European Research Studies,
Volume XVII, Issue (3), 2014 *pp. 91-102*

Is There a Heteroscedasticity due to Gender in the Adult Learning? A View from Europe

Marco Biagetti [1], Sergio Scichittano[2]

Abstract:

In this article the cross section nature of the European Union Statistics on Income and Living Conditions (EU-SILC), the new homogenized European panel survey, is exploited to check for unobserved heteroskedasticity due to gender in the adult learning. We focus our analysis on a whole sample of 21 European countries. We find a significant heteroskedasticity due to sex in the adult learning in Europe.

Key Words: *Heteroskedasticity, Adult Learning, Human Capital, Europe*

JEL Classification : *I20, J24*

1 Ministry for Economic Development, Department for the Development and the Economic Cohesion, via Sicilia 162, 00187, Rome, Italy.
2 Corresponding author. E-mail: sergio.scicchitano@dps.gov.it.it, Tel: +39-06-47619923, fax +39-06-42086946

1. Introduction

In this article we check for unobserved heteroskedasticity due to gender in adult learning in Europe. It is largely accepted by the economic literature that human capital is a continuous process starting at school and keeping being diffused in the labour market through adult learning. Indeed, skills can be accumulated not only before getting a job, through pre-occupational education but also during working life by fostering continuous learning and/or training.

Although the European Jobs Strategy's emphasizes adult education during working life, empirical literature, by focusing on the growth effects of the initial education, does not seem to take sufficiently into account the contribution of workforce adult learning as an additional source of human capital and growth. As for the determinants of adult learning, some empirical regularities have been found: young and better educated workers, involved in highly-skilled occupations and in large firms enjoy greater learning opportunity[2]. These findings can be easily defined as stylized facts.

However, there is no accepted evidence of which gender is more likely to receive any adult learning. When the training definition is considered, some papers (i.e. Bassanini et al. 2007) show that being female is associated with a higher probability of being involved in training. Arulampalam et al. (2004) find these results in 4 countries; conversely, in the other 6 countries there is not a significant difference between males and females. Oppositely, Pischke (2001) estimates that men in Germany are more likely to access to training. When considering a broader learning activity, Drewes (2008) finds that female are more likely to participate in educational programs, but less likely to take training courses. Thalassinos et al. (2009) have analysed gender inequalities in shipping. For the UK, in Jenkins et al. (2002) females are six percentages points more likely to undertake lifelong learning, while Sargant et al. (1997) show that men are more likely to be involved in training and education. Also Simonsen and Skipper (2008) find that men and women have different enrollment patterns: women are more likely to attend basic or post-secondary training courses, whereas men are more likely to get enrolled in vocational ones.

Our results shows that in Europe there exists a significant unobserved hetetoskedasticity due to gender in the adult learning. In the empirical model we used the European Union Statistics on Income and Living Conditions (EU-SILC),

[2] See in particular Sargant et al. (1997), Jenkins et al. (2002), Ok and Tergeist (2003), Arulampalam et al. (2004) and Bassanini, et al. (2007) and Drewes (2008).

the new homogenized European panel survey. Since 2005[3] EU-SILC has succeeded the European Community Household Panel (ECHP): there, the number of countries is increased, indicators are updated, and common guidelines, definitions, and procedures are used.

This empirical paper is organized as follows. The next section describes the data. Section 3 reports the results of the empirical model. In the last section we present our main conclusions.

2. The Data

Our data are from the 2005 first wave, of EU-SILC, the new homogenized panel survey that has replaced ECHP. Similarly to the ECHP, EU-SILC is an attractive source of information because it adopts the same "community" questionnaire used by the national data collection units in each included country, which obviously makes comparisons across nations easier.

EU-SILC has three main advantages with respect to other similar datasets. Firstly, the set of economies is fully comparable. This desirable feature is obtained through the use of common guidelines, definitions and procedures. Secondly, all the old and the new European member states are surveyed, while the ECHP covered only 14 economies. The new dataset, thus gives information on many of the so-called new entrants. Finally it updates the whole of the indicators. Indeed, our comparison involves 21 European member states whose labor market institutions and adult learning systems are known to have different characteristics.

