the review of income and wealth

Review of Income and Wealth
Series 61, Number 3, September 2015
DOI: 10.1111/roiw.12094

THE "MAKE-UP" OF A REGRESSION COEFFICIENT: GENDER GAPS IN THE EUROPEAN LABOR MARKET

BY M. GRAZIA PITTAU*

Department of Statistical Science, Sapienza University of Rome

SHLOMO YITZHAKI

The Hebrew University of Jerusalem and Central Bureau of Statistics

AND

ROBERTO ZELLI

Department of Statistical Science, Sapienza University of Rome

We provide a comprehensive picture of the relationship between labor market outcomes and age by gender in the 28 European countries covered by the European Statistics on Income and Living Conditions. The analysis is based on a somewhat unconventional approach that refers to concentration curves in the Gini regression framework. It allows identification of ranges in the explanatory variables where local slopes change sign and/or size, i.e. the components that "make up" a regression coefficient. Gender is a crucial factor differentiating participation among workers, although employment–age profiles do not substantially differ. Relevant differences in age profiles concern working-hours patterns: some countries are characterized by an almost specular behavior in men and women; other countries instead show similar patterns. Generally, earnings increase with age for both men and women. However, local regression coefficients are not monotonic over the entire age range and can even be locally negative in some countries.

JEL Codes: C30, J16, J21

Keywords: age profile, concentration curves, gender gap, Gini regression, labor market

1. INTRODUCTION

Although in Europe there is a documented trend toward a narrowing gender gap in the labor market, there are still significant gender inequalities in terms of participation, working time, and earnings (OECD, 2002; European Commission, 2006). These gender inequalities could depend on differences in education, skill experiences, underestimation, and discrimination. A large body of literature has tried to explain gender pay gap as differences related to human capital and other "unexplained" characteristics potentially associated with discrimination (Blau and Kahn, 2000). Age is a key factor in the labor market, not only as a proxy for experience but also in the shaping of long-term dynamics in the structure of paid labor force and distribution of working hours. Age structure influences participation in

Note: We thank two anonymous referees and the editor for helpful comments. Financial support from Sapienza research grants is gratefully acknowledged.

*Correspondence to: M. Grazia Pittau, Dipartimento di Scienze Statistiche, Sapienza Università di Roma, P.le Aldo Moro 5, 00185 Roma, Italy (grazia.pittau@uniroma1.it).

© 2014 International Association for Research in Income and Wealth

the labor market according to a three-part division: younger individuals are less likely to participate because they are involved in formal schooling, senior individuals are likely to leave the labor-force around their sixties for retirement, and the intermediate age group represents the "working age" population. Of course participation of individuals in the labor market is also affected by a variety of other circumstances such as cultural traditions, productive structure, and labor regulation, which can significantly influence average working hours. As women and men age, their willingness to work, their amount of time set aside for work, and ultimately their earnings differentiate: age carries important implications for gender disparities in the labor market (Lee *et al.*, 2007). Notwithstanding a tremendous literature on gender gap in European labor markets, there is still a scarcity of cross-country studies of gender gap profiles in Europe. Based on the European Statistics on Income and Living Conditions (EU-SILC) dataset, this paper provides a comprehensive picture of the relationship between labor market outcomes and age by gender in Europe. The EU-SILC is a collection of annual national surveys of socio-economic conditions of individuals and households in the EU countries. All national surveys in EU-SILC have standard questionnaires and procedures for data processing and yield *ex-ante* harmonized microdata that allow homogeneous inter-country comparisons using a uniform protocol.

This study embodies a somewhat unconventional approach to identify age profiles of labor market outcomes. This approach, originally illustrated in Yitzhaki and Schechtman (2012), is based on the use of concentration curves in the context of a Gini regression framework. Unlike simple OLS regression, this method allows us to investigate non-monotonicity between two variables. It also has advantages over conventional non-parametric regression methods since it goes beyond the visual impression of non-parametric curves, making possible the identification of ranges in the explanatory variables where local slopes change sign and/or size. In other words, we are able to investigate the components that "make up" a regression coefficient.

The paper is articulated as follows: after the present introduction, Section 2 describes the data and provides a comprehensive cross-country summary of the labor market variables used in the profile analysis. Section 3 presents a brief review of our method based on concentration curves and Gini regression coefficients. Section 4 reports the main empirical results on the relationship between labor market outcomes (namely labor participation, hours of work, and earnings) and age by gender in Europe. We cluster European countries according to their age profiles and for each cluster a representative country is more deeply investigated. Section 5 offers conclusions and suggestions for further research.

2. Data, Definitions, and Early Evidence

The lack of comparable data has often prevented comprehensive analysis of gender disparities across a large set of countries. Furthermore, problems in sample coverage, quality of data, and the fact that labor market outcomes are often based on different definitions, hampered comparability even within Europe. EU-SILC overcomes these concerns since the national surveys are based on a standard questionnaire and provide *ex-ante* harmonized micro-data on European countries.

© 2014 International Association for Research in Income and Wealth

We used data from the latest release (cross-sectional UDB SILC 2008 Rev.4, March 2012) of the 2008 wave, which collects information on household activities related to the year 2007.

Since our analysis focuses on the structure of employment, hours of work, and earnings by age and gender, we selected, for each respondent, age, personal cross-sectional sampling weight, activity status, average weekly hours of work in main and secondary activities, employee income, and self-employed income. For most of the countries we were able to identify age in terms of years and quarters. We selected individuals aged between 18 and 65, which corresponds to the typical end of high school and the most frequent year of retirement in Europe.

In the following we give a description of the variables (for more details see the reference manual: Eurostat, 2009). Each household respondent (of any age) is labeled by her activity status calculated on the basis of questions concerning number of months spent at full-time work, at part-time work, in unemployment, in retirement, studying, and in inactivity. The respondent is identified as employed (full-time and part-time), unemployed, retired, or in other inactivities if the respondent has spent more than six months in that status. If the respondent has not spent at least six months in none of those, the variable is coded as missing/undefined.

Disparities between women and men in the labor market are evident in two primary measures: labor force participation rates (the gender participation gap) and average earnings (the gender wage gap). Participation rate can be defined as the percentage of employed and unemployed over the total population aged 18–65. To better relate labor participation and wages, we follow the literature in estimating employment rates by gender, as the ratio between employed and total population aged 18–65.

Table 1 lists the countries covered by EU-SILC, the number of workers in the sample, and the estimated employment rate by gender in 2008. The employment gap is measured as the difference between men's and women's employment rates as a percentage of men's employment rate. It ranges from 6.3 percent in Iceland to 34.8 percent in Italy. In line with the findings of previous studies (e.g., OECD, 2002), Mediterranean countries as well as most Eastern European countries are characterized by wide gaps. For Italy, Hungary, and Poland this alarming scenario is further amplified by relatively low male employment rates. Nordic countries, instead, display the lowest employment gaps among the EU economies, accompanied by high levels of employment rates for both men and women.

Earnings are defined as the total gross annual remuneration, in cash or in kind, payable by an employer to an employee for work done by the latter during the income reference period.[1] Employee income is composed of gross employee cash or near cash income, gross non-cash employee income, and employer's social insurance contribution. Self-employment income includes gross cash benefits or

[1]The use of gross income is due to data availability. EU-SILC follows the Canberra group recommendation on income that suggests to collect gross incomes at component level, leaving to the single national statistical offices the decision to implement the collection also of net incomes. In 2008, the year of our analysis, several countries, like Germany, Denmark, Finland, Netherlands, Norway, and the U.K., do not display data on net incomes. Gender differences may be slightly overestimated where measurement is based on gross incomes because of the inclusion of taxes and social security contributions (OECD, 2012). For example, second earners, who are often women, will be subject to different tax thresholds than their partners in many countries.

