University of Lisbon

Lisbon School of Economics & Management

Master in Applied Econometrics and Forecasting

2016/2017

ESTIMATING GENDER DIFFERENCES IN THE PROBABILITY OF

UNEMPLOYMENT: EVIDENCE FROM PORTUGAL

Author:

Joana Monteiro Passinhas

Advisor:

Isabel Proença

ABSTRACT

By Joana M. Passinhas

Using a dynamic random effects probit model we estimate the probability of unemployment in Portugal in order to assess gender differences in average partial effects and in unemployment persistence, with data from four waves of the Survey on Income and Living Conditions (ICOR), for the period between 2010 and 2013. The estimation occurs while controlling for unobserved individual heterogeneity and for the "initial conditions" problem, which arises from not knowing the stochastic process which originated the observed state of unemployment. We find strong evidence of persistence in unemployment, with some, although weak, evidence that men suffer more from the negative implications of previous unemployment. Simultaneously, we found evidence of higher probabilities of unemployment for women through a fixed effect that aimed to capture gender discrimination in an unstable labor market. The main contributions of the present work lie in the study of the determinants of the probability of unemployment, which represents a shortage in the current literature in labor economics, during a period of high unemployment in Portugal, and by having a special focus on unemployment persistence and gender discrimination.

RESUMO

Por: Joana M. Passinhas

Através de um modelo dinâmico probit de efeitos aleatórios, estimou-se a probabilidade de desemprego em Portugal de forma a avaliar se existem diferenças entre géneros nos efeitos parciais médios e na persistência do desemprego. Os dados utilizados provêm do Inquérito ao Rendimento e Condições de Vida (ICOR) para o período entre 2010 e 2013. A estimação é feita ao mesmo tempo que se controla pela heterogeneidade individual não observada e pelo problema das condições iniciais, que ocorre pelo fato de não se conhecer o processo estocástico que originou o estado de desemprego observado. Encontrámos forte evidência empírica de persistência do desemprego, e alguma evidência de que esta persistência é mais pronunciada para os homens. Através da inclusão de um efeito fixo especifico para as mulheres, que pretende captar o efeito da discriminação de género num período de instabilidade no mercado de trabalho, concluímos que existe evidência estatística de maior probabilidade de desemprego para as mulheres. Este trabalho tem como principais contributos o estudo dos determinantes da probabilidade de desemprego, que representa uma carência da literatura em economia do trabalho, no fato de o estudar num período de grande desemprego em Portugal, e no especial enfoque que dá à persistência do desemprego e à discriminação de género.

ACKNOWLEGDMENTS

I would like to thank my advisor, Professor Isabel Proença for her important suggestions and support throughout this journey. I am also grateful to Daniel Abreu, Francisco Mendonça and João Cruz for their comments and general contributions in my academic development that led to the present work.

TABLE OF CONTENTS

LIST OF TABLES

1. INTRODUCTION

In 2009, Portugal was facing its biggest government budget deficit to date (9.4% of the Gross Domestic Product (GDP)), one of the highest in all Euro Zone, and the pressure from the European Commission to reduce it, as well as to reduce the public debt (130.4% of GDP), was unbearable. In September of 2010, the Portuguese government announced an austerity package, with measures that focused on public administration pay cuts and raising taxes. In the same year, the deficit went to be the highest in Portuguese history, with a soaring value of 11% of GDP, much higher than what the *Maastricht Treaty* (MT) and the *Stability Growth Pact* (SGP) established (3% of GDP). The year of 2011 was marked by both financial and political instability. With the country near bankruptcy, following the rejection of the Stability and Growth Pact IV (SGPIV) and the consequent resignation of the Prime Minister, the opposition parties asked for financial help to the "Troika" of the European Commission (EC), the European Central Bank (ECB) and the International Monetary Fund (IMF). As part of the deal, the Portuguese government agreed in reducing the public deficit to 3% until 2013 through another package of austerity measures[1].

Over the period of 2010-2013[2] the GDP growth rate was negative, despite the positive growth of 2010 (1.9%) that was greatly influenced by a temporary growth of the private consumption driven by expectation of higher taxes over goods that ultimately led to an anticipation of acquiring durable goods. The falling growth rate of GDP also seemed to affect other economic indicators, with the unemployment rate following an abrupt increasing trend, reaching 16.2% in 2013. As a sub product of the crisis, the differential

---

[1] See Pereira &Weeman (2015) for more on the impact of the Global Financial Crisis in Portugal.
[2] The data referenced in this section was retrieved from Statistics Portugal. Statistical Yearbook of Portugal (2011, 2012, 2013, 2014).

in unemployment rates between men and women actually fell, from - 2.1 percentage points (p.p.) in 2010 to - 0.4 p.p. in 2013. The same did not happen with the participation rate that seemed to be unaffected by the crisis, maintaining the difference between men and women in approximately between 8 to 9 p.p., in all four years. This is consistent with evidence found in Albanesi & Şahin (2013) where the unemployment rate increased more for men than for women during the recent recessions, resulting from gender differences in industry distribution, due to the impact of those recessions on the construction and financial sectors, where the majority of the workforce is male.

As it has been for some time in Portugal, this was also a period marked by a demographic crisis, with the population falling at a -1.3% rate per year mostly because of emigration, especially high quality at working age emigration, and with a continued upward trend of the elderly population.

In this background of socio-economic crisis, we will focus on studying differences in the probability of unemployment between genders over a period of declining employment in Portugal, while controlling for relevant exogenous variables and unobserved individual heterogeneity. The causal effect of previous unemployment on current unemployment or the "scaring" effect of unemployment as it is known, will also be studied, especially to see if there is a difference between genders.

Moreover, we will seek to answer whether there are relevant differences in the probability of being unemployed and in the persistence of unemployment between men and women given a set of characteristics, in a period of high unemployment.

The motivation behind this work is twofold. First and foremost, Azmat et al. (2006) found that gender gaps in unemployment rates had risen in the past two decades, in some European countries, especially when the overall unemployment rate was high,

even when the attachment of women to the labor market had simultaneously increased. They conclude that this could be explained by labor market institutions, the impact of human capital differences and from the general easiness of indulging in prejudice against women, consequence of the excess of labor supply. These results reinforce the need to study a possible effect of gender discrimination in the probability of unemployment, especially in this high unemployment period for Portugal.

Secondly, to investigate the causes of the rise in European unemployment, that were perceived as not driven by exogenous shifts on the supply side, Arulampalam et al. (2000) found evidence of a casual effect of an individual's previous unemployment experience on his future labor market condition. Past unemployment may be interpreted by employers as a signal of lower expected productivity, which eventually turns unemployment into a recurring cycle, therefore justifying the special interest in studying the effects of previous unemployment experience in the probability of unemployment.

Consequently, we explore the determinants of unemployment, of discrimination in the labor market and of persistence in unemployment by estimating dynamic panel data models of unemployment that allow to control for the effects of unobserved individual heterogeneity, after controlling for observable characteristics as education level, experience, age and number of kids.

For this purpose, it will be used a random effects dynamic probit model that will use 2010 as the initial condition for unemployment to account for "initial conditions" problem, using data from four waves of the Survey on Income and Living Conditions (ICOR), from 2010 to 2013, provided by Statistics Portugal (*Instituto Nacional de Estatística* – INE). Our results suggest that there is evidence of higher probabilities of

unemployment for women, relatively to men, in spite of women having stronger presence in the higher levels of education.

We were also able to find strong state dependence effects with respect to previous unemployment incidence, during this period of high unemployment in Portugal, and weak evidence that unemployment persistence has a higher effect on the probability of unemployment for men.

The remainder of this paper is set out as follows. Section 2 reviews some important past literature on unemployment, discrimination in the labor market and unemployment persistence, Section 3 presents the econometric model, Section 4 describes the data set and Section 5 presents our estimates and results. The final section concludes and provides suggestions for further research.