Each wave includes a household and a personal file. In the 2005 wave 197,657 nationally representative households and 422,040 individuals from EU-25 countries were interviewed. The use of a cross-sectional survey is not likely to do any harm in terms of bias, as the greatest part of the variables used in this paper (and of the whole dataset) are time invariant. This means there is no chance to exploit useful

3 Indeed, EU-SILC was launched gradually between 2003 and 2005 in all EU Member States and has become the source of data for the analysis of income distribution and social inclusion at EU level. More precisely, EU-SILC was first brought out in 2003 on the basis of a gentlemen's agreement in six Member States (Belgium, Denmark, Greece, Ireland, Luxembourg and Austria) as well as in Norway. In 2004, under Regulation N° 1177/2003 of the EP and European Council, EU-SILC was implemented in twelve EU-15 countries (Germany, Netherlands and the United Kingdom delayed the launch for one year) as well as in Estonia, Iceland and Norway. In 2005, EU-SILC was operating in all EU-25 countries, plus Iceland and Norway, all with available cross-sectional data. Bulgaria, Turkey and Romania launched EU-SILC in 2006, and Switzerland followed suit in 2007. Former Yugoslav Republic of Macedonia and in Croatia are evaluating its start as well.

information from the time dimension, making a panel data approach quite redundant.

We have focused our analysis on 21 countries, the only ones having available data for our interest variables. The countries are: Austria (AT), Belgium (BE), Cyprus (CY), Czech Republic (CZ), Germany (DE), Denmark (DK), Estonia (EE), Spain (ES), Finland (FI), France (FR), Greece (GR), Hungary (HU), Italy (IT), Lithuania (LT), Luxembourg (LU), Latvia (LV), Netherlands (NL), Portugal (PT), Slovenia (SI), Slovak Republic (SK), United Kingdom (UK).

In our study we only consider workers aged between 16 and 64, who are employed full or part-time according to their current self defined economic status. Thus, we choose to drop unemployed and retired individuals[4], pupils, students, people with unpaid work experience, permanently disabled or/and unfit to work, people in compulsory military community or service, those fulfilling domestic tasks and caring responsibilities and other inactive persons. After doing so, 103,588 people remained, 54,796 of which are men and 48,792 women.

The observed dependent variable (Adult learning) is binary, taking value one if the individual is currently involved in some learning (education or training) program defined under ISCED-97 as *"an array or sequence of educational activities, which are organised to accomplish a pre-determined objective or a specified set of educational tasks*" (UNESCO, 1999, p. 5). Unfortunately we cannot observe the specific typology of learning because this variable covers regular education and training systems which are normally intended to lead to a certification recognized by national authorities qualifying for a specific education/program[5]. All we can observe with these data is whether or not an individual participates in any formal adult learning process, over its own working-life cycle. Thus, similarly to Simonsen and Skipper (2008) and Drewes (2008), we empirically identify adult learning as education and training formal enrolment over the entire working life-cycle.

4 Early retired too are included in this definition: consequently they have been kept out of this empirical survey.

5 The individual's participation in this program may be on a full-time attendance basis, a part-time attendance one or by correspondence. The variable also includes modules (short programs/courses) which may be part of a longer regular education program and are taken and completed, giving to their graduates the corresponding academic credit, independent of whether the person continues to complete it fully or not. The level of the short programs/courses will be the same as the program they form part of. Furthermore, if the interviewed individual is enrolled as a student or an apprentice in a program within the regular educational system the answer will be 1. For apprentices who are in a period of only 'on-the-job training' or alternate 'on-the-job' and 'in-school learning' within the framework of an alternate (e.g. dual) program, the answer is coded 1 as well, since the person is enrolled in a qualifying scheme. The following adult programs cannot be classified using ISCED-97: i) vocational education organized by a firm without leading to an official award or certification ii) any non-formal education without leading to an official award or certification iii) individual cultural activities for leisure.