© 2014 International Association for Research in Income and Wealth

TABLE 1

Employment Rates, Annual Earnings, and Average Weekly Hours of Work by Gender and Relative Gap in the European Countries Covered by EU-SILC, 2008

	Country	Sample Size	Employment Rate			Earnings (Euros*)			Hours of Work			Hourly Earn. Gap (%)
			Women	Men	Gap (%)	Women	Men	Gap (%)	Women	Men	Gap (%)	
At	Austria	8,298	59.8	77.7	23.1	41,619	28,770	30.9	44h30′	35h10′	20.8	12.8
Be	Belgium	9,352	59.1	70.9	16.7	50,284	36,012	28.4	42h45′	33h50′	20.9	9.5
Bg	Bulgaria	7,392	59.3	72.8	18.5	4,278	3,266	23.6	47h45′	45h40′	4.1	20.4
Cy	Cyprus	6,214	61.0	78.8	22.6	29,555	19,917	32.6	44h25′	37h55′	14.6	21.1
Cz	Czech Republic	16,900	57.5	76.8	25.1	13,688	9,828	28.2	45h25′	41h05′	9.6	20.5
Dk	Denmark	8,928	68.0	77.3	12.0	54,229	41,464	23.5	40h40′	35h55′	11.5	13.6
Ee	Estonia	7,911	70.5	78.2	9.9	13,698	9,313	32.0	41h30′	39h50′	3.8	29.3
Fi	Finland	16,594	68.6	75.2	8.8	43,770	33,027	24.5	41h15′	36h50′	10.5	15.7
Fr	France	15,138	63.2	71.8	12.0	39,636	27,824	29.8	42h10′	35h20′	16.2	16.2
De	Germany	17,677	61.1	73.3	16.6	37,732	21,486	43.0	44h05′	33h25′	25.5	23.5
Gr	Greece	9,992	51.0	76.5	33.3	26,918	19,260	28.4	47h35′	38h30′	17.5	13.3
Hu	Hungary	14,077	52.0	65.9	21.1	9,322	8,157	12.5	42h15′	40h10′	4.9	8.0
Is	Iceland	5,380	84.1	89.8	6.3	63,074	41,119	34.8	48h40′	38h50′	20.1	18.4
Ie	Ireland	6,978	58.1	70.8	18.0	47,708	31,802	33.3	42h10′	30h30′	27.5	8.1
It	Italy	32,420	47.6	73.0	34.8	37,776	28,549	24.4	42h45′	35h30′	16.9	9.1
Lv	Latvia	7,671	67.4	77.4	12.9	10,989	8,625	21.5	44h15′	41h40′	5.8	16.7
Lt	Lithuania	7,288	66.8	74.1	9.8	7,935	5,787	27.1	41h35′	39h30′	5.1	23.1
Lu	Luxembourg	6,486	56.9	77.9	26.9	67,469	41,421	38.6	43h40′	34h55′	19.9	23.3
Nl	The Netherlands	16,090	63.2	76.1	16.9	52,238	29,369	43.8	40h05′	28h00′	30.1	19.6
No	Norway	8,478	72.3	82.8	12.6	65,738	42,285	35.7	42h30′	35h55′	15.7	23.7
Pl	Poland	24,094	54.1	69.9	22.6	9,551	7,889	17.4	44h50′	40h00′	10.8	7.4
Pt	Portugal	7,141	63.2	75.8	16.7	19,335	14,940	22.7	42h50′	38h25′	10.1	14.0
Ro	Romania	12,033	52.5	72.3	27.4	4,387	3,289	25.0	43h00′	40h40′	4.8	21.3
Sk	Slovakia	11,149	61.5	75.3	18.3	9,684	7,511	22.4	43h15′	40h00′	7.7	15.9
Si	Slovenia	20,051	58.1	68.4	15.1	20,968	19,242	8.2	42h15′	40h25′	4.5	3.9
Es	Spain	22,383	57.4	79.1	27.4	27,001	20,488	24.1	43h00′	37h05′	13.6	12.2
Se	Sweden	11,187	75.6	81.7	7.4	42,061	30,802	26.8	36h55′	30h10′	18.3	10.4
UK	United Kingdom	10,971	69.5	80.6	13.7	43,650	24,800	43.2	42h40′	32h40′	23.4	25.8

Notes: *For non-Euro area countries, earnings are converted into Euros.
Authors' calculation on weighted data from EU-SILC 2008, rev.4. Cross-sectional weights used.

© 2014 International Association for Research in Income and Wealth

losses from self-employment including royalties, and value of goods produced for personal consumption. An individual can earn her labor income from multiple activities. Therefore an individual can earn simultaneously wages and self-employment income.

Most European countries are characterized by a high incidence of part-time employment and flexibility in working hours among women, which should be accounted for in order to fully evaluate gender gaps in the labor market. Hours of work are the number of hours a person normally works in a week in her job. In the EU-SILC survey, this covers all hours including extra hours, whether paid or unpaid, which the person normally works, but does not cover travel time to work and main meal breaks. If multiple jobs are held, respondents are asked to indicate the number of working hours in the main job and in the subsidiary jobs. The main job is the one with the greatest number of hours usually worked. Some respondents, particularly the self-employed and family workers, may not have usual hours, in the sense that their hours vary considerably from week to week or month to month. When the respondent is unable to provide a figure for usual hours for this reason, the average of the hours actually worked per week over the past four weeks is used as a measure of usual hours (Eurostat, 2009).

Table 1 also reports annual earnings and average weekly hours of work by gender in the European countries, along with the corresponding gender percentage gaps. The last column displays the percentage gender gap of hourly earnings. Hourly earnings are measured as the ratio of annual earnings over the estimated number of annual working hours. Total annual earnings provide a better idea of how much individuals "take home" (OECD, 2002), while hourly wage is the price of labor and consequently can be considered the most appropriate measure of the gender pay gap.

For all countries, women's earnings are lower than men's, in terms of both hourly earnings and annual earnings, in line with the literature (e.g., Blau and Kahn, 1996, 2003). Women's hourly rates of pay are, on average, 16.3 percent less than men's, with large variation across countries. The gap in terms of annual earnings gap is more marked, 28.1 percent, reflecting also the gap in working hours that is equal to 14.1 percent. The gender gap in hourly earnings is less than 10 percent (which roughly corresponds to the mean minus one standard deviation of the gap distribution) in Belgium, Hungary, Ireland, Italy, Poland, and Slovenia, while Estonia, Germany, Lithuania, Luxembourg, Norway, and the U.K. report a gap greater than 23 percent (roughly the mean plus one standard deviation of the distribution).

For OECD countries, several studies (Barth *et al.*, 2002; OECD, 2002; Olivetti, 2008) have observed a strong negative correlation between wage and employment gaps, that is, countries with the lowest wage gaps are accompanied by the highest employment gap. As argued in Olivetti and Petrongolo (2008), such negative correlation reveals a different selection process into employment across countries. Countries with a higher rate of female labor participation are those with a high gender wage gap since the additional participants are essentially women with low labor market "attachments" (low skills and low work experience). Instead, low wage gaps are found in countries where mainly women with high-wage characteristics enter the labor force. When we extend the analysis to extra

