.518	0.119	0.317	0.361	0.352	0.403
20	0.341	0.226	0.577	0.153	0.401	0.355	0.360	0.427
21	0.506	0.366	0.429	0.153	0.307	0.320	0.354	0.392
22	0.280	0.243	0.288	0.243	0.389	0.335	0.382	0.391
23	0.222	0.242	0.499	0.283	0.410	0.329	0.314	0.354
24	0.367	0.208	0.352	0.211	0.366	0.317	0.312	0.354
25	0.408	0.294	0.475	0.246	0.378	0.295	0.306	0.324
26	0.258	0.236	0.322	0.241	0.446	0.253	0.283	0.354
27	0.295	0.282	0.369	0.274	0.410	0.257	0.282	0.343
28	0.314	0.249	0.281	0.228	0.420	0.281	0.262	0.361
29	0.236	0.237	0.302	0.251	0.390	0.252	0.249	0.332
30	0.330	0.293	0.329	0.241	0.486	0.255	0.273	0.349
31	0.312	0.213	0.331	0.261	0.411	0.287	0.258	0.330
32	0.242	0.255	0.400	0.265	0.587	0.255	0.243	0.344
33	0.269	0.275	0.292	0.235	0.425	0.267	0.253	0.359
34	0.281	0.313	0.396	0.354	0.484	0.258	0.245	0.382
35	0.338	0.298	0.303	0.254	0.437	0.294	0.225	0.377
36	0.303	0.322	0.300	0.273	0.503	0.299	0.282	0.340
37	0.342	0.317	0.330	0.251	0.493	0.232	0.272	0.356
38	0.291	0.274	0.311	0.287	0.411	0.285	0.266	0.350
39	0.308	0.251	0.304	0.275	0.429	0.343	0.256	0.357
40	0.383	0.296	0.281	0.266	0.451	0.284	0.282	0.368
41	0.303	0.271	0.312	0.262	0.456	0.278	0.267	0.376
42	0.259	0.237	0.321	0.200	0.511	0.259	0.243	0.364
43	0.275	0.310	0.363	0.240	0.458	0.387	0.285	0.353
44	0.272	0.302	0.322	0.262	0.459	0.258	0.299	0.387
45	0.249	0.305	0.288	0.248	0.494	0.266	0.289	0.368
46	0.325	0.425	0.344	0.253	0.472	0.322	0.263	0.381
47	0.314	0.385	0.332	0.274	0.542	0.294	0.257	0.382
48	0.266	0.325	0.334	0.244	0.554	0.367	0.283	0.404
49	0.339	0.317	0.349	0.242	0.522	0.300	0.269	0.377
50	0.287	0.348	0.327	0.255	0.466	0.339	0.302	0.403

Age group	Greece	Ireland	Italy	Luxembourg	Mexico	Norway	Sweden	United States
51	0.286	0.292	0.299	0.233	0.536	0.309	0.276	0.377
52	0.326	0.323	0.355	0.308	0.475	0.372	0.286	0.387
53	0.332	0.281	0.332	0.254	0.430	0.304	0.242	0.387
54	0.372	0.361	0.300	0.309	0.471	0.347	0.254	0.401
55	0.335	0.392	0.361	0.297	0.525	0.292	0.298	0.406
56	0.378	0.335	0.397	0.319	0.475	0.379	0.298	0.400
57	0.457	0.391	0.376	0.308	0.536	0.332	0.389	0.409
58	0.410	0.348	0.368	0.316	0.496	0.328	0.288	0.414
59	0.405	0.354	0.383	0.309	0.530	0.295	0.305	0.452
60	0.334	0.363	0.384	0.308	0.610	0.342	0.274	0.394
61	0.328	0.374	0.380	0.310	0.457	0.393	0.255	0.432
62	0.411	0.349	0.379	0.362	0.522	0.378	0.309	0.416
63	0.379	0.373	0.344	0.291	0.469	0.316	0.256	0.417
64	0.410	0.422	0.450	0.250	0.502	0.460	0.273	0.447
65	0.377	0.398	0.340	0.287	0.509	0.368	0.296	0.414
66	0.417	0.376	0.388	0.291	0.547	0.328	0.291	0.412
67	0.348	0.472	0.369	0.295	0.427	0.304	0.251	0.418
68	0.373	0.466	0.418	0.250	0.481	0.267	0.289	0.394
69	0.360	0.396	0.421	0.215	0.501	0.260	0.275	0.400
70	0.376	0.407	0.399	0.268	0.552	0.264	0.279	0.414
71	0.314	0.397	0.367	0.251	0.486	0.275	0.246	0.396
72	0.343	0.368	0.357	0.225	0.567	0.300	0.256	0.406
73	0.341	0.420	0.354	0.250	0.608	0.268	0.255	0.420
74	0.403	0.368	0.342	0.316	0.435	0.325	0.260	0.407
75	0.388	0.343	0.320	0.208	0.534	0.360	0.253	0.428
76	0.354	0.369	0.321	0.248	0.417	0.371	0.252	0.394
77	0.371	0.449	0.383	0.310	0.591	0.280	0.301	0.404
78	0.404	0.349	0.305	0.246	0.510	0.284	0.252	0.420
79	0.367	0.306	0.379	0.210	0.624	0.273	0.220	0.443
80	0.349	0.444	0.360	0.231	0.541	0.257	0.256	0.410
81	0.368	0.169	0.310	0.222	0.513	0.502	0.274	0.417
82	0.334	0.376	0.384	0.238	0.650	0.271	0.252	0.381
83	0.478	0.418	0.320	0.193	0.436	0.254	0.306	0.425
84	-	0.241	0.293	0.420	0.609	0.232	0.229	0.415
85	-	0.322	0.347	0.161	0.762	0.296	0.230	0.388
Avg	**0.295**	**0.291**	**0.329**	**0.285**	**0.294**	**0.282**	**0.341**	**0.342**
0.325	**0.356**	**0.259**	**0.487**	**0.308**	**0.278**	**0.389**		
Var	0.004	0.007	0.002	0.003	0.001	0.002	0.003	0.004
0.006	0.003	0.002	0.006	0.003	0.001	0.001		