In the regressions we include among the explanatory variables: age (*age*) and squared age (*age squared*[6]); a dummy for marital status (*Marital status)* which equals 0 for workers who were never married, or got separated, became widow or divorced in 2005, and 1 if they were married in 2005); a dummy for the level of education attained (*Past education*=0 for workers with at most upper secondary education, 1 at least post secondary non tertiary education).

All these variables are included in the individual-specific group. More specifically marital status and the level of education attained can be determined by the respondent, while the other regressors of that group cannot be controlled by her/him. Further, we insert a dummy for self defined current economic status (*full time contract*, taking value 0 for individuals working full time, and 1 if they work part time), another for the type of contract (*permanent contract* which equals 0 if workers signed a permanent contract, and 1 if they signed a temporary one[7]) and a third one for recent job changes (*job change* which equals 0 if workers did not change job since last year, 1 otherwise). These three dummies are included in the job-specific group. Finally we build two dummies for the local unit size (*Unit size*=0 if it has between 0 to 10 persons, 1 between if it comprised between 11 to 49, 2 if in the local unit there are 50 persons and more) and one for the type of worker's sector of activity (*High-skilled*=0 if un-skilled, 1 if skilled). These are included in the firm specific group.
The summary statistics of these variables for both men and women are shown in table 1. Some of these need a better explanation, pertaining to EU-SILC definitions.

[6] As is well known from the classical econometrics, the squared of continuous independent variables is a method to control for possible non linear relations between the dependent and the squared regressors.

[7] EU-SILC makes clear that in the majority of Member States most jobs are based on written work contracts. However in some countries such contracts exist only for specific cases (for example in the public sector, for apprentices, or for other persons undergoing some formal training within an enterprise). Taking account of these different institutional arrangements the notions "temporary job" and "work contract of limited duration" (likewise "permanent job" and "work contract of unlimited duration") describe situations which under different institutional frameworks can be perceived as similar. A job may be regarded as temporary if it is understood by both the employer and the employee that the termination of the job is determined by objective conditions such as reaching a certain date, completing an assignment or the return of another employee who has been temporarily replaced. In the case of a work contract of limited duration the condition for its expiration is generally mentioned in the contract. To be included in these groups are: i) persons with a seasonal job ii) persons engaged by an employment agency or business and hired out to a third party in order to execute a "work mission" (unless there is a work contract of unlimited duration with the employment agency or business) iii) persons with specific training contracts. If there exists no objective criterion for the end of a job or work contract these should be regarded as permanent or of unlimited duration. What is involved is the actual employment being time-limited under an agreement - not that he/she has, for example, considered stopping work in order to travel or attend college. Respondents who have a contract to do their job, which may be renewed, for example, once a year, should be coded according to whether or not the respondents themselves consider their job to have an indefinite term.

Table 1. Summary statistics of the variable used in the model

	Men				Women			
	Mean	S. D.	Min	Max	Mean	S. D.	Min	Max
Adult Learning	0.05	0.21	0	1	0.06	0.24	0	1
Age	41.39	11.36	16	64	41.19	10.84	16	64
Marital Status	0.64	0.48	0	1	0.62	0.49	0	1
Past education	0.28	0.45	0	1	0.36	0.48	0	1
Full-time contract	0.05	0.22	0	1	0.24	0.43	0	1
Permanent contract	0.13	0.34	0	1	0.15	0.36	0	1
Job change	0.08	0.27	0	1	0.08	0.28	0	1
Unit size	0.94	0.86	0	2	0.94	0.85	0	2
High-skilled	0.38	0.48	0	1	0.42	0.49	0	1

The *full-time contract* variable captures the individual's own perception of her main activity at present. It differs from the ILO concept to the extent that people's own views of their main status differ from the strict definitions used in the ILO manuals. For instance, many people who would regard themselves as full-time students or homemakers may be classified by the ILO as employed if they have a part-time job. Similarly, some people who consider themselves unemployed may not meet the strict ILO criteria of taking active steps to look for a job and being immediately available. The self-declared main activity status is, in principle, determined on the basis of the most time spent, but no criteria have been specified explicitly.