© 2014 International Association for Research in Income and Wealth

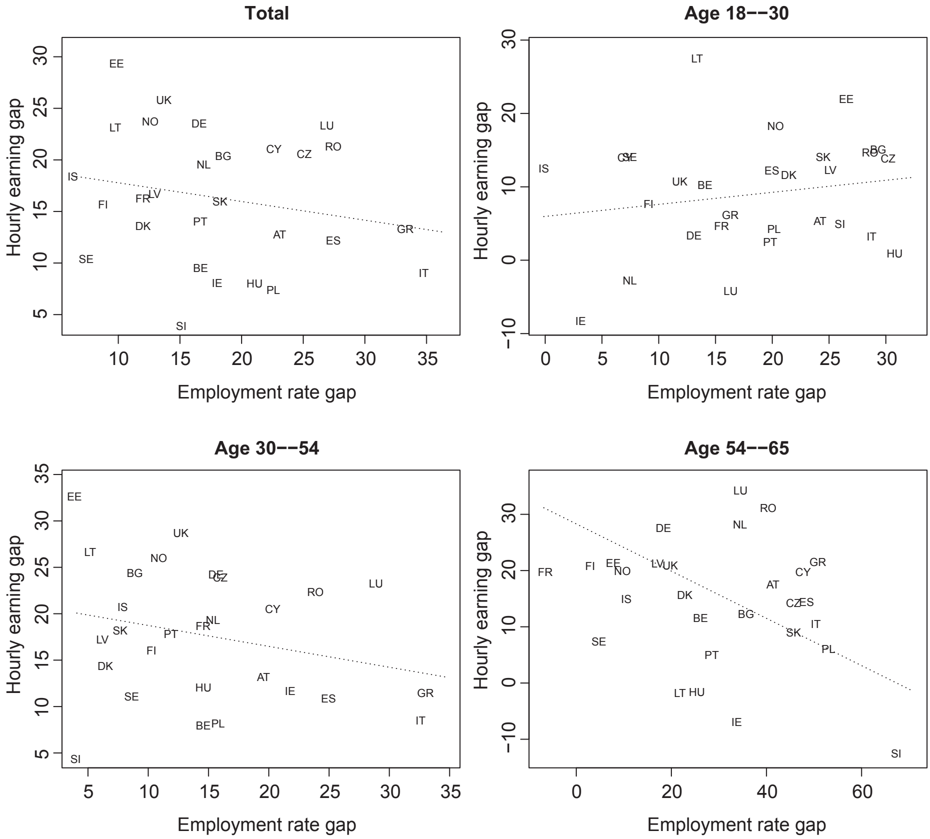

Figure 1. Gender Gaps Across European Countries 2008: Hourly Earnings Gap versus Employment Rate Gap

Note: A small amount of jittering (Gelman *et al.*, 1995) is added to the points, to make any multiplicities transparent. The dotted lines are regression lines fit to the observed data.

OECD countries, we find a different picture, at least in Europe. Figure 1 (top left panel) shows a small negative correlation between wage and employment gaps. However, when we split the working population by age groups, correlation between employment rate and hourly wage gaps varies according to age: it becomes positive for the youngest group aged 18–30 (top right panel), slightly negative for workers aged 30–54 (bottom left panel), and negative for 54–65 year-olds (bottom right panel). This evidence suggests no selection effect for young generations entering the labor market, while selection occurs as age increases, becoming substantial in the pre-retirement period, probably due to early exits from employment of low-skilled female workers. Similar conclusions can be drawn when we substitute hourly pay gap with annual earnings gap. It is also interesting to discuss the relationship between employment rate and hours worked gaps (Figure 2) and how it varies by age group. For the total population there is no correlation between gaps in employment and in hours of work. However, this

© 2014 International Association for Research in Income and Wealth

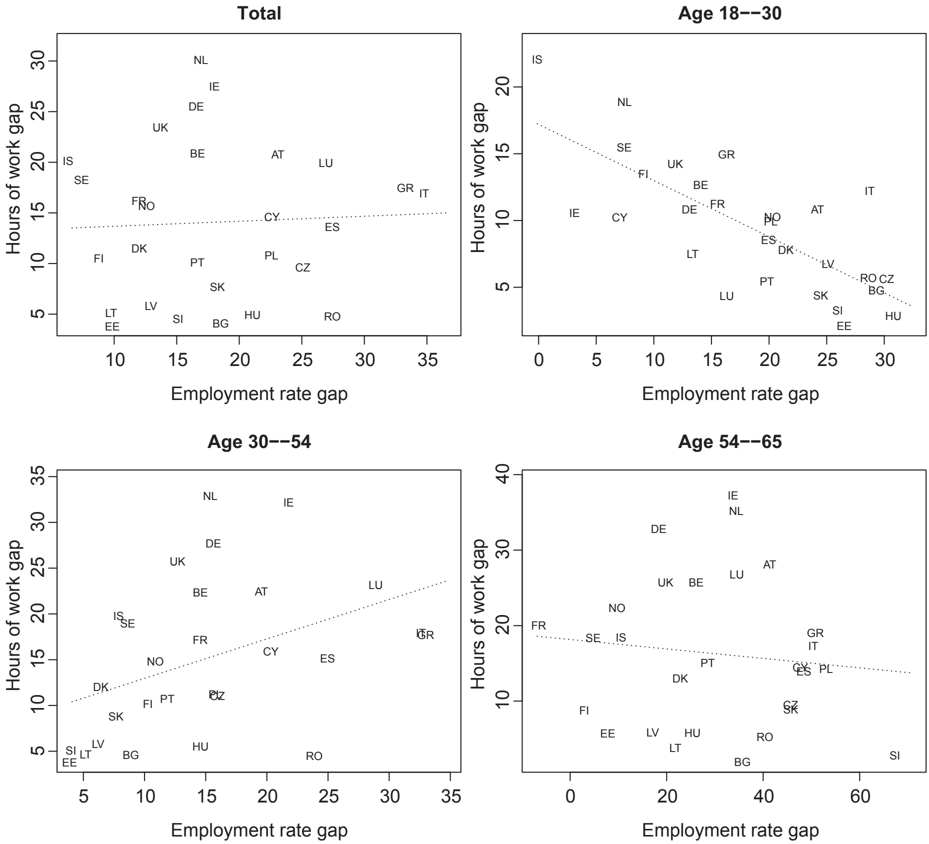

Figure 2. Gender Gaps Across European Countries 2008: Gap in Hours of Work versus Employment Rate Gap

Note: A small amount of jittering (Gelman *et al.*, 1995) is added to the points, to make any multiplicities transparent. The dotted lines are regression lines fit to the observed data.

absence of correlation hides a strong negative correlation for younger workers and a strong positive correlation for the 30–54 year-old group, indicating different behaviors of male and female workers as they age. These preliminary results confirm the importance of age in shaping labor market dynamics. Our analysis seeks to understand the different ways in which age affects labor market outcomes across European countries. It is well-documented that in many areas age does not have a monotonic relationship with other variables (Deshpande and Suresh, 1990). Therefore, for each country and for each labor market outcome, age profiles are analyzed by concentration curves that make it possible to estimate local changes in regression slopes, within the Gini regression framework. The next section outlines the main features of our method.

3. A Brief Review of the Method

Our method uses concentration curves within a Gini regression framework (Yitzhaki and Schechtman, 2013) to study the relationship between two variables

© 2014 International Association for Research in Income and Wealth

Y and X through visual tools. One of the main advantages of these curves is the possibility to graphically check whether the relationship between Y and X is monotonic over the entire range of the explanatory variable X. More generally, it allows one to distinguish within the overall regression coefficient the contribution of different sections of the explanatory variable.

3.1. *NLMA Curves and Gini Regression*

A concentration curve can be viewed as an extension of the Lorenz curve to include a two-variable case. The idea behind it is similar to that of the absolute or generalized Lorenz curve (Shorrocks, 1983), but in this case the horizontal axis and the vertical axis represent two different variables. Similarly to the presentation of Gini by a Lorenz curve,[2] it is possible to use a concentration curve, the so called (N)LMA curve, to graphically present the Gini regression coefficient. The NLMA curve is the Normalized LMA (Line of independence Minus the Absolute concentration curve) curve, defined as the the Line Of Independence (LOI) minus the Absolute Concentration Curve (ACC) of Y with respect to X. The ACC portrays on the horizontal axis the cumulative distribution of the explanatory variable X, and on the vertical axis the correspondent expected cumulative values of the dependent variable Y; while the LOI describes the expected cumulative values of Y when X and Y are independent. The properties of the ACC are useful in several areas.[3] In the area of regression there are two potential uses for concentration curves. The first is to investigate how the association between two variables changes along the range of the explanatory variable, and the second use is to study the weighting scheme of the regression curve (see Heckman *et al.*, 2006, for the derivations of the weighting schemes for many different econometric models). In this paper we concentrate on the first use.

Let $g(x) = E(Y|X = x)$ be the conditional expectation of Y given X, that is the regression curve. The ACC of Y with respect to X can be formally defined as:

$$\text{ACC}(p) = \int_{-\infty}^{x_p} g(t)dF(t), \tag{1}$$

where x_p is implicitly defined by $p = F(x_p)$, and F is the cumulative distribution function of X.

In case of independence between Y and X the ACC curve collapses to the LOI equal to:

$$\text{LOI}(p) = \mu_Y \cdot p \tag{2}$$

where μ_Y is the expected value of the dependent variable Y.