2. LITERATURE SURVEY

The present survey of the literature will be divided in two sub-sections. The first aims to familiarize the readers of the relevant literature on theoretical models of labor discrimination while the second focus on providing useful insight on the causal relationships between unemployment and the considered determinants of unemployment.

2.1. THEORETICAL MODELS OF LABOR MARKET DISCRIMINATION

The topic of discrimination in labor markets has been given a lot of attention in the last couple of decades. This attention has created two different approaches on the subject, taste-based discrimination and statistical discrimination.

The goal of this section is to identify the reasons behind labor discrimination, based on the most popular theoretical models.

2.1.1. TASTE-BASED DISCRIMINATION

Becker (1971) defines that someone has a "taste for discrimination" if he or she acts as if he or she was willing to give up on some of his income in order to avoid interacting or to be associated with a member of a certain group.

When employers have a particular distaste in hiring someone from a particular group, say women, the cost of hiring a person from this group can be determined as the sum between the cost in terms of her wage and a measure of the disutility of hiring her. The measure of disutility will be interpreted as a discrimination coefficient, hereinafter referred to as d_c.

The equilibrium happens when wages adjust to $w_M = w_W + d_c$, with w_W and w_M being, respectively, the wage of a woman and of a man, so that the market absorbs all female and male workers. If d_c is left to differ across employers, d_c^j with $(j = 1, \dots, J)$, and $d_c^j \neq 0$, then some employers would only hire men, if $w_M - w_W < d_c^{j_1}$, and some would only hire women, if $w_M - w_W > d_c^{j_2}$, where $j_1 \neq j_2$. Facing a constraint of parity of wages, and assuming that $\forall j \in \{1, \dots, J\}: d_c^j > 0$, then no woman would be employed.

2.1.2. STATISTICAL DISCRIMINATION

In Arrow (1973), employers' discrimination is also considered as a result of their perception of reality, reflecting the way they "perceive", in an unequal manner, expected performance of men relatively to women. This asymmetric perspective of the employers is based solely on which group each individual belongs to, male or female, and will mirror their experience or the social consensus regarding women's average ability relatively to men's, and vice versa. If there is a cost to identify one's

true productivity, the fact that some employer views male workers as having a higher probability of being qualified for the job, will result in higher unemployment rates for women. This problem is intensified with the existence of law enforcements that demand parity of wages, not enabling women to lower their wage rate in order to undertake the costs induced by statistical discrimination.

Phelps (1972) is the first to develop a model meant to explain statistical discrimination. This contribution was useful for building the foundations and inspiring a large number of models on discrimination. Notwithstanding, it did not, as Aigner & Cain (1977) state, focused on explaining economic discrimination as it assumed differences in ability for each group of individuals. For that reason, we will focus on Aigner & Cain's (1977) model for explaining statistical discrimination.

This type of discrimination happens simply because employers can't observe actual marginal productivity of an individual until hiring them. Hence employers are left to make a judgment based on some sort of evaluation criteria (through tests, interviews, and curriculum) which is, inevitably, a noisy signal of an individual marginal productivity. If the reliability of this signal is smaller for women than for men, Aigner & Cain (1977) find a differential in expected productivity between the two groups of people with the same true productivity, which will lead to a preference in employing men followed by an increase of the unemployment rate for women. It could also lead to occupational segregation, just as taste based discrimination. For example, if women are perceived as more likely to quit or to perform poorer in certain types of occupations (e.g. male dominant occupations) then a profit-maximizing employer will be led to favor men over women, for already male dominant occupations, even if both candidates seem equal in all relevant characteristics. Bielby & Baron (1986) found

evidence for this particular result, with employers reserving some jobs for men and others especially for women.

The distinction between these two types of discrimination has particularly important policy implications. For taste-based discrimination, policies should work on raising the cost of engagement in discriminatory behavior while for statistical discrimination policies should focus on providing and improving tools for obtaining information on job candidates.

2.2. UNEMPLOYMENT DETERMINANTS

According to INE, the national statistical institute in Portugal, someone is unemployed if he or she has been searching for work over the past four weeks, is available for work and, currently, doesn't have work. This follows the International Labour Organization (ILO) measure of unemployment, which is meant to include, not the individuals who actively do not want to be in the labor market, but those who want to work but are not able to find any.

From this definition two possible reasons for unemployment emerge. One stemming from productivity problems, where the person's expected productivity is low and, therefore, delivers a marginal product that is expected to be less than the current wage rate. And one situation where the employer would be willing to hire, but the worker is not willing to supply labor at the current wage rate. This implies that both expected and true productivity are important factors that weigh in on both decisions: the employers' decision of hiring or not and the workers' decision to accept, or not, an offer.

The expected productivity of individuals is highly related to the image they give the employer in the process of applying for a job. This image is mostly based on the

candidates acquired skills, the training and education levels that they have, their relevant work experience, and non-human-capital factors such as gender, race, age, if they have children or not and the age of the children.

Moreover, this expected productivity can be very different from the true productivity. It can also be overshadowed by a disutility that the employer might have in hiring from one group of people (taste based discrimination), or by an unjustified perception based on their personal characteristics (statistical discrimination).

2.2.1. GENDER DIFFERENCES IN UNEMPLOYMENT

There are many possible explanations for an existing gender gap regarding unemployment rates. An economic approach suggests examining both sides of the labor market to better evaluate these reasons.

On the demand side, discrimination, be it taste-based or statistical, is the obvious determinant, as we have previously seen, that would foster higher female unemployment rates. One result stated in Jackman (2002) was that countries with overall low unemployment had very similar unemployment rates for men and women, while countries with high unemployment had a higher unemployment rate for women. Both types of discrimination can explain this situation. If an employer has distaste towards women, he finds it easier to hire more men when the supply of labor is larger than demand. Moreover, if women are viewed as less productive and, given that supply exceeds demand, to avoid incurring in any more costs, the employer will be more reluctant to hire women, therefore raising the female unemployment rate.

Notwithstanding, when considering the business cycle, Albanesi & Şahin (2013) found that the unemployment rate increased more for men than for women during recessions, resulting from gender differences in industry distribution, mostly due to

the impact that the recent recessions had on the construction and financial sectors, where the majority of the workforce is male. They also found that the unemployment gender gap was highly affected by the labor attachment of each gender, and that the recent convergence of unemployment rates was related to an increase in labor attachment of women while, simultaneously, the labor attachment of men declined. Şahin et al. (2010) also found evidence of higher unemployment rates for men in the recession of 2007, resulting from men dominating the most affected industries but also from the fact that a higher percentage of men tried to rejoin the labor force but were unable to find a job, therefore, transitioning from inactive status to unemployment status.

On the supply side, human capital accumulation, previous employment status, the number of kids, as well as different personal characteristics may be behind the gender gap in unemployment rates. Azmat et al. (2006) focused on explaining the cross-country differences in the gender unemployment gap. One important conclusion that they arrived was that differences in human capital accumulation were one of the most important part in explaining the flows from employment to unemployment, and vice-versa.

2.2.2. PERSISTENCE IN UNEMPLOYMENT

Arulampalam et al. (2000) estimate dynamic panel data models of unemployment to separate the effects of unobserved individual heterogeneity and unemployment persistence. Their results were consistent with the ''scaring theory'' of unemployment, i.e., that a previous state of unemployment affects future employment status. This effect is found in other studies such as Ahmad (2014) and Arulampalam et al. (2001) who also found that previous unemployment experience brings future unemployment.

Arulampalam (2001) also found that previous unemployment had a negative impact on wages when individuals re-entered the job market. These effects are usually justified by the fact that past unemployment experience might result in depreciation of human capital, therefore raising one's probability of unemployment, and/or by the fact that past unemployment is seen by employers as a signal of low productivity.