Table 3: Combined One-sample Kolmogorov-Smirnov test

Country	D	P-value	Corrected
Austria	0.0736	0.861	0.826
Belgium	0.1577	0.071	0.053
Canada	0.1811	0.034	0.024
Switzerland	0.1059	0.440	0.382
Germany	0.0676	0.920	0.895
Denmark	0.1544	0.082	0.062
Spain	0.0910	0.645	0.588
Greece	0.0612	0.968	0.956
Ireland	0.0646	0.942	0.923
Italy	0.1171	0.317	0.266
Luxembourg	0.0983	0.536	0.477
Mexico	0.0605	0.967	0.955
Norway	0.1181	0.307	0.257
Sweden	0.0957	0.572	0.513
United States	0.0878	0.681	0.625

Table 4: Kruskal-Wallis equality-of-populations rank test - Country

Country	Observations	Rank Sum
Austria	67	26345.50
Belgium	67	23359.50
Canada	62	34934.50
Switzerland	67	22886.50
Germany	67	26293.00
Denmark	67	21612.50
Spain	66	40908.50
Greece	65	38898.50
Ireland	67	35796.00
Italy	67	44465.50
Luxembourg	67	15566.00
Mexico	67	62259.00
Norway	67	30073.50
Sweden	67	19874.50
United States	67	54230.00
chi-squared = 447.060 with 14 d.f.		
probability = 0.0001		
chi-squared with ties = 447.060 with 14 d.f.		
probability = 0.0001		

Table 5: Kruskal-Wallis equality-of-populations rank test - Age Group

Age Group	Observations	Rank Sum	Age Group	Observations	Rank Sum
19	15	7429.00	53	15	7961.50
20	15	7390.00	54	15	9077.00
21	15	8905.00	55	15	8534.00
22	15	7779.50	56	15	10111.50
23	15	7962.00	57	15	10360.50
24	15	6680.00	58	15	9585.50
25	15	6931.50	59	15	9927.00
26	15	5309.00	60	15	9517.50
28	15	4688.00	61	15	9402.50
29	15	4921.00	62	15	10962.50
30	15	5759.00	63	15	9413.50
31	15	5150.00	64	15	9805.00
32	15	4786.00	65	15	9258.00
33	15	5409.00	66	15	9061.50
34	15	6322.50	67	15	8566.00
35	15	6499.00	68	15	8572.00
36	15	5716.00	69	15	7829.00
37	15	5339.00	70	15	8414.50
38	15	6318.50	71	15	8302.00
39	15	5674.00	72	15	8310.50
40	15	5789.50	73	15	7461.00
41	15	5765.00	74	15	8742.50
42	15	5498.00	75	15	7422.50
43	15	6683.00	76	15	8275.00
44	15	5944.50	77	15	8228.00
45	15	6175.00	78	15	7409.00
46	15	6916.00	79	15	8112.00
47	15	7456.50	80	15	8424.00
48	15	7597.00	81	15	6683.00
49	15	7778.00	82	15	7078.50
50	15	7574.00	83	15	5984.50
51	15	6691.50	84	15	6473.00
52	15	8482.00	85	15	5161.00
chi-squared = 121.805 with 66 d.f..					
probability = 0.0001					
chi-squared with ties = 121.805 with 66 d.f.					
probability = 0.0001					