[2]For an exhaustive survey of Gini's contributions, see Giorgi (2005). This article also provides a picture of the lively debate about the scientific contributions of Gini and other Italian scholars over the last century.

[3]For a comprehensive set of definitions and properties of the Absolute Concentration Curves, see Yitzhaki and Olkin (1991). For their use in performing statistical analyses compatible with stochastic dominance in the area of tax reforms and in finance, see Yitzhaki and Schechtman (2013).

© 2014 International Association for Research in Income and Wealth

The LMA curve is defined as the difference between these two curves:

(3) $$\text{LMA}(p) = \text{LOI}(p) - \text{ACC}(p).$$

The properties of the LMA curve that are useful in our context are the following:

(a) The LMA curve starts at (0, 0) and ends up at (1, 0). It can take any shape depending on properties of $g(x) = E(Y|X = x)$.

(b) The derivative (slope) of the LMA curve with respect to p is equal to:

$$\frac{d(\text{LMA})}{dp} = \mu_Y - g(x_p).$$

As a consequence, the LMA curve has a positive (null, negative) slope if μ_Y is greater (equal, smaller) than $g(x_p)$.

(c) The second derivative of the LMA curve with respect to p is equal to:

$$\frac{d^2(\text{LMA})}{d^2 p} = -\frac{d(g(x_p))}{dp} = -\frac{d(g(x_p))}{dx_p} \cdot \frac{d(x_p)}{dp}.$$

As a consequence, the LMA curve is concave at p (convex, straight line) if and only if $\frac{d(g(x_p))}{dx_p}$ is greater (smaller, equal) than 0, being $\frac{d(x_p)}{dp}$ always positive.

(d) The area between the LMA curve and the horizontal axis is equal to cov (Y, F(X)) (see Yitzhaki, 2003, for a formal proof). This property implies a direct relationship with the Gini regression coefficient, defined as (Olkin and Yitzhaki, 1992):

(4) $$\beta^G = \frac{\text{cov}(Y, F(X))}{\text{cov}(X, F(X))}.$$

(e) If the LMA intersects the horizontal axis, then there exist monotonic increasing transformations of X, that can change the sign of the regression coefficient in an OLS regression but not in a Gini regression.

By dividing the LMA curve by cov (X, F(X)) the area enclosed between the curve and the horizontal axis becomes equal to the Gini regression coefficient. This curve is referred to as the NLMA curve. The covariance cov (X, F(X)) is (one fourth of) the well known Gini's mean difference (GMD) introduced by Gini in 1912.

The upper panel of Figure 3 illustrates the NLMA curve for a linear regression curve, while the bottom panel shows the NLMA for a piecewise regression curve. If Y is a linear function of X, the LMA curve looks like a bell curve with the peak at the median. On the left-hand side of the figure the curve is increasing, an indication that the dependent variable is below its average, while on the right-hand side of the figure the curve is declining, which indicates that Y is above its average. Since the linear relationship is always positive, the total regression coefficient is

© 2014 International Association for Research in Income and Wealth

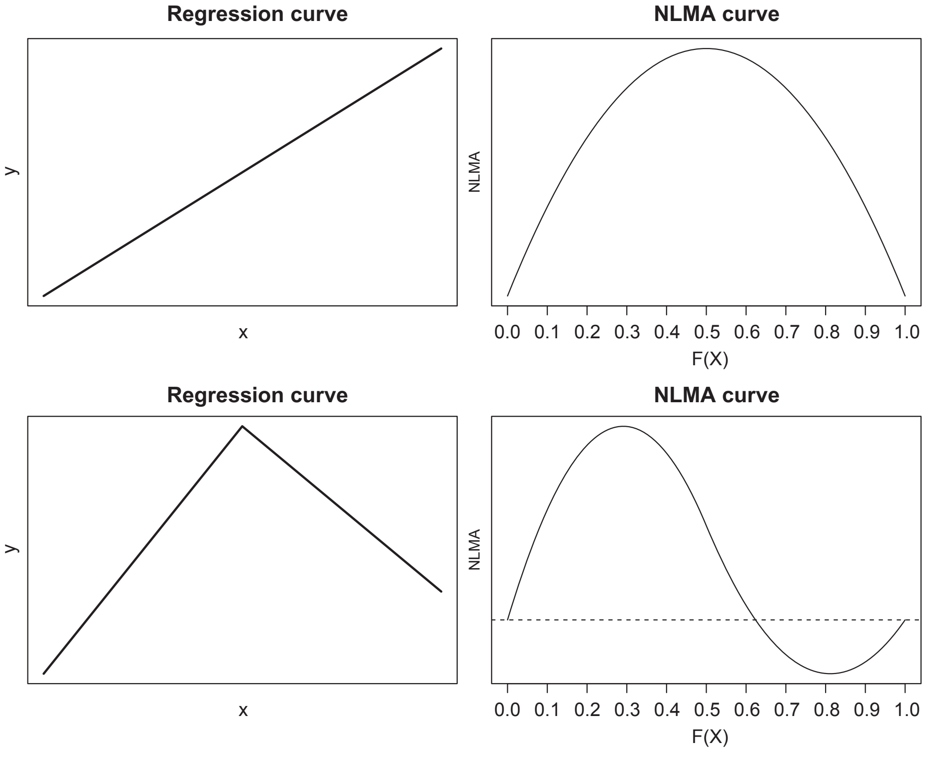

Figure 3. Examples of NLMA Curves

positive and the NLMA curve is always above the horizontal axis. A linear regression curve with a negative slope will result in a mirror image NLMA curve that will be always below the horizontal axis.

The bottom panel presents instead a segmented linear regression curve in which the first local regression coefficient is positive and the second is negative. This behavior is reflected in the corresponding LMA curve that changes curvature from concavity to convexity. The curve crosses the horizontal axis, therefore the sign of cov(Y, F(X)) (and consequently the sign of the regression coefficient) depends on the magnitudes of the areas above and below the horizontal axis. The inspection of the NLMA curve allow for the decomposition of the Gini regression coefficient into intra- and inter-group components as we will illustrate in the next section.

3.2. *Making Up the Regression Coefficients*

Assume that observations are partitioned into M disjoint (non-overlapping) groups, according to different levels of X, denoted by $m = 1, \dots, M$. Let $\pi_m = n_m/n$ be the relative size of group m and $\bar{Y}_m$ and $\bar{X}_m$ the group's averages. The Gini

© 2014 International Association for Research in Income and Wealth

regression coefficient β^G of the overall population can be decomposed as follows (Yitzhaki and Schechtman, 2013):[4]

$$\beta^G = \sum_{m=1}^{M} w_m \beta_m^G + w_B \beta_B^G. \tag{5}$$

This decomposition is based on four types of components: the group's weight (w_m), the group's regression coefficient (β_m), the between-group weight (w_B), and the between-group regression coefficient (β_B). The first term of the right-hand side of equation (5) can be interpreted as the intra (within)-group component, while the second term is the inter (between)-group component.

The term w_m is the contribution of group m to the overall variability and it is defined as $w_m = \pi_m \dfrac{G_m}{G}$. The term G_m is (one fourth of) the GMD of X in group m while G is (one fourth of) the GMD of X. The term w_B is defined as

$$w_B = \frac{\operatorname{cov}(\bar{X}, \bar{F}(X))}{\operatorname{cov}(X, F(X))},$$

where $\bar{F}(X)$ is the vector of the M averages of the ranks of the members of the groups when they are ranked according to the distribution of X. The coefficients β_m^G and β_B^G are defined as follows:

$$\beta_m^G = \frac{\operatorname{cov}_m(Y, F(X))}{\operatorname{cov}_m(X, F(X))};$$

$$\beta_B^G = \frac{\operatorname{cov}_B(\bar{Y}_m, \bar{F}(X))}{\operatorname{cov}_B(\bar{X}_m, \bar{F}(X))}.$$

A similar decomposition also holds for the OLS regression coefficient. However, we focus on the Gini regression given its direct relationship with NLMA curves.[5]

4. Analysis of Age Profiles by Gender

4.1. *Evidence from the NLMA Curves*

Appendix A provides, for each country, men's and women's estimated NLMA curves by age for the following labor economic variables: employment rate, hours of work, and annual and hourly earnings. At the risk of oversimplifying, we classify countries according to similar patterns of the NLMA curves. It turns out that with few exceptions, similarity in age patterns reflects geographical

[4]The decomposition of the OLS regression coefficient follows the same structure, except that the weights are derived from the decomposition of the variance.