They also conclude that unemployment persistence during a demand contraction – during less job opportunities – may be longer because on one hand, perceived average quality of the unemployed is lower and, on the other hand, the poor economic conditions will result in fewer job vacancies being open, which will enhance the demand contraction. Some studies (e.g. Elmeskov & MacFarlan, 1993) have also stated that persistence of higher than usual unemployment could actually be a result of an increase of the natural rate of unemployment and therefore could actually never fully correct itself to the previous level. The same article also studies the "hysteresis" phenomenon as an alternative explanation to the persistence of unemployment where, according to this view, the structural unemployment rate depends fully, or partially, on the current unemployment rate. The authors use the "scarring" effect of unemployment as a source of the "hysteresis" phenomenon citing that long periods of unemployment lead to less training opportunities, to depreciation of human capital. Additionally, past unemployment may be interpreted by employers as a signal of lower expected productivity, which eventually turns unemployment into a recurring cycle.

Because our work consists in trying to estimate the probability of unemployment during a period of increasing unemployment in Portugal, the analysis of persistence is of utter importance.

2.2.3. UNEMPLOYMENT AND HUMAN CAPITAL

EDUCATION

If we think of the theory of the firm, we recall that the simplest assumption is that firms try to maximize their profits. This assumption is then used to explain firms' decisions regarding their economic activity, which also contemplates the decision regarding the amount of labor being used. From the employer's perspective, hiring is an investment decision. They need to evaluate both the expected returns of labor, measured as expected productivity, as well as the risk of this expected productivity being less than the true productivity, while trying to maximize profits.

In his model of signalization in the labor market, Spence (1973) exemplifies how further investments in education signalize employers that an individual has higher ability. While individuals might grasp the real value of their ability, the employer cannot fully scope this information. The asymmetry in information, as well as the positive association that employers make between higher investments in education and higher ability, imply that education is a powerful signal of higher, or lower, productivity. This result is empirically confirmed in Arulampalam et al. (2000) where qualifications have statistical significance in determining unemployment incidence.

In fact, as it was mentioned earlier, unemployment can arise from the gap between expected marginal product and the wage rate set by the employer. This expected marginal product is smaller for those with low levels of human capital, all other characteristics equal, and, therefore, low levels of human capital are often related to greater probability of unemployment, while higher levels are related to the opposite.

The increase of the average level of education and the emphasis on the importance of education as a tool for empowering women, has contributed to reduce the gender

education gap and, in some countries, to even reverse it. Regardless, this similarity in educational levels was not followed by an equal choice of educational field, with women having greater presence in the areas of education, health and welfare, and humanities and arts while men still dominate engineering, manufacturing and construction, in OECD countries[3]. One important sub product of this asymmetry is gender segregation in occupation that has been suggested as the main cause in the existing gender pay gap.

These past years of some educational transformation suggest that the study of gender discrimination in the labor market should not continue to emphasize so much on the differences in education levels but instead relate it to other factors, for example, the impact of choosing male or female dominant fields of education.

EXPERIENCE

Employers, as most economic agents, make decisions based on limited information. Therefore, the greatest amount of available information allows for better predictions of expected productivity and for reducing the risk that arises when employing someone. Experience plays an important role in providing valuable information to the employer by signaling, through the candidates' previous work experience, if he or she would be a good fit to both the company and the job. In fact, according to human capital theory, the skills accumulated through experience raise the probability of being employed in the future.

As gender differences in educational levels become less and less relevant, with both men and women attaining similar levels of schooling, differences in actual experience have narrowed less (possibly from the differences in chosen fields of study, women's

---

[3] OECD Education database and OECD (2006), Education at a Glance: OECD Indicators, OECD, Paris.

seemingly preference for part time or temporary jobs and the fact that domestic work is still performed by women). As a result, their impact on unemployment incidence is especially important in developed countries that for have no gender gap in education.

For example, according to Blau & Kahn (1997), women's lower levels of human capital (especially lower levels of full-time experience) explain close to one third of the pay gap. Manning & Swaffield (2005) also found that, despite the approximately nonexistent pay gap, in the UK, when entering the labor market, there is a significant disparity past 10 years of entering it. They conclude that, although a large component of this difference is unexplained, human capital accounts for half of it, mainly thorough gender differences in on-the-job training and in accumulated experience.

2.2.4. UNEMPLOYMENT AND CHILDREN

Until now, we have focused our study on the impact that human-capital characteristics have on the probability of being unemployed. We have left an important part to take in consideration, which involves the aspects of raising kids, or having them, while searching for work. When studying the cross-country differences in the unemployment gender gap, Azmat et al. (2006) found that the gender gap in unemployment rates is larger for those who are married and those who have young children.

This might result from the fact that women can most likely be over-represented in part-time jobs (see Petrongolo (2004)). This job allocation is said to reflect women's preferences specially their need to combine work with child care, although some cases this part time employment is involuntary.

The association between fertility and female employment is usually found to be negative with Kögel (2004) finding that, although there has been a reduction since

1985, this association in OECD countries remains negative. This result can hinder future employment for women as they will have lower levels of experience than their male counter parts and may suffer more from the scarring effect of unemployment. The possibility that women at this age decide to have kids and voluntarily stay out of the labor force, but also from the general perception that a long maternity would lead to depreciation of the human capital stock and costs in temporary replacement, which eventually works as an obstacle for employment. These conclusions enhance the importance of estimating different impacts that family attributes have on the probability of a woman to be unemployed.

2.2.5. UNEMPLOYMENT AND OCCUPATIONAL SEGREGATION

Preferences over future occupations become more complex and narrower as a person grows up. One of the first criteria for shorting the list of possible future occupations is eliminating the ones that are perceived to be socially inappropriate for the person's sex (Gottfredson, 1981). This, combined with an image of who they would like to be, are some of the reasons that explain occupational segregation, as well as educational segregation.

Therefore, it is plausible to assume that the fact that society perceives some roles as strictly female or strictly male, have an extremely important impact creating gender differences both on the pursued fields of education, as well as the chosen economic occupation (see Cejka & Eagly (1999) or Wright et al. (2015)). This may, in itself, create a "snowball" effect if it strengthens the gender-occupation stereotype, which would, most likely, result in more statistical discrimination.

In fact, as we have seen with both theoretical models of discrimination, a sub product of them is occupational segregation. Therefore, if an occupation is male-dominant, it

is likely that some type of discrimination, either taste-based or statistical, might be contributing to the high employment rate of men. An analogous conclusion comes from female-dominant occupations. Consequently, differences in employment rates (and, therefore, unemployment rates) regarding gender must be correlated with prevailing gender segregation in economic occupations. In fact, some studies conclude that gender stereotypes might be preventing women from being hired and/or promoted in particular occupations, hindered by gender roles (e.g., Eagly & Karau, 2002).

Another matter that makes occupational gender segregation an important factor when studying gender discrimination is the fact that it is thought to be one of the main explanations for the gender gap in earnings of the current time, where women are as much educated as men (Gauchat et al. 2012). It is also perceived as economical inefficient as the lack of gender representation might keep talented individuals from occupying roles that they would be a great fit. This will result in lower overall productivity and economic growth that otherwise could come from unconstrained choices.

3. THE MODEL

The econometric specification is based on a binary dependent variable, y_{it}, which takes on the value one if the individual i is unemployed at time t and zero otherwise.

Consider the following dynamic model assuming that y_{i0}, the initial condition of y_{it}, is the value for y for each individual i in 2010

$$y_{it} = \mathbf{1}\big[\, \mathbf{z_{it}}\boldsymbol{\gamma^1} + \mathbf{z_{it}}\boldsymbol{\gamma^2}\mathbf{d^f}_i + \rho^1 y_{it-1} + \rho^2 y_{it-1}d_i^f + \varphi_i + v_{it} \geq 0 \big] \quad (1)$$

$$i = 1, \dots, n; t = 1, \dots, T$$

Where $\mathbf{1[.]}$ is the usual indicator function, $\mathbf{z_{it}}$ is a $1 \times K$ vector of contemporaneous

explanatory variables (such as education, experience, number of kids and age), d_i^f is a dummy variable that equals one if the individual i is a woman, $\boldsymbol{\gamma^1}$ is a $K \times 1$ vector of parameters associated with $\mathbf{z_{it}}$, $\boldsymbol{\gamma^2}$ is a $K \times 1$ vector of parameters associated with the interaction between d_i^f and $\mathbf{z_{it}}$, ρ^1 is the parameter that reflects the persistence of unemployment for men, ρ^2 is the parameter that reflects differences between genders for the persistence of unemployment, v_{it} is the idiosyncratic error term and φ_i is the unobserved heterogeneity term, constant in time.