The distinction between full-time and part-time work should be made on the basis of a spontaneous answer given by the respondent. It is impossible to establish a more exact distinction between part-time and full-time work, due to variations in working hours between Member States and also between branches of industry. By checking the answer with the number of hours usually worked, it should be possible to detect and even to correct implausible answers, since part-time work will hardly ever exceed 35 hours, while full-time work will usually start at about 30 hours.

The "local unit" to be considered is the geographical location where the job is mainly carried out or, in the case of itinerant occupations, can be said to be based. Normally it consists of a single building, part of a building, or, at the largest, a contained group of buildings. The "local unit" is therefore the group of employees working for the enterprise who are geographically located at the same site[8].

[8] An enterprise executes one or more activities at one or more locations. The EU-SILC dataset clearly states that the local unit is "an enterprise or part thereof (e.g. a workshop, factory, warehouse, office, mine or depot) situated in a geographically identified place. At or from this place economic activity is

The *highly skilled* variable refers to the main job (current main job for people at work or last main job for people who do not have a job). If multiple jobs are held or were held, the main job should be the one with the greatest number of hours usually worked. The variable is coded according to the ISCO-88 classification. In this paper we identify as unskilled the occupations between 1 to 34 according to ISCO-88 classification. Occupations between 41 to 93 are identified as skilled.

3. The Empirical Model

In the model we control for possible heteroskedasticity of error variance across genders which may cause parameter estimates to be biased, inconsistent and inefficient (Yatchew and Griliches 1985).

Indeed, the more heterogeneous career patterns for women is a widely recognized fact in labour economics and econometrics (see for instance Allison 1999, Williams 2009). That is to say unmeasured variables affecting adult learning - like attitudes, abilities or wish to learn - may have a different impact on women and men.

As a matter of fact, Arulampalam et al. (2004) recognize that the probability of training for men may differ from that of women for several reasons: for instance, the former gender is more likely to be offered training opportunities by their employers, either because men are less likely to quit or due to discrimination. Their analysis is carried out in a panel data framework: after the estimation of a RE model, they use Bayesian techniques to estimate for each individual of each gender the unobserved individual specific component.

Here we take a different route, partly compelled by the cross-sectional nature of our data. Indeed, as is widely known, in a probit model the residual variance is assumed to be Var (ε)=1 while in the logit model it is set to $\pi^2/3$. That is equivalent to say that in binary regression models coefficients are inherently standardized. More precisely, while in the OLS model the standardization is implemented by rescaling

carried out for which - save for certain exceptions - one or more persons work (even if only part-time) for one and the same enterprise". Further explanatory rules include the following. i) A geographically identified place is interpreted on a strict basis: two units belonging to the same enterprise at different locations (even within the smallest administrative unit of the Member State) are regarded as separate local units. ii) If a person works in more than one place or at home, the local unit is taken to be place from which instructions emanate or from where the work is organized. The concept of local unit relates to the operational definition of the establishment in ISIC Rev.3 as follows. A single local unit may carry out, in a unique location, more than one kinds of activities. The operational definition of the establishment corresponds to the local kind-of-activity unit (local KAU), i.e. the part of the enterprise KAU which corresponds to a local unit. As for the definition of enterprise KAU, it "groups all the parts of an enterprise contributing to the performance of an activity at class level (four digits) of NACE (REV 1.1) Rev. 1 and corresponds to one or more operational subdivisions of the enterprise".

all variables to have a variance of 1, in a probit or logit model it is accomplished by scaling the variables and residuals so that the residual variances are 1 or $\pi^2/3$ (Long and Freese 2006):

$$\Pr(y_i = 1) = \Phi\left(\frac{x_i \beta}{\sigma}\right) \quad (1)$$

If (as it is normally assumed) $\sigma=1$, we get the usual homoskedastic probit. But if we are in presence of heteroskedasticity in the residual variance, a problem arises in modeling a possible equation variance to get rid of biases in the estimates.
Thus, following Alvarez and Brehm (1995) we model the following equation to account for a possible heteroskedasticity:

$$\Pr(y_i = 1) = \Phi\left(\frac{x_i \beta}{\exp(z_i \gamma)}\right) \quad (2)$$

where $\exp(z_i\gamma) = \sigma_i = f(z_i)$ and z's are a set of regressors.