[5]Note also that a monotonic transformation of the explanatory variable, X does not affect F(X). Therefore, unlike the OLS, a monotonic transformation of X cannot change the sign of the Gini regression coefficient but only may affect its magnitude (see last property of LMA curves).On the properties of the Gini regression and its advantages, see Yitzhaki and Schechtman (2004).

© 2014 International Association for Research in Income and Wealth

TABLE 2

Classification of the Countries in Terms of Their Age Profiles Characteristics

Group	Countries	Age Profiles Characteristics
1	Denmark Iceland Norway Sweden	Labor participation profile for men and women almost identical. Very different profiles in hours of work for men and women: inverse U-shape for men, M-shape for women. Similar earnings profiles: increase until late forties and settle later on.
2	Cyprus Greece Italy Portugal Spain	Similar profile between men and women for earnings (almost linear). Very different behavior in participation and hours of work.
3	Austria Belgium Czech Republic France Germany Ireland Luxembourg Netherlands Slovakia United Kingdom	Similar age profiles in employment rate. Large differences in terms of hours of work between men and women. Earnings positively correlated with age for men and weakly correlated for women.
4	Estonia Finland Latvia Poland Romania Slovenia	Similar age patterns by gender in all labor market outcomes.
5	Bulgaria Hungary Lithuania	Similar age patterns by gender in employment and hours of work. Female age profiles in earnings characterized by a dip during the thirties.

proximity. Table 2 shows the classification of the countries along with the main characteristics in terms of age profiles.

The first group of countries includes Sweden, Norway, Denmark, and Iceland, all the Nordic countries but Finland. Although some differences exist, this group is characterized by similar estimated employment profiles. Employment increases with age for both men and women with a peak at around 20 percent of the age distribution (corresponding to late twenties/early thirties depending on the country) and essentially plateauing after that until it starts decreasing at around 85 percent, which roughly corresponds to the pre-retirement age. Wage profiles, both in terms of total income and in terms of pay per hour, are also similar between men and women. The NLMA curves indicate that men's and women's earnings are expected to increase until their mid to late forties, after which they tend to settle. Obviously it is difficult to establish how age itself affects earnings because of convolution of age, cohort, and selection effects. Therefore, any evidence can be interpreted only as suggestive of relevant association rather than to reach firm conclusions. With the exception of Iceland, the remarkable difference between men and women in this group occurs in NLMA curves of weekly hours of work. Men's working hours over life course tend to assume an inverse U-shaped distribution. This is revealed by the alternance between concave and convex regions. On

© 2014 International Association for Research in Income and Wealth

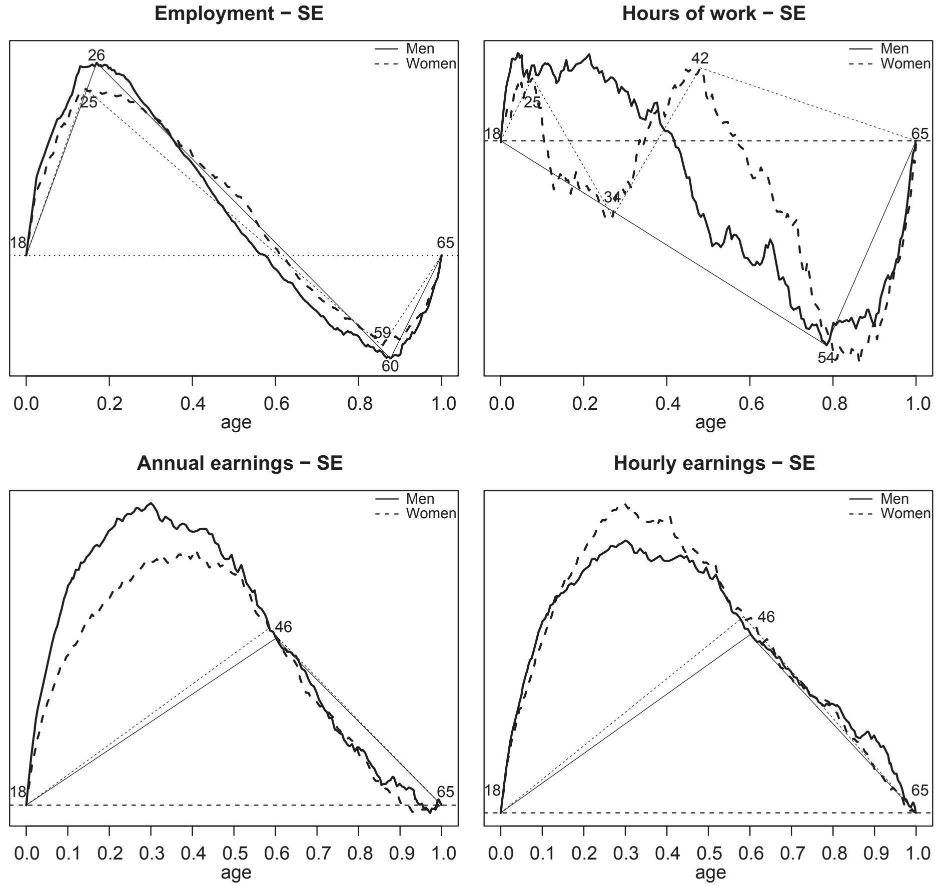

Figure 4. NLMA Curves of Labor Market Outcomes by Age and Gender in Sweden

the other hand, women's hours of work show striking variation over the life cycle, eventually forming an M-shaped distribution that reflects the decrease in working hours during periods of childbirth and childbearing.

Among the countries belonging to the first group, we selected Sweden. The choice of one country among others is not related to any theoretical fundamental but rather has the aim of enhancing the potential of the method and, in this vein, Sweden can be considered a representative country. Figure 4 shows NLMA curves of labor market variables by gender in Sweden. In each plot we added dotted lines for both men and women in order to connect the points which identify sub-intervals that make up the overall regression coefficient. The decomposition of the Gini regression coefficients will be explored in the next section.

The second group of countries comprises Italy, Spain, Portugal, Greece, and Cyprus. A common characteristic of these countries is the presence of a similar profile between men and women in regard to earnings, in contrast with a very different behavior in employment and hours of work. We picked Italy as the