In order to obtain $D(y_{it}|\mathbf{z_{it}}, y_{it-1}, \varphi_i)$, the following two assumptions are made

Assumption 1:

$$D(y_{it}|\mathbf{z_{it}}, y_{it-1}, \varphi_i) = D(y_{it}|y_{it-1}, \dots, y_{i0}, \mathbf{z_{it}}, \varphi_i) \qquad (2)$$

Assumption 2: $f_t(y_t|\mathbf{z_t}, y_{t-1}, \varphi; \boldsymbol{\theta})$ is a correctly specified probability function for the conditional distribution on the left side of equation (2), where $\boldsymbol{\theta}$ is a vector of parameters.

These two assumptions imply that the dynamics are correctly specified and that $\mathbf{z_{it}} = \{\mathbf{z_{i1}}, \dots, \mathbf{z_{iT}}\}$ is strictly exogenous, conditional on φ_i. Moreover, because it doesn't restrict the distribution of φ_i, it allows for dependence between the unobserved effects and $\mathbf{z_{it}}$.

Furthermore, these assumptions imply that the probability function of $(y_{i1}, \dots, y_{iT})$ given $(y_{i0} = y_0, \mathbf{z_{it}} = \mathbf{z_t}, \varphi_i = \varphi)$ is $\prod_{t=1}^{T} f_t(y_t|\mathbf{z_t}, y_{t-1}, \varphi; \theta_0)$, with $\boldsymbol{\theta_0}$ equal to the true value of $\boldsymbol{\theta}$. Therefore, this probability function depends on an unobservable term, φ, and because of that, is not useful for inference. One solution to overcome this problem would be to consider φ as a fixed effect, resulting in the estimation of n parameters (with nT observations) which would lead to the incidental parameter

problem[4] when T is fixed, resulting in highly biased estimates of γ.

Wooldridge (2005) solves this problem by using the density of $(y_{i1}, \dots, y_{iT})$ conditional on $(y_{i0}, \boldsymbol{z_i})$. Because we already know the density of $(y_{i1}, \dots, y_{iT})$ conditional on $(y_{i0} = y_0, \boldsymbol{z_{it}} = \boldsymbol{z_t}, \varphi_i = \varphi)$, we only need to specify the density of φ_i conditional on $(y_{i0}, \boldsymbol{z_{it}})$ and assume that it is correctly specified, which leads to Assumption 3.

Assumption 3: $h(\varphi|y_0, \boldsymbol{z}; \boldsymbol{\alpha})$ is a correctly specified model for the density of $D(\varphi_i|y_{i0}, \boldsymbol{z_i})$ with respect to a σ-finite measure $\eta(d\varphi)$, where $\boldsymbol{\alpha}$ is some vector of parameters with the true value equal to $\boldsymbol{\alpha_0}$.

Under assumptions 1,2 and 3, the density of $(y_{i1}, \dots, y_{iT})$ given $(y_{i0} = y_0, \boldsymbol{z_i} = \boldsymbol{z}, \varphi_i = \varphi)$ is

$$\int_{\mathbb{R}^J} \prod_{t=1}^{T} f_t(y_t| \boldsymbol{z_t}, y_{t-1}, \varphi; \boldsymbol{\theta_0}) \, h(\varphi|y_0, \boldsymbol{z}; \boldsymbol{\alpha_0}) \, \eta(d\varphi)$$

Leading to the following log-likelihood function for each observation

$$l_i(\boldsymbol{\theta}, \boldsymbol{\alpha}) = \log\left[\int_{\mathbb{R}^J} \prod_{t=1}^{T} f_t(y_t| \boldsymbol{z_t}, y_{t-1}, \varphi; \boldsymbol{\theta_0}) \, h(\varphi|y_0, \boldsymbol{z}; \boldsymbol{\alpha_0}) \, \eta(d\varphi) \right] \quad (3)$$

To estimate $\boldsymbol{\theta_0}$ and $\boldsymbol{\alpha_0}$, Wooldridge (2005) uses the conditional Maximum Likelihood Estimator by summing the log-likelihood in equation (3) across $i = 1, \dots, N$ and maximizing with respect to $\boldsymbol{\theta}$ and α. This yields $\sqrt{N}$-consistent and asymptotically normal estimators, under standard regularity conditions.

Let us now consider that the unobserved heterogeneity term (φ_i) has a common part for both genders (φ_{0i}) and a fixed effect for women (θ). Therefore, we can write φ_i as $\varphi_i = \varphi_{0i} + \theta d_i^f$.

Consequently, we can rewrite equation (1) as the following

---

[4] More on this in Wooldridge (2010)

$$y_{it} = \mathbf{1}\left[\mathbf{z_{it}\gamma^1} + \mathbf{z_{it}\gamma^2 d_i^f} + \rho^1 y_{it-1} + \rho^2 y_{it-1} d_i^f + \varphi_{0i} + \theta d_i^f + v_{it} \geq 0 \right] (4)$$

From the literature survey, we are aware that one of the reasons behind the gap in unemployment rates is taste-based discrimination and statistical discrimination. The latter happens when the employer discriminates women, not because he or she has any disutility towards them, but because his/her general perception is that women have less ability than their male counterparts. Therefore, it is not restrictive to assume that women would have higher probabilities of unemployment in occupations where their perceived ability is lower, e.g. on activities that are usually thought to be reserved for men and, consequently, this fixed effect for women θ, could also reflect that interesting result. In fact, because we aim to control for true individual ability, a difference between genders in the probability of unemployment should be associated with some type of discrimination which, has we have seen previously, usually leads to occupational segregation, that is said to be one of the causes behind the still existing wage gap. Then, this fixed effect aims to capture all types of discrimination that contribute to a possible gender gap in the probability of unemployment.

Finally, we are left to choose the density for the random unobserved heterogeneity term. We will adopt Mundlak's version of Chamberlain's assumption as in Wooldridge (2010), with the addition of conditioning on y_{i0} as well.

To specify the density of φ_{0i} conditional on $(y_{i0}, \mathbf{z_i})$, consider the following specification

$$\varphi_{0i} = \alpha_0 + \alpha_1 y_{i0} + \overline{\mathbf{z_i}}\mathbf{\alpha_2} + a_i \text{ where } a_i|y_{i0}, \mathbf{z_i} \sim N(0, \sigma_a^2) \ (6)$$

and $\overline{\mathbf{z_i}}$ is a vector that contains the average value, for each individual, for all covariates that vary over time.