Table 6: One-way Error Component Model (I)

Gini index	**Country**	**Age group**
family95	-0.270**	-0.270**
family00	0.322**	0.322**
health95	0.195**	0.195**
health00	-0.201**	-0.201**
housing95	0.081**	0.081**
housing00	0.006**	0.006
labourprogramme95	0.086**	0.086†
labourprogramme00	0.071**	0.071
oldage95	-0.085**	-0.085**
oldage00	0.108**	0.108**
attw67	-1.796**	-1.796
attw100	7.239**	7.239*
attw133	-15.818**	-15.818**
attw167	9.228**	9.228**
Intercept	0.340**	0.340**
N	997	997
R^2	0.483	0.483
F	98.085	

Significance level: †: 10% *: 5% **: 1%

Table 7: One-way Error Component Model (II)

Gini index	**Country**	**Age group**
family95	0.138†	0.138**
family00	-0.267*	-0.267**
health95	0.245**	0.245**
health00	-0.319**	-0.319**
housing95	-1.009**	-1.009**
housing00	0.866**	0.866**
labourprogramme95	0.259*	0.259**
labourprogramme00	0.047*	0.047**
oldage95	0.072*	0.072**
oldage00	-0.069*	-0.069**
cpi97	-0.180**	-0.180**
cpi98	0.271**	0.271**
cpi99	-0.164**	-0.164**
Intercept	1.040**	1.040**
N	997	997
R^2	0.45	0.45
F	457.18	61.494

Significance level: †: 10% *: 5% **: 1%

Table 8: One-way Error Component Model (III)

Gini index	**Country**	**Age group**
family95	0.035	0.035*
family00	-0.009	-0.009
health95	-0.011	-0.011
health00	0.035	0.035†
housing95	-0.032	-0.032†
housing00	0.106†	0.106**
labourprogramme95	-0.091*	-0.091**
labourprogramme00	0.125*	0.125**
oldage95	-0.096**	-0.096**
oldage00	0.072**	0.072**
gdpgr97	-0.029*	-0.029**
gdpgr98	-0.033*	-0.033**
gdpgr99	0.088*	0.088**
Intercept	0.226†	0.226**
N	997	997
R^2	0.448	0.448
F	39.237	76.903

Significance level: †: 10% *: 5% **: 1%

Table 9: Two-way Error Component Model

Gini index	**(I)**	**(II)**	**(III)**
family95	-0.270**	0.138†	0.035
family00	0.322**	-0.267*	-0.009
health95	0.195	0.245**	-0.011
health00	-0.201	-0.319**	0.035
housing95	0.081	-1.009**	-0.032
housing00	0.006	0.866**	0.106†
labourprogramme95	0.086**	0.259**	-0.091*
labourprogramme00	0.071	0.047**	0.125*
oldage95	-0.085	0.072*	-0.096**
oldage00	0.108	-0.069*	0.072**
attw67	-1.796**		
attw100	7.239**		
attw133	-15.818**		
attw167	9.228		
cpi97	-0.180**		
cpi98	0.271**		
cpi99	-0.164**		
gdpgr97		-0.029*	
gdpgr98		-0.033**	
gdpgr99		0.088**	
Intercept	0.340**	1.040**	0.226†
N	997	997	997
R^2	0.483	0.45	0.448
F	75.942	58.454	62.23

Significance level: †: 10% *: 5% **: 1%

Table 10: 3-Level Mixed Effects (I)

Gini index	**Coefficient**	**(Std. Err.)**
Equation 1 : gini		
family95	-0.270**	(0.032)
family00	0.322**	(0.032)
health95	0.194**	(0.032)
health00	-0.201**	(0.020)
housing95	0.081**	(0.025)
housing00	0.006	(0.031)
labourprogramme95	0.086*	(0.043)
labourprogramme00	0.070	(0.055)
oldage95	-0.084**	(0.022)
oldage00	0.108**	(0.017)
attw67	-1.809	(1.151)
attw100	7.278*	(3.104)
attw133	-15.829**	(2.055)
attw167	9.218**	(0.799)
Intercept	0.340**	(0.028)
Equation 2 : lns1_1_1		
Intercept	-3.918**	(0.131)
Equation 3 : lnsig_e		
Intercept	-2.915**	(0.023)
N	997	
Log-likelihood	1411.686	
$\chi^2_{(14)}$	1042.703	

Significance level: †: 10% *: 5% **: 1%

Table 11: 3-Level Mixed Effects (II)