© 2014 International Association for Research in Income and Wealth

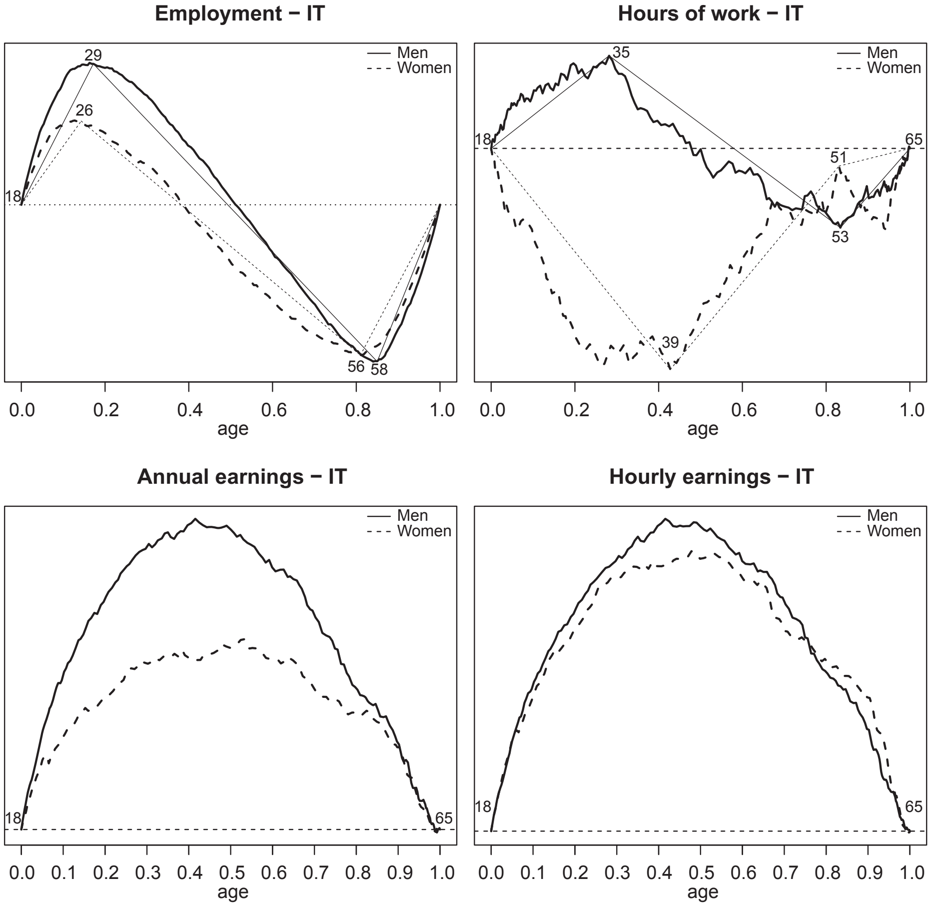

Figure 5. NLMA Curves of Labor Market Outcomes by Age and Gender in Italy

representative country for this group. Figure 5 shows the NLMA curves for this country. Although the pattern in employment for men and women looks the same, there are some important differences. To some extent, men seem to follow the "turning points" of women: for women aged 18–26 and 56–65 employment rate is below average, while male employment rate is below average for men aged 18–29 and 58–65 (ages that vary little among the countries in this groups). For middle-aged workers employment rate is quite stable for men (the NLMA curve is almost linear) and decreasing for women (the NLMA curve is slightly convex). This trend is particularly prominent in the case of Spain and Greece. The behavior of women in terms of hours of work is essentially the reverse of that of men. The estimated effect of age on hours of work is positive for men aged less than 35 (corresponding to the 25th percentile). After that, work effort starts to gradually decrease until the age of 50–55. For women the estimated effect of age is negative until the age of 35–40. After this, hours worked for women rise until pre-retirement age. The inverse U-shape of both the earnings NLMA curves indicates that men's and women's earnings are expected to constantly increase until the age of retirement,

© 2014 International Association for Research in Income and Wealth

even though the upwards sloping is steeper for men than for women. This evidence may suggest that both male and female workers are protected by seniority rules, union bargaining, and employment protection legislation. An alternative explanation would be that low-wage employees tend to quit working at earlier age. These age–earnings profiles are in line with the theoretical framework developed by Lazear (1979), according to which the worker is initially underpaid—that is wage lower than her marginal productivity—but later on in the working life the worker is overpaid. Such a delayed compensation contract discourages workers from "shirking" and this is true for both men and women. The decomposition of the Gini regression coefficients for Italy will be presented in the following section.

France, Germany, the Netherlands, the United Kingdom, Luxembourg, Belgium, Austria, and, to a lesser extent, Ireland, the Czech Republic, and Slovakia belong to a third group. Generally, labor participation profiles for men and women are almost identical (with the exception of Ireland and the Czech Republic). Male and female employment rate increases with age, peaking at around 20 percent of the age distribution and remaining almost constant for middle-aged workers and then gradually decaying. Age appears to be more of a discriminating factor for working hours, where it amplifies large differences between male and female profiles. In terms of hours worked, men show small correlation with age, with the exception of France where a significant positive correlation is evident. With little differences among countries, men in this group work less than average when they are young, increase their work time until the early thirties, and slightly decrease it until retirement. The variability in female working hours by age is more relevant: women work more than average when they are young. There is evidence of a significant reduction in working hours for women in the thirties, which probably reflects cultural attitudes toward child rearing in these countries, as witnessed by the convexity in the curve. An increase in working hours is present for women in the 35–55 age group, followed by a substantial reduction for older female workers. The main feature of this group occurs in earnings profiles. The resultant picture for earnings shows them to be positively correlated with age for men and weakly correlated for women. Germany was selected as the representative country of this group (see Figure 6) and the Gini regression decomposition was performed on it.

Another group of countries characterized by similar age patterns by gender in all labor market variables is comprised by Poland, Latvia, Estonia, Slovenia, Finland and, Romania. The patterns are perhaps best illustrated by the case of Poland (see Figure 7). Male age profile for employment is close to women's profile and the same is true of the hours of work pattern. Male earnings are characterized by an initial rise for the lower 40 percent observations (up to age 35 in Poland) and then a gradual fall off. Female earnings behave in a somewhat similar manner, while showing a less substantial increase. For these countries, annual and hourly earnings profiles are in accordance with the profiles in working hours, with high positive correlation between wages and hours worked. This empirical evidence can be related, at least for the Eastern countries, to the employment structure, which is still more oriented toward manufacturing rather than advanced services.

The last group comprises Lithuania, Bulgaria, and Hungary. The distinctive feature of this group is essentially related to the earnings profile, which shows, for men, an increase until the thirties and then a gradual drop until retirement. In the

© 2014 International Association for Research in Income and Wealth

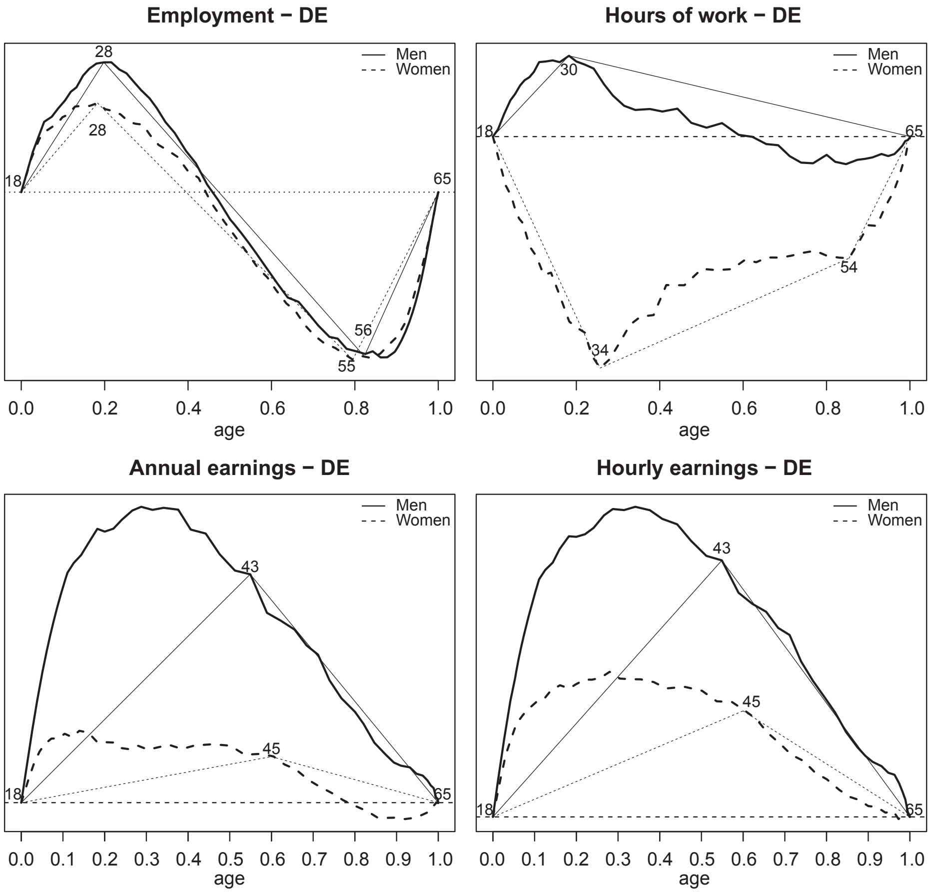

Figure 6. NLMA Curves of Labor Market Outcomes by Age and Gender in Germany

case of women instead we have an initial increase, as shown by the concavity in the NLMA curve. The curvature then goes from concave to convex, which indicates a "dip" associated with the childbearing period, which is followed by a slight increase or by stability, until women leave the labor market. Lithuania is the representative country of this group. Figure 8 shows NLMA curves for Lithuania. As we can see, there is similarity in the employment patters of men and women. This similarity becomes more visible when we ignore minor differences due to small sample sizes in the NLMA curves.