Then, the density for φ_{0i} conditional on $(y_{i0}, \mathbf{z_i})$ can be specified as the following

$$\varphi_{0i}|y_{i0}, \mathbf{z_i} \sim N(\alpha_0 + \alpha_1 y_{i0} + \overline{\mathbf{z_i}}\boldsymbol{\alpha_2}, \sigma_a^2) \quad (7)$$

In conclusion, y_{it} given $(y_{it-1}, \dots, y_{i0}, \mathbf{z_i}, a_i, d_i^f)$ follows a dynamic probit model with unobserved effects with the following specification

$$P(y_{it} = 1|y_{it-1}, \dots, y_{i0}, \mathbf{z_{it}}, \varphi_{it}, d_i^f) = \Phi\Big[\mathbf{z_{it}}\boldsymbol{\gamma^1} + \mathbf{z_{it}}\boldsymbol{\gamma^2}\boldsymbol{d_i^f} + \rho^1 y_{it-1} + \rho^2 y_{it-1}d_i^f + $$

$$\theta d_i^f + \alpha_1 y_{i0} + \overline{\mathbf{z_i}}\boldsymbol{\alpha_2} + a_i\Big] \quad (8)$$

3.1. ESTIMATING AVERAGE PARTIAL EFFECTS

Our aim is to estimate the partial effect of each explanatory variable in equation (8). Because it is a nonlinear function, it will depend on the observed values of the variables, therefore, we will need to estimate the average partial effects (APEs). However, the APEs depend on unobserved variables (φ_i), therefore Wooldridge (2005, 2010) suggests to average APEs across the distribution of φ_i. The expectation of model (8) gives,

$$\mu(\mathbf{z}, y_{-1}) = E\Big[\Phi[\mathbf{z}\,\boldsymbol{\gamma^1} + \mathbf{z}\,\boldsymbol{\gamma^2}d_i^f + \rho^1 y_{-1} + \rho^2 y_{-1}d_i^f + \theta d_i^f + \alpha_1 y_{i0} + \overline{\mathbf{z_i}}\boldsymbol{\alpha_2}$$

$$+ a_i]\Big] \quad (9)$$

where the expectation is with respect to the distribution of a_i, $\mathbf{z}$ and y_{-1} are possible values of $\mathbf{z_t}$ and y_{t-1}, respectively.

It is possible to obtain a consistent estimator of equation (9) using iterated expectations leading to,

$$E\{E[\Phi(\mathbf{z}\boldsymbol{\gamma^1} + \mathbf{z}\,\boldsymbol{\gamma^2}d_i^f + \rho^1 y_{-1} + \rho^2 y_{-1}d_i^f + \theta d_i^f + \alpha_1 y_{i0} + \overline{\mathbf{z_i}}\boldsymbol{\alpha_2} + a_i|y_{i0}, \mathbf{z_i}]\} \quad (10)$$

The conditional expectation inside equation (10) given (7) is

$$\Phi(\mathbf{z}\boldsymbol{\gamma_a^1} + \mathbf{z}\boldsymbol{\gamma_a^2}d_i^f + \rho_a^1 y_{-1} + \rho_a^2 y_{-1}d_i^f + \theta_a d_i^f + \alpha_{a1}y_{i0} + \overline{\mathbf{z_i}}\boldsymbol{\alpha_{a2}}) \quad (11)$$

where the 'a' subscript denotes the original parameter multiplied by $(1 + \sigma^2{}_a)^{-1/2}$.

A consistent estimator of the expected value of (11) with respect to the distribution of $(y_{i0}, \mathbf{z_i})$ at time t is

$$N^{-1}\sum_{i=1}^{N} \Phi\left(\mathbf{z_{it}}\widehat{\boldsymbol{\gamma}_a^1} + \mathbf{z_{it}}\widehat{\boldsymbol{\gamma}_a^2}d_i^f + \hat{\rho}_a^1 y_{it-1} + \hat{\rho}_a^2 y_{it-1}d_i^f + \hat{\theta}_a d_i^f + \hat{\alpha}_{a1}y_{i0} + \overline{\mathbf{z}_i}\widehat{\boldsymbol{\alpha}_{a2}}\right) \quad (12)$$

where the 'a' subscript now denotes multiplication by $(1 + \hat{\sigma}^2{}_a)^{-1/2}$ and the estimates of the parameters may be obtained with a random effects probit or a pooled probit, where the former estimates the "unscaled" parameters and $\sigma^2{}_a$, while the pooled probit estimate directly the scaled parameters.

Therefore, the average partial effect of a generic dummy variable, say the k-th variable of the vector $\mathbf{z_{it}}$, will be, for a man and for a woman, respectively, equal to the following

$$\widehat{APE}_{kt} = N^{-1}\left\{\sum_{i=1}^{N}\left[\Phi\left(\mathbf{z_{it}}^{k-1}\widehat{\boldsymbol{\gamma}_a}^{1,k-1} + \hat{\gamma}_{ka}^{1,k-1} + \hat{\rho}_a^1 y_{it-1} + \hat{\alpha}_{a0} + \hat{\alpha}_{a1}y_{i0} + \overline{\mathbf{z}_i}\widehat{\boldsymbol{\alpha}_{a2}}\right) - \right.\right.$$

$$\Phi\left(\mathbf{z_{it}}^{k-1}\widehat{\boldsymbol{\gamma}_a}^{1,k-1} + \hat{\rho}_a^1 y_{it-1} + \hat{\alpha}_{a0} + \hat{\alpha}_{a1}y_{i0} + \overline{\mathbf{z}_i}\widehat{\boldsymbol{\alpha}_{a2}}\right)\left.\left.\right]\right\} \text{ if } i \text{ is a man}$$

$$\widehat{APE}_{kt} = N^{-1}\left\{\sum_{i=1}^{N}\left[\Phi\left(\mathbf{z_{it}}^{k-1}(\widehat{\boldsymbol{\gamma}_a}^{1,k-1} + \widehat{\boldsymbol{\gamma}_a}^{2,k-1}) + \hat{\gamma}_{ka}^{1,k-1} + \hat{\gamma}_{ka}^{2,k-1} + (\hat{\rho}_a^1 + \hat{\rho}_a^2)y_{it-1} + \right.\right.\right.$$

$$\hat{\theta}_a + \hat{\alpha}_{a1}y_{i0} + \overline{\mathbf{z}_i}\widehat{\boldsymbol{\alpha}_{a2}}\Big) -$$

$$\Phi\left(\begin{matrix}\mathbf{z}^{k-1}(\widehat{\boldsymbol{\gamma}_a}^{1,k-1} + \widehat{\boldsymbol{\gamma}_a}^{2,k-1}) + (\hat{\rho}_a^1 + \hat{\rho}_a^2)y_{it-1} + \hat{\theta}_a + \hat{\alpha}_{a1}y_{i0} + \\ \overline{\mathbf{z}_i}\widehat{\boldsymbol{\alpha}_{a2}}\end{matrix}\right)\left.\left.\right]\right\} \text{ if } i \text{ is a woman}$$

where $\mathbf{z^{k-1}}$ is a $1 \times (K-1)$ vector containing the first $K-1$ variables of $\mathbf{z}$ and $\widehat{\boldsymbol{\gamma}_a}^{g,k-1}$ is the correspondent vector of parameters for $g = 1,2$.

While for some variable z^j (continuous) it will simply be,

$$\widehat{APE}_{jt} = \widehat{\boldsymbol{\gamma}_{aj}^1}\frac{1}{N}\sum_{i=1}^{N}\phi(\mathbf{z_{it}}\widehat{\boldsymbol{\gamma}_a^1} + \hat{\rho}_a^1 y_{it-1} + \hat{\alpha}_{a0} + \hat{\alpha}_{a1}y_{i0} + \overline{\mathbf{z}_i}\widehat{\boldsymbol{\alpha}_{a2}}) \text{ if } i \text{ is a man}$$

$$\widehat{APE}_{jt} = (\widehat{\boldsymbol{\gamma}_{aj}^1} + \widehat{\boldsymbol{\gamma}_{aj}^2})\frac{1}{N}\sum_{i=1}^{N}\phi(\mathbf{z_{it}}\widehat{\boldsymbol{\gamma}_a^1} + \mathbf{z_{it}}\widehat{\boldsymbol{\gamma}_a^2} + \hat{\rho}_a^1 y_{it-1} + \hat{\rho}_a^2 y_{it-1} + \hat{\theta}_a + \hat{\alpha}_{a1}y_{i0} +$$

$$\overline{\mathbf{z}_i}\widehat{\boldsymbol{\alpha}_{a2}}) \text{ if } i \text{ is a woman}$$

Then, APEs are $\sqrt{N}$ consistent and asymptotically normal distributed. Their standard error can be obtained using the delta method or bootstrap.