In table 2, we report the results for a) the homoskedastic unweighted form; b) the homoskedastic weighted specification; c) the heteroskedastic model. The second differs from the first because the inverse of the probability for an individual to be included in the sample due to the survey design has been used. In the third, following Allison (1999), a variance equation $\sigma_i = \exp(sex_i\gamma)$ - i.e. one depending only on the gender regressor - is built up, because, as it has been shown in the literature and briefly reported above, unmeasured variables correlated to lifelong learning decisions may be strongly affected by gender differences.

Table 2 shows the marginal effects (computed at the mean values of the regressors) measuring the change in the probability of adult learning for an infinitesimal change in each independent, continuous variable and the change in the probability for discrete changes in dummy variables measured with respect to the base.

We find that individual characteristics are statistically significant predictors for the inception of an adult learning activity, although the magnitude of the effects is relatively small: the results show that young, better educated and unmarried workers are more likely to receive adult learning. By examining job characteristics, workers with temporary and part time contracts, who did not change job in the last year show a significant higher probability to get in adult learning. Firm specific characteristics are also relevant because workers in small local units and in low-skilled occupations are less likely to undertake adult learning. The relationship between unit size and learning probability is also monotonic: the predicted odds to get in adult learning

are, in particular, higher for workers in large local units compared to workers employed in medium local ones. It should be also noted that being a part-time and temporary employee, having a skilled occupation are the features with the strongest effects.

Concerning heteroskedasticity, the results show that, on the one hand, in the homoskedastic weighted specification (see table 5, column 2) there is no huge change in any explanatory variable compared to the homoskedastic unweighted model (see tables 2, column 1) except for the gender variable which becomes statistically not significant. In particular the homoskedastic weighted model shows that women seem to be significantly more likely to undertake adult learning. On the other hand, while no variation can be noted in any other regressor, the corresponding sign for gender on the heteroskedastic model becomes negative and quite significative with respect to the homoskedastic weighted specification (see table 2, column 3). Following this specification, gender heteroskedasticity appears a serious issue for adult learning incidence.

Table 2. Regression results. Dependent variable: Adult Learning participation. Un-weighted homoscedastic model, weighted homoscedastic model and heteroscedastic model.

	Unw. Homosk . [1]		W. Homosk.[1] [2]		Het. (gender) [3]
Gender	0.085***		0.016		-0.106**
	[5.65]		[0.71]		[-1.67]
Age	-0.122***		-0.166***		-0.173***
	[-25.60]		[-24.94]		[-22.76]
Age squared	0.001***		0.001***		0.001***
	[17.75]		[19.68]		[18.57]
Marital status	-0.138***		-0.109***		-0.104***
	[-8.11]		[-4.07]		[-3.71]
Past education	0.034*		0.102***		0.108***
	[2]		[3.78]		[3.80]
Full-time contract	0.232***		0.208***		0.215***
	[11.57]		[6.66]		[6.50]
Permanent contract	0.350***		0.564***		0.586***
	[19.32]		[22.1]		[20.46]
Job change	-0.087***		-0.218***		-0.227***
	[-3.69]		[-5.81]		[-5.73]