4.2. *Decomposition of the Gini Regression Coefficients*

In this section we try to "translate" what we have learned from the NLMA curves into the decomposition of non-overlapping age groups and their contribution to the overall Gini regression coefficients.[6]

[6]To check the robustness of our results we also run OLS regressions. The signs of all the regression coefficients are identical and the values do not differ significantly.

© 2014 International Association for Research in Income and Wealth

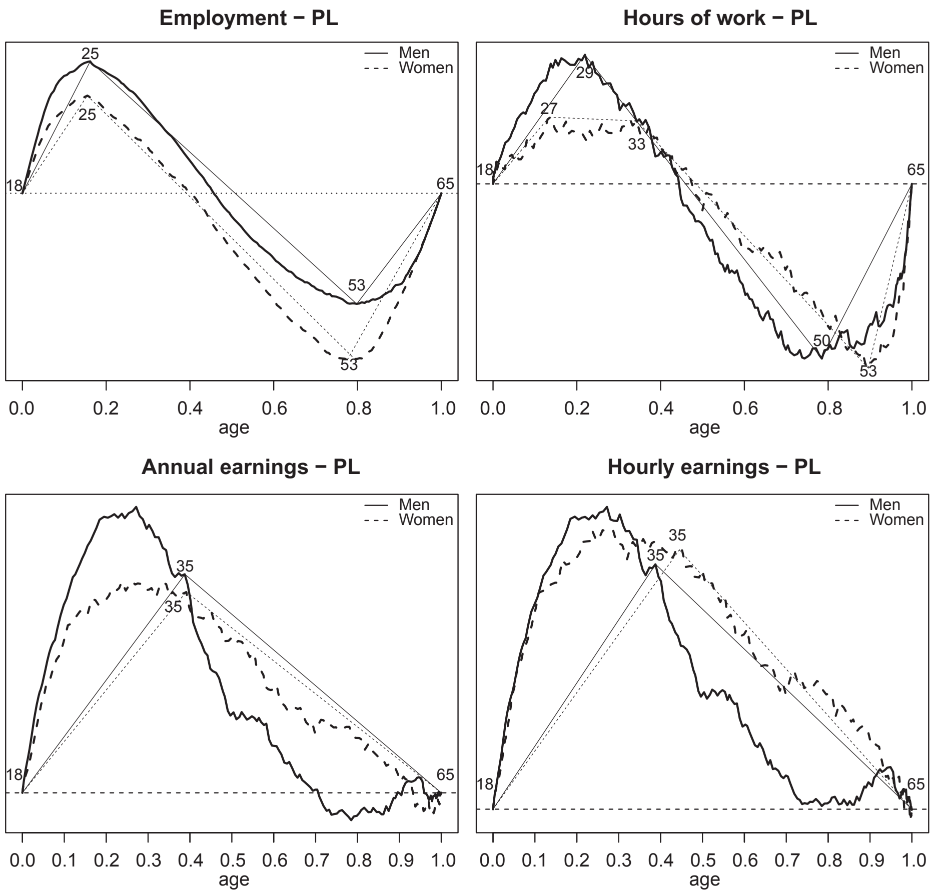

Figure 7. NLMA Curves of Labor Market Outcomes by Age and Gender in Poland

We report the Gini decomposition, along with the ingredients for the regression "make up," for the five representative countries.[7]

Table B.1 in Appendix B presents the contributions of different age groups to the Gini regression coefficient for Sweden, according to the grouping identified by the NLMA curves as shown in Figure 4. In terms of employment rate, the expected annual increase is 7.4 percent for men and 6.2 percent for women until the ages of 26 and 25, respectively, and employment rate is below average. The estimated β coefficient is equal to –0.14 for men aged 26–60 and equal to 0.18 for women aged 25–59, indicating a rather stable employment rate within those ranges, which is above average. After 60, the rapid increase of the NLMA curves (see upper left panel in Figure 4) translates into negative coefficients for both men and women accompanied with below average employment rate. The estimated β coefficients of

[7]A similar analysis has been performed for all the European countries but is not reported for the sake of space. The results do not significantly differ across countries of same groups and are available upon request.

© 2014 International Association for Research in Income and Wealth

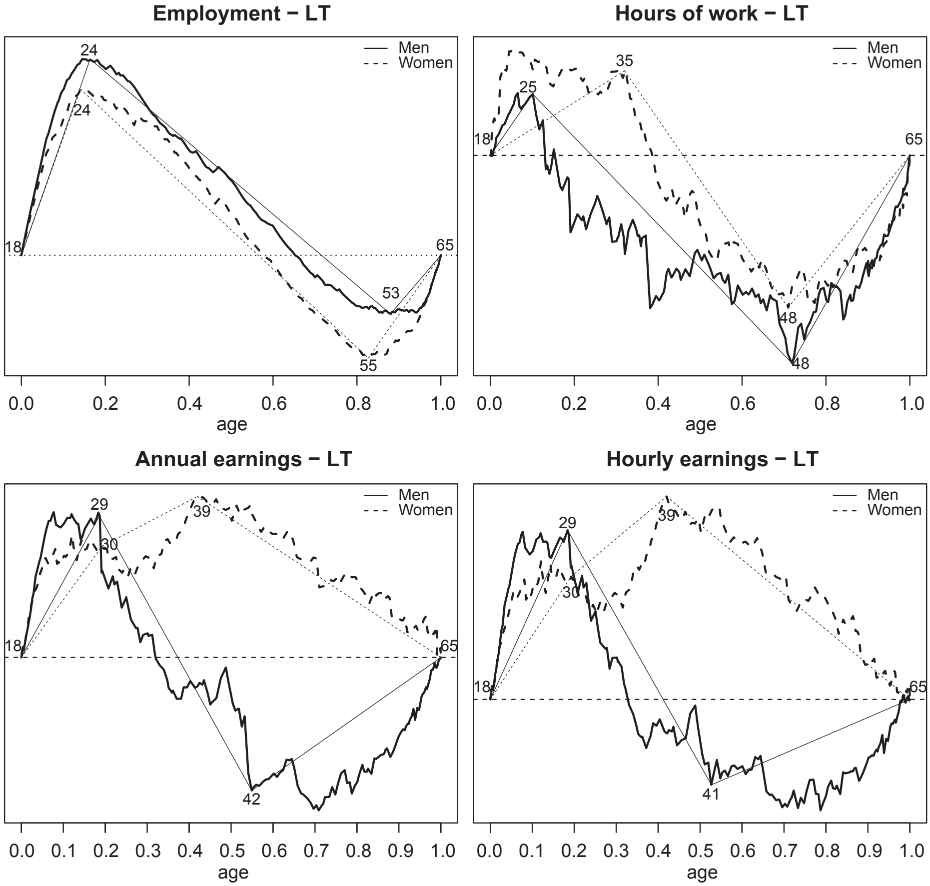

Figure 8. NLMA Curves of Labor Market Outcomes by Age and Gender in Lithuania

weekly hours of work is equal to 0.054 for men aged less than 54 years (corresponding to 78 percent of the entire number of working men) and equal to –0.348 for the older group. This roughly means that men increase their effort at work by less than 5 minutes per year until the age of 54; after that the time weekly spent on work decreases at a pace of approximately 20 minutes per year. The Gini decomposition for women confirms the M-shape already detected by the NLMA curve corroborated by estimated alternative signs in the regression coefficients. The decomposition of annual earnings clearly shows an expected annual increase of 948 euros for men and 693 euros for women until the age of 46; after that age annual earnings for both men and women start to decrease at a rate of of 237 and 309 euros per year, respectively. In terms of hourly earnings the expected increase of 0.50 cent is similar for men and women until the age of 46, followed by stability until retirement.

The Gini decomposition for Italy is presented in Table B.2. The most relevant features are: a significant reduction of 0.25 percent per year in the employment rate

© 2014 International Association for Research in Income and Wealth

for women aged 26–56; the almost specular behavior of men and women in terms of working hours; and the linearity of annual and hourly earnings for both men and women.