3.2. ESTIMATING THE EFFECTS OF PREVIOUS UNEMPLOYMENT

To study the effects of previous unemployment on current unemployment, we will estimate the probability of unemployment at time t of the individual i if he or she was unemployed at time $t - 1$. This probability will be obtained for both genders and, subsequently, analyzed for the existence of significant dissimilarities between genders. This results in estimating the following probabilities for men and women, respectively

$$P(y_{it} = 1 | y_{it-1} = 1, d^f{}_i = 0, \dots, y_{i0}, \boldsymbol{z_{it}}) = \Phi[\,\boldsymbol{z_{it}\gamma^1} + \rho^1 + \alpha_0 + \alpha_1 y_{i0} + \boldsymbol{\overline{z_i}\alpha_2} + a_i\,]$$

$$P(y_{it} = 1 | y_{it-1} = 1, d^f{}_i = 1, \dots, y_{i0}, \boldsymbol{z_{it}}) = \Phi(\,\boldsymbol{z_{it}\gamma^1} + \boldsymbol{z_{it}\gamma^2} + \rho^1 + \rho^2 + \theta + \\ \alpha_1 y_{i0} + \boldsymbol{\overline{z_i}\alpha_2} + a_i)$$

Another important estimate that enables the study of unemployment persistence, is the APE of the lagged dependent variable which consists in obtaining the following estimates for men and women, respectively

$$\widehat{APE}_{UNt} = \frac{1}{N}\sum_{i=1}^{N}[P(y_{it} = 1 | \boldsymbol{z_{it}}, y_{it-1} = 1, d^f{}_i = 0) - P(y_{it} = 1 | \boldsymbol{z_{it}}, y_{it-1} = 0, d^f{}_i = 0)]$$

$$\widehat{APE}_{UNt} = \frac{1}{N}\sum_{i=1}^{N}[P(y_{it} = 1 | \boldsymbol{z_{it}}, y_{it-1} = 1, d^f{}_i = 1) - P(y_{it} = 1 | \boldsymbol{z_{it}}, y_{it-1} = 0, d^f{}_i = 1)]$$

4. THE DATA AND VARIABLES

The data was provided by Statistics Portugal (*Instituto Nacional de Estatística* − INE) and is from four waves of the Survey on Income and Living Conditions (ICOR), a Portuguese representative survey of households that allows for the analysis of living conditions, income and employment. This survey is included in the European program - Statistics on Income and Living Conditions - EU-SILC and was conducted annually from 2010 to 2013 with a quarter of the sample being randomly replaced each year. The

final sample has been chosen so that all individuals remain in every wave. This is an important constraint as the econometric analysis we are going to perform requires a lag of unemployment in order to study its persistence. In addition, to ease the estimation of the initial conditions, it was chosen a common date of entry 2010 in order to assume that y_{i2010} is the initial condition of y for each individual i. Furthermore, we removed individuals who had missing relevant information in any year and, since our aim is to study the active population (employed or unemployed status), the individuals who were not in the labor market for some year, at the time of the interview, were also dropped. The final sample resulted in 774 individuals observed in all 4 waves, from 2010 to 2013, with 384 women and 390 men.

The measure of unemployment used throughout this work is based on the ILO definition of unemployment, i.e., a person is considered unemployed if he or she does not have a job, has looked for work in the past four weeks and is available for work.

The vector of covariates which will be used in the estimation of the model and their expected signal are listed in Table A1 in Appendix. The descriptive statistics of those covariates can be seen in the Appendix, in Table A2 and Table A3 for all of the sample and for women, respectively. The information on the real number of kids of each individual was not available. This variable was computed, for individual i, as the sum of individuals in the original sample who had reported i to be his or her parent. Therefore, it is a proxy for the number of kids of the individual i. The status of employment of the spouse was included to condition for family conditions. In the beginning it was used the variable $married_{it}$ which was 1 if the individual i was married at time t, but it turned out to be extremely insignificant from a statistical point of view. To somehow replace it, we used the status of employment of the spouse that was built in a similar way than

number of kids. As a limitation, it was not possible to identify the unemployment status of his or her spouse for 1,4% of the individuals who were married in the sample.

The raw unconditional probability of being unemployed for each wave is presented in Table 1. It also displays the distribution of employment status, over the four-year period.

This table shows evidence of higher female participation on part time employment, for all years. As mentioned in the literature review, some women tend to search and accept jobs that enables for temporary leave or less working hours, so that they can take care of children, while men don't generally search for the same conditions. It is also possible to notice that 2012 was a turning point regarding unemployment: until there, women were the dominant group, where in 2012 and 2013 this place was occupied by men.

Table 1 – Distribution of employment status

	Wave 1 (2010)	Wave 2 (2011)	Wave 3 (2012)	Wave 4 (2013)
Total				
% Unemployed	14%	14%	16%	17%
% Employed Full Time	81%	81%	80%	79%
% Employed Partial Time	5%	5%	4%	4%
Men				
% Unemployed	12%	13%	17%	18%
% Employed Full Time	84%	84%	81%	80%
% Employed Partial Time	4%	3%	2%	2%
Women				
% Unemployed	15%	14%	14%	16%
% Employed Full Time	79%	79%	80%	78%
% Employed Partial Time	6%	7%	6%	6%

Source: INE - ICOR, author calculations

This could be traced to the sectoral composition of job losses in Portugal for this time as it affected the craft, industry and construction qualified workers[5] the most, which, is a male dominant occupation.

---

[5] See Table A4 from the Appendix.

From Table A5 in Appendix we can see that, the probability of being unemployed in 2011, conditioned on being unemployed in 2010, was of approximately 70% for men and for women. In 2013, this probability increases to values of 81% and 78% for men and women, respectively. This suggests an increase of persistence in unemployment, which reinforces the need to study it, especially on the account of how high the unemployment rate was (17% of the active population was unemployed in 2013).

Table 2 shows the distribution of the levels of education. This is meant to unveil the expected productivity for both genders, based only on education (excluding other relevant aspects, such as experience and training). From this we can see that women in the sample have a dominant presence on higher levels of education (high school and college), indicating higher expected productivity for women based only on education.

Table 2 – Distribution of levels of education

% For women relatively to men				
	Wave 1 (2010)	*Wave* 2 (2011)	*Wave* 3 (2012)	*Wave* 4 (2013)
%Sixth grade	45%	44%	44%	45%
% Ninth grade	43%	43%	44%	41%
% High school	54%	54%	54%	54%
%College	63%	63%	63%	63%

Source: INE - ICOR, author calculations

In figures 1.a and 1.b, we consider three levels of experience, that differ with each age range. The first one is meant to represent higher attatchment to the labor market (in green), the second one that is meant to represent average attatchment (in the middle) and the last one represents low attatchment (red). For example, an individual who is aged between 20 to 30 years is expected to have, if his/her attatchment to the labor market is high, more than 5 years of experience or, if his/her attatchment is low, lower than 2 years of experience.

From this, we can observe that, except for the early stages of ones career, men in the

sample appear to have higher attatchment to the labor market. This conclusion comes from higher percentages for men, for each age range, in the levels of experience linked to higher attatchment, while women seem to focus on the middle levels of experience. Notwithstanding, when comparing 2013 to 2010, it appears as there has been a convergence, especially for the higher ranges of age. This is consistent with the idea that women are increasing their labor attatchment which in result will eventually reduce a gap in unemployment rates between genders.

Therefore, the data does replicate some of the results we expected from the literature survey, such as relatively lower labor attachment for women, with a convergence happening in this period of economic recession; higher presence in the lower levels of education for men; higher participation in part time employment for women relatively to men; higher increase of men's unemployment rate during a recession and persistence of high unemployment rates.

Market labor attatchment[6]
Figure 1.a – Relating age with experience[7] (2010)
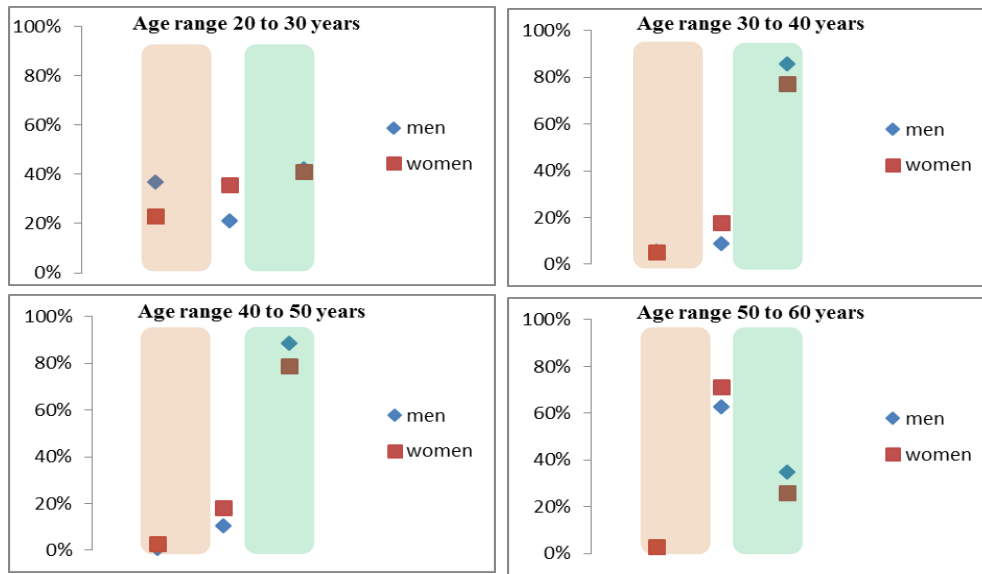
Source: INE - ICOR, author calculations

Figure 1.b– Relating age with experience (2013)
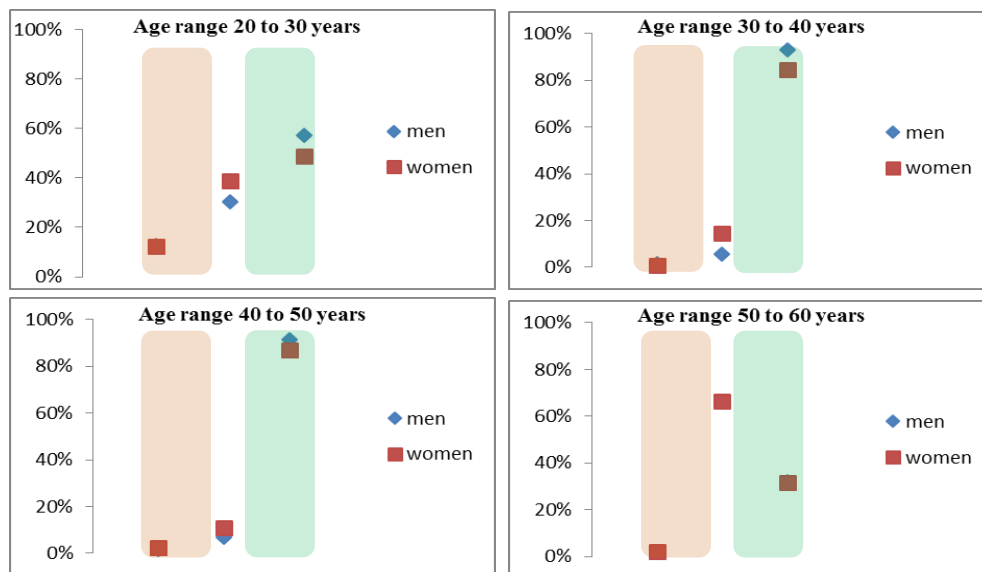
Source: INE - ICOR, author calculations

---

[6] For the age range between 20 to 30 years old it was assumed that having two or less years of experience implied low attachment, while having more than five years implied high attachment; for the age range between 30 to 40 years old low attachment is having less than five years of experience while high attachment is having more than ten years of experience; for the age range between 40 to 50 years, low attachment is having less than ten years of experience while high attachment is having more than twenty years of experience; for the age range between 50 to 60 years old, low attachment is having less than twenty years of experience and high attachment is having more than forty years of experience.

[7] Information regarding experience was built with information regarding how many years they have been in paid work.

5. THE RESULTS

In our model, we include the usual set of control variables that reflect individual characteristics and family variables that affect expected productivity, and also a fixed effect of time that aims to control for macroeconomic effects.

This section presents the main results obtained using the methodology developed in section 3 applied to the data described in section 4. It will be divided into three sub sections, the first sub section starts with the unrestricted random effects probit model, the second develops on the restricted model and presents the estimates for the average partial effects. Finally, the final section studies the persistence in unemployment.

5.1. RANDOM EFFECTS PROBIT MODEL

Estimates of panel data models for the probability of unemployment using both dynamic random effects probit (REP), modeled with the Mundlak device as in (8), and the pooled probit (PP), estimated for comparison purposes only, given the fact that its estimates are inconsistent, are given in Table A6 in Appendix. For these models the variables *married* and the square of both *experience* and *age* were not included because they were highly insignificant. The variable *married* was replaced by the variable *unemp spouse* that is meant to reflect the employment status of the spouse of the individual which in result can control for family conditions.

After some likelihood ratio tests, the final model that will be used is presented in Table A7 in Appendix. The variable *numberkids* only appears to be statistical insignificant in these unrestricted models because the mean of the variable was being controlled for, simultaneously. All of the variables meant to extract the relation between covariates and unobserved heterogeneity were withdrawn from the model, i.e., the final model does not include the Mundlak device because it was not statistically significant at the usual level

of 10%. The only regressors which interact with the variable *female* that seem to have statistical significant (at a 10% level) between genders were *age* and *numberkids*. Notwithstanding, the differences between the models in Table A6 and Table A7, generally speaking, are not very significant, although the latter model beneficiates from higher efficiency that derives from being more parsimonious. All variables affect positively or negatively the probability of unemployment as expected from the literature survey, with the exception being the fact that *age* and unemployment condition of the last period seemingly affect women's probability of unemployment in a less negative way than they affect men's.

5.2. ESTIMATES OF AVERAGE PARTIAL EFFECTS

The contribution in the probability of unemployment of each exogenous variable considered is obtained through the estimation of the average partial effect (APE) of each variable averaged across the distribution of unobserved heterogeneity. Table 3 gives the average partial effects for each model.

Observing the exogenous variables, the fixed effect of being a woman has the highest impact on the probability of unemployment with an APE of 0.112 probability points (pr.p.) for the REP model, although not statistically significant at the usual 10% (p-value equal to 13,7%). Notwithstanding, for the PP model it still has a high value of 0.091 pr.p. and it is statistically significant. *Age* and *experience* seem to affect the probability the least, with an average partial effect of only 0.004 pr.p. and -0.004 pr.p., respectively. For women, *age* has an even lower APE (0.002 pr.p.), meaning that when at the same age, everything else constant, men have a higher probability of being unemployed than women. This is an interesting result as the general perception is that, with age, it is harder to find a job for women than for men.

Having an educational level assumed as high[8], reduces the overall probability of unemployment in 0.030 pr.p. while having a spouse that is unemployed raises it in 0.058 pr.p.. The first result is consistent with the literature, as it was expected that individuals with higher education would appear to have higher levels of productivity to employers. For the second result, there was no relevant literature that could explain this positive association between the unemployment status of an individual and of his or her spouse. Therefore, this comes as an especially interesting result that could be further developed in the future, perhaps as an association between higher probability of unemployment and employment unstableness of the household. This could also represent how most economic crisis affect the population differently, resulting in higher levels of inequality. The number of kids has different effects on the probability of unemployment when we consider both genders. When comparing a woman with the same number of kids as a man, while controlling for all other covariates, she will, on average, have a higher probability of being unemployed of 0.028 pr.p..

The positive and significant coefficient of the lag of the dependent variable suggests that there is persistence in unemployment. Therefore, our results provide favorable evidence that past unemployment raises the probability of current unemployment. For the considered time period, women seem to suffer less from this situation. Although surpassing the usual 10% significance level, with a p-value of 10,7%, the interaction between *female* and the lag of the dependent variable reaches an APE of -0.019 pr.p. in the REP model. Notwithstanding, if we consider the PP model, this APE turns out to be statistically significant, reducing the effect of unemployment persistence for women in 0.032 pr.p..

---

[8] An individual is said that he or she has high educational levels if he or she has a high school degree or above.

Table 3 – Average partial effects

	Random Effects Probit (1)	Pooled Probit (2)
Unemp at t-1	0.055*** (0.021)	0.127*** (0.014)
Unemp at t-1×Female	-0.019 (0.012)	-0.032* (0.018)
Age	0.004*** (0.001)	0.006*** (0.001)
Age×Female	-0.002*** (0.001)	-0.003** (0.001)
Unemp Spouse	0.058*** (0.017)	0.067*** (0.153)
Female	0.112 (0.071)	0.091* (0.047)
Number Kids	-0.021** (0.010)	-0.026** (0.012)
NumberKids×Female	0.028** (0.014)	0.034* (0.017)
Experience	-0.004*** (0.001)	-0.006*** (0.001)
Higher Educ	-0.030*** (0.010)	-0.044*** (0.013)

Notes:
1. Standard errors are in brackets. For both models the standard error was computed with the Delta Method.
2. Significance levels: *10%, **5%, ***1%.
3. The APE of the variable Unemp at t-1×Female has a p-value of 10,7% for the RE model.
4. The APE of the variable Female has a p-value of 11,2% for the RE model.

This effect might be related to the fact that, in this period of crisis and high unemployment, women were more flexible to the new labor market conditions, thus, accepting more jobs than men. The effects that long-term unemployment can have on individuals, such as loss in human capital accumulation, higher financial instability and the discouragement in actively looking for work resulting from repeated failures in the search and application process, are part of the reasons that makes studying past unemployment as a determinant of current unemployment so important. According to

the literature related to this topic, unemployment persistence can reflect increases in the natural rate of unemployment, i.e., can reflect increases in the long run equilibrium of the unemployment rate. Consequently, if the unemployment rate has been persistently high in result of an increase in the long-term equilibrium unemployment rate, labor market policies should focus on structural labor reforms rather than only focusing on increasing short term employment. Because this was a period marked by high unemployment rates and, simultaneously, a reduction in public spending, the lack of policies aimed to fix this crisis-aggravated problem could eventually translate into a slow adjustment of the unemployment rate that might actually never achieve the same level as in the period before the crisis. Therefore, labor market reforms, especially in the form of creating stable employment and in the form of increasing human capital of long term unemployed individuals, need to take place in order to prevent the results that come from permanently higher unemployment rates.

The estimated average partial effect for persistence is also higher for men than women, with men who were unemployed in the previous period having a probability of unemployment higher in 0.055 pr.p. than men who were employed in the previous period and in 0.019 pr.p. higher than women who were unemployed in the previous period. Despite having a lower effect than men, women who were unemployed in the previous period also have a higher probability of being unemployed in the present period in 0.036 pr.p. than women who were employed at the time of the previous period. This reflects the "scarring" effect of unemployment, which leaves the unemployed with fewer opportunities to change their condition towards employment.

The fact that men appear to suffer more from the negative implications of previous unemployment could be explained by various factors such as the fact that the recession

affected male dominant industries the most and that men might be more reluctant to accept any job, especially if it's not in the field of their previous occupation.

6. CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

The present work provides some answers on important questions regarding gender discrimination and unemployment persistence in the Portuguese labor market. We estimated a binary panel data model for the probability of unemployment that simultaneously controls for unobserved individual heterogeneity. This unobserved heterogeneity, in this context, could represent important individual characteristics that are not observable, such as individual ability and unobserved discrimination, both taste-based and statistical discrimination.

Our results suggest that there is evidence of higher probabilities of unemployment for women, relatively to men, in spite of women having stronger presence in higher levels of education. Notwithstanding, it appears that the economic crisis helped closing the gender gap in the probability of unemployment, with the unconditional unemployment rate of men surpassing women's, replicating some empirical evidence which found that, in periods of economic recession, men's unemployment rate rises faster than women's. This could reflect the effect of higher female labor force attachment[9] fueled by higher financial necessities and labor instability, a result from the economic crisis affecting Portugal during this time. Higher education and experience appear to have negative effects on the probability of unemployment, contributing to its reduction. Thus, the importance of human capital in reducing the probability of unemployment is reinforced. By controlling for ability, which is assumed to be included in the unobserved heterogeneity, these human capital effects become independent of differences in ability,

---

[9] Albanesi & Şahin (2013) found that the rise in female labor force attachment and the decline in male attachment could account for the most part of the closing of the gender unemployment gap.

which strengthens the idea that the attainment of higher levels of education and of higher labor attachment are reliable signals of high marginal productivity to employers. Both age and the number of kids seem to influence the probability of unemployment differently between genders, with the increase of the number of kids raising the probability of unemployment for women, while reducing it for men. This is consistent with the theory that taking care of kids is still a job predominately done by women. In particular, having kids might affect women's presence in the labor market in a twofold way: by signaling employers that women might need to leave work more often or by reducing women's desire in full time work experience. Some policies regarding this particular result should take place, as childbearing is especially important for Portugal, a country that has been suffering by the complications that arise from population ageing, in particular hindering the foundation of social security. This could be done e.g. by reducing the non-wage cost of labor, in particular by offering day care benefits to newly parents and by forcing both genders to take an equal amount of days in parental leave.

When we tried to control for discrimination using a fixed effect for women, we obtained strong statistical evidence that it increases women's probability of unemployment. This indicates that labor reforms should focus on trying to reduce both taste-based and statistical discrimination, e.g. focusing on attaining gender parity in occupations, as it may change society's perception on gender roles by not socially restricting particular jobs to one specific gender. Gender parity in occupations could simultaneously spread information of the productivity of the other gender, which could eventually lead to a reduction in statistical discrimination by closing a specific informational gap. It can also spread the risk that each gender carries when focusing on one specific set of occupations, such as the risk of an economic crisis strongly affecting a "one-gender-

dominated" industry. In the long run, if a parity is attained for most occupations then, eventually, it could translate into a change in both occupational and educational gender segregation, as well as into eliminating all statistical discrimination that come from distorted perceptions of expected productivity based on gender.

We were also able to find strong state dependence effects with respect to previous unemployment incidence, during this period of high unemployment in Portugal. This finding is consistent with the theory that previous unemployment experience has a sizeable impact on future employment. It can also be a result from a higher structural unemployment rate, which suggests that, despite the focus on fiscal consolidation, Portugal is concurrently in need of deep labor reforms that aim to provide better conditions for individuals that want to work. Therefore, if employment instability has such high implications on future employment, labor policies should focus on offering higher assistance in job-search and training programs to individuals who have been unemployed for some time. Hence, it would be possible to contradict the trend of human capital depreciation and could eventually lead to higher employability.

Further research on the topic of the present work could focus on capturing the impact that being in a one-gender-dominant occupation has on the probability of unemployment, e.g. focusing on estimating if women who in the previous period were employed in the science, technology, engineering, and mathematics (STEM) related occupations have higher or lower probabilities of being unemployed in the next period. This has an ambiguous expected association. The fact that entering those occupations is harder for women might either reduce the probability of staying employed in the next period, or, because they were made to ensure higher expected productivity to enter said occupations, their state of employment might be more stable than men in the same

occupation and women in other occupations. Some other relevant aspects could be studied in order to reveal the extent that discrimination can have on the labor market, such as educational segregation, e.g. the impact that choosing a STEM field of