Unit size: 11-49	0.102***		0.066**		0.070**
	[5.38]		[2.24]		[2.26]
Unit size: 50 or more	0.120***		0.132***		0.140***
	[6.55]		[4.63]		[4.67]
High-skilled	0.424***		0.312***		0.327***
	[23.50]		[11.59]		[11.35]
Constant	0.924***		1.582***		1.857***
	[10.86]		[14.42]		[13.57]
lnsigma2_gender					0.083**
					[2.21]
Country effects	Yes		Yes		Yes
Log pseudolikelihood	-17813.838		-18139.887		-17810.54
Wald chi2(11)	6580.16		3192.88		2614.56
Prob > chi2	0.000		0.000		0.000
Pseudo R-sq	0.1559		0.2041		
Number of obs.	103.588		103.588		103.588

Notes: values of *z* statistics in brackets. Probability weighting used in columns 2 and 3 entails the use of robust standard errors
Significance: * p<0.10, ** p<0.05, *** p<0.01.
[1]Observations are weighted by EU-SILC personal cross-sectional weights.

4. Conclusion

In this article we have shown possible confounding effects of the heteroskedasticity due to gender in adult learning in Europe, by using the latest European dataset, EU-SILC - which has never been used yet, to our knowledge in studying this topic - for a sample drawn from individuals of 21 European countries. Indeed gender is the most likely source of heteroskedasticity as suggested by the empirical literature. We model it through a particular variance equation. The outcome of our computations gives further insight to the role of the unobserved heteroskedasticity due to gender. Its effect on the probability to take on adult learning can even be so strong to reverse the sign of the coefficient associated to the same variable in the main probit model.
The policy implications revealed by our findings are relevant for the soundness of measures implemented in the labour market. Indeed, should heteroskedasticity due to gender be not well accounted for, those policy tools would end up being very poorly addressed and misguided, mainly when it comes to the apportionment of limited financial resources.

References

Allison P., 1999. Comparing Logit and Probit Coefficients across Groups, *Sociological Methods and Research*, 28(2), 186-208

Alvarez R. M., and Brehm J., 1995. American Ambivalence towards Abortion Policy: Development of a Heteroskedastic Probit Model of Competing Values, *American Journal of Political Science*, 39, 1055-1082

Arulampalam W., Booth A., Bryan M., 2004. Training in Europe, *Journal of the European Economic Association*, 2 (2-3), pp. 346-360

Bassanini A., Booth A. L., Brunello G., De Paola M., and Leuven E., 2007. Workplace Training in Europe, in Brunello, Garibaldi and Wasmer (eds.), Education and Training in Europe, Oxford University Press, Ch. 8 - 13

Drewes T., 2008. Adult Education and Training from a Longitudinal Perspective, Trent University

Jenkins A., Vignoles A., Wolf A., Galindo-Rueda F., 2002. The determinants and effects of lifelong learning', CEEDP Centre for the Economics of Education, London School of Economics and Political Science, London, UK, 19.

Long J., Freese S & J., 2006. Regression Models for Categorical Dependent Variables Using Stata', Stata Press, Second Edition, College Station, Texas,

Ok W., Tergeist P., 2003. Improving Workers' Skills: Analytical Evidence and the Role of the Social Partners, OECD Social, Employment and Migration Working Papers

Sargant N., Field J., Francis H., Schuller T., Tuckett A., 1997. The Learning Divide: A Study of Participation in Adult Learning in the United Kingdom, NIACE

Simonsen M., Skipper L., 2008. The Incidence and Intensity of Formal Lifelong Learning', Economics Working Papers 2008-07, School of Economics and Management, University of Aarhus.

Thalassinos, I.E., Stamatopoulos, V.T. and Arvanitis, E.S. (2009), "Gender Wage Gap: Evidence from the Hellenic Maritime Sector 1995-2002," *Bulletin of Economic Sciences, Transylvania University of Brasov, Vol. 2**(51),*** 197-210.

Yatchew A., Griliches Z., 1985. Specification Error in Probit Models, *Review of Economics and Statistics*, 18, 134-139

Williams R., 2009. Using Heterogeneous Choice Models to Compare Logit and Probit Coefficients across Groups', Sociological *Methods & Research*, 37(4), 531-559.