The third group of European countries is represented by Germany, whose Gini regression decomposition is shown in Table B.3. What characterizes this group is essentially a different pattern for men and women in terms of working hours and earnings. On average, employed men annually increase their weekly working hours by about 20 minutes per year until the age of 30. After that, they slightly reduce their working hours by 5 minutes a week per year. Women instead reduce their weekly hours at a rate of about 15 minutes per year until the age of 34, after which they increase it by about 6 minutes a week per year until the age of 54 and rapidly decrease it after that. Men's annual earnings increase by 1520 euros per year until the age of 43 in contrast with an increase of only 270 euros for women. A similar behavior is detected for hourly pay. Moreover, after the mid forties the reduction in annual earnings is much more pronounced for women than for men: 388 euros per year versus 113.

Poland represents the fourth group of European countries described above. Gini decomposition for Poland is shown in Table B.4. It confirms the close association between the annual earnings pattern and the working hours pattern for both men and women. Young men tend to increase their work hours by 15 minutes a week per year with a consequent increase in their annual earnings of 425 euros per year until the age of 35. For young women, the increase in working hours is less pronounced (around 5 minutes a week) and thus the annual earnings are expected to increase less (about 250 euros per year). In contrast, age profiles for hourly pay of men and women are more similar, and the same holds for age–employment profiles.

Lithuania was selected as representative for the last group. As shown in Table B.5, the pattern of male annual earnings is compatible with the age–working hours profile. For men, an initial increase in both weekly hours and earnings when they are in their twenties is followed by a substantial stability until their forties and a moderate decrease after that. Women, instead, increase their earnings until 30 by a rate of about 300 euros per year, but in the 30–39 range their expected annual earnings decrease by 220 euros per year. Similarly, the estimated negative coefficient for hourly earnings ($\beta_{30-39} = -0.119$) implies that each additional year of age corresponds to an expected reduction of 12 cents an hour. A slight increase in earnings is estimated as women reach 40 years old until retirement, despite a reduction in terms of hours of work.

5. Conclusions and Further Research

In the present paper we have used concentration curves within a Gini regression framework to represent European gender differences in labor market age profiles. The NLMA curves for each key variable in the labor market showed sections where the curves were convex and sections where they were concave. The Gini regression enabled us to quantify the contribution of each section to the overall regression coefficient. This connection between Gini regression coefficients and concentration curves allowed us to verify the monotonicity of regression

© 2014 International Association for Research in Income and Wealth

curves between age and employment rate and working hours, annual and hourly earnings. The search for a monotonic relationship is important because conclusions that are based on non-monotonic regression curves may be sensitive to the selection of the range of variables in the model and to monotonic transformation of the variables.

We analyzed the European labor market age file in all the 28 countries covered by EU-SILC. These countries were grouped based on similar behavior for a better interpretation. We detected five different groups according to their age patterns in the labor market. Generally country classification reflects geographical proximity with few exceptions: for example, Finland does not belong to the group of the other Nordic countries, and Lithuania is not in the same group with the other two Baltic states. For each group, a representative country was selected. For the selected countries, we first estimated the NLMA curve and then identified age sections that showed positive (or above average) slopes in the regression curve and sections with negative (or below average) values, so that adding up the slopes yielded the appropriate Gini regression coefficient. As a result we were able to identify turning points in labor market variables according to age, to outline age profiles according to gender and check whether patterns differed between men and women and/or across countries. Our results show that both gender and age carry important implications for labor market outcomes. Gender is a crucial factor differentiating participation in the labor market among workers, although employment–age profiles do not substantially differ between men and women in almost all the European countries. The most relevant differences in age profiles concern working-hours patterns: some countries are characterized by an almost specular behavior in men and women; other countries instead show similar patterns. Generally, earnings increase with age for both men and women. However, local regression coefficients are not monotonic over the entire age range and can even be locally negative in some countries.

The method used in this paper allows one to decompose a simple regression coefficient. The main advantage of this method relies on graphical tools provided by the NLMA curves. Obviously, it is limited to bivariate analysis. As pointed out by Yitzhaki and Schechtman (2013), the method can be extended to deal with multiple regression. The regression coefficients in a multiple-regression framework are derived as solutions of sets of linear equations, with the simple regression coefficients serving as parameters in this set of equations. However, due to interaction effects, complications and complexities arise. In fact, transformations applied to an independent variable may also affect other regression coefficients signs either directly (i.e., through the change in the sign or magnitude of the simple regression coefficient) or indirectly (through the change in the correlations with other regression predictors). Multiple Gini regression represents a challenging topic for further research.

References

Barth, E., M. Røed, and H. Torp, *Towards a Closing of the Gender Pay Gap*, European Commission Report, Oslo, 2002.

© 2014 International Association for Research in Income and Wealth

Blau, F. D. and L. M. Kahn, "Wage Structure and Gender Earnings Differential: An International Comparison," *Economica*, 63, 29–62, 1996.
———, "Gender Differences in Pay," *Journal of Economic Perspective*, 14, 75–99, 2000.
———, "Understanding International Differences in the Gender Pay Gap," *Journal of Labor Economics*, 21, 106–44, 2003.
Deshpande, J. V. and R. P. Suresh, "Non-Monotonic Ageing," *Scandinavian Journal of Statistics*, 17, 257–62, 1990.
European Commission, *A Roadmap for Equality Between Men and Women 2006–2010*, EC, Brussels, 2006.
Eurostat, *Eurostat Community Statistics on Income and Living Condition*, Study Documentation, Luxembourg, 2009.
Gelman, A., B. P. Carlin, H. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman & Hall, London, 1995.
Giorgi, G. M., "Gini's Scientific Work: An Evergreen," *Metron*, LXIII, 299–315, 2005.
Heckman, J., S. Urzua, and E. Vytlacil, "Understanding Instrumental Variables in Models with Essential Heterogeneity," *Review of Economics and Statistics*, 88, 389–432, 2006.
Lazear, E. P., "Why Is There Mandatory Retirement?" *Journal of Political Economy*, 87, 1261–84, 1979.
Lee, S., D. McCann, and J. C. Messenger, *Working Time Around the World: Trends in Working Hours, Laws and Policies in a Global Comparative Perspective*, International Labor Organization, Geneva, 2007.
OECD, *Employment Outlook*, OECD, Paris, 2002.
———, *OECD Family Database*, OECD, Paris, 2012.
Olkin, I. and S. Yitzhaki, "Gini Regression Analysis," *International Statistical Review*, 60, 185–96, 1992.
Olivetti, C., "Gender and the Labour Market: An International Perspective and the Case of Italy," *Rivista di Politica Economica*, V–VI, 3–32, 2008.
Olivetti, C. and B. Petrongolo, "Unequal Pay or Unequal Employment? A Cross-Country of Gender Gaps," *Journal of Labor Economics*, 25, 621–54, 2008.
Shorrocks, A. F., "Ranking Income Distribution," *Economica*, 50, 3–17, 1983.
Yitzhaki, S., "Gini's Mean Difference: A Superior Measure of Variability for Non-Normal Distributions," *Metron*, LXI, 285–316, 2003.
Yitzhaki, S. and I. Olkin, "Concentration Indices and Concentration Curves," in K. Mosler and M. Scarsini (eds), *Stochastic Orders and Decisions under Risk*, IMS Lecture Notes-Monograph Series, 19, 380–92, 1991.
Yitzhaki, S. and E. Schechtman, "The Gini Instrumental Variable, or the 'Double Instrumental Variable' Estimator," *Metron*, LXII, 287–313, 2004.
———, "Identifying Monotonic and Non-Monotonic Relationships," *Economics Letters*, 116, 23–25, 2012.
———, *The Gini Method: A Primer on a Statistical Method*. Springer, New York, 2013.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix
Table 1: Gini decomposition: Sweden
Table 2: Gini decomposition: Italy
Table 3: Gini decomposition: Germany
Table 4: Gini decomposition: Poland
Table 5: Gini decomposition: Lithuania

© 2014 International Association for Research in Income and Wealth