Gini index	Coefficient	(Std. Err.)
Equation 1 : gini		
family95	0.138**	(0.025)
family00	-0.267**	(0.034)
health95	0.245**	(0.023)
health00	-0.319**	(0.030)
housing95	-1.008**	(0.104)
housing00	0.866**	(0.087)
labourprogramme95	0.259**	(0.034)
labourprogramme00	0.047**	(0.013)
oldage95	0.071**	(0.012)
oldage00	-0.069**	(0.011)
cpi97	-0.180**	(0.019)
cpi98	0.271**	(0.029)
cpi99	-0.164**	(0.022)
Intercept	1.040**	(0.067)
Equation 2 : lns1_1_1		
Intercept	-3.938**	(0.136)
Equation 3 : lnsig_e		
Intercept	-2.877**	(0.023)
N	997	
Log-likelihood	1364.573	
$\chi^2_{(13)}$	900.183	

Significance level: †: 10% *: 5% **: 1%

Table 12: 3-Level Mixed Effects (III)

Gini index	**Coefficient**	**(Std. Err.)**
Equation 1 : gini		
family95	0.035	(0.023)
family00	-0.009	(0.021)
health95	-0.011	(0.010)
health00	0.035*	(0.015)
housing95	-0.032	(0.023)
housing00	0.106**	(0.024)
labourprogramme95	-0.091**	(0.022)
labourprogramme00	0.125**	(0.025)
oldage95	-0.096**	(0.010)
oldage00	0.072**	(0.008)
gdpgr97	-0.029**	(0.004)
gdpgr98	-0.033**	(0.004)
gdpgr99	0.088**	(0.010)
Intercept	0.226**	(0.040)
Equation 2 : lns1_1_1		
Intercept	-3.939**	(0.136)
Equation 3 : lnsig_e		
Intercept	-2.875**	(0.023)
N	997	
Log-likelihood	1360.598	
$\chi^2_{(13)}$	891.897	

Significance level: †: 10% *: 5% **: 1%

Table 13: 2-Level Mixed Effects (I)

Gini index	**Coefficient**	**(Std. Err.)**
Equation 1 : gini		
family95	-0.270**	(0.096)
family00	0.322**	(0.097)
health95	0.195*	(0.097)
health00	-0.201**	(0.059)
housing95	0.081	(0.075)
housing00	0.006	(0.093)
labourprogramme95	0.086	(0.132)
labourprogramme00	0.071	(0.167)
oldage95	-0.085	(0.067)
oldage00	0.108*	(0.050)
attw67	-1.796	(3.492)
attw100	7.239	(9.414)
attw133	-15.818*	(6.253)
attw167	9.228**	(2.416)
Intercept	0.340**	(0.086)
Equation 2 : lns1_1_1		
Intercept	-3.969	(0.000)
Equation 3 : lnsig_e		
Intercept	-2.851	(0.000)
N	997	
Log-likelihood	1385.795	
$\chi^2_{(14)}$	112.502	

Significance level: †: 10% *: 5% **: 1%

Table 14: 2-Level Mixed Effects (II)

Gini index	**Coefficient**	**(Std. Err.)**
Equation 1 : gini		
family95	0.138	(0.202)
family00	-0.267	(0.283)
health95	0.245	(0.192)
health00	-0.320	(0.243)
housing95	-1.014	(0.848)
housing00	0.870	(0.707)
labourprogramme95	0.261	(0.278)
labourprogramme00	0.047	(0.103)
oldage95	0.072	(0.100)
oldage00	-0.069	(0.086)
cpi97	-0.181	(0.155)
cpi98	0.272	(0.232)
cpi99	-0.165	(0.179)
Intercept	$1.041^{\dagger}$	(0.551)
Equation 2 : lns1_1_1		
Intercept	-2.882^{**}	(0.718)
Equation 3 : lnsig_e		
Intercept	-2.851^{**}	(0.023)
N	997	
Log-likelihood	1370.925	
$\chi^2_{(13)}$	13.34	

Significance level: †: 10% *: 5% **: 1%

Table 15: 2-Level Mixed Effects (III)

Gini index	**Coefficient**	**(Std. Err.)**
Equation 1 : gini		
family95	0.036	(0.195)
family00	-0.010	(0.176)
health95	-0.011	(0.088)
health00	0.035	(0.127)
housing95	-0.033	(0.196)
housing00	0.106	(0.199)
labourprogramme95	-0.090	(0.185)
labourprogramme00	0.125	(0.208)
oldage95	-0.096	(0.085)
oldage00	0.072	(0.064)
gdpgr97	-0.029	(0.037)
gdpgr98	-0.033	(0.038)
gdpgr99	0.088	(0.084)
Intercept	0.227	(0.342)
Equation 2 : lns1_1_1		
Intercept	-2.846**	(0.718)
Equation 3 : lnsig_e		
Intercept	-2.851**	(0.023)
N	997	
Log-likelihood	1369.033	
$\chi^2_{(13)}$	12.37	

Significance level: †: 10% *: 5% **: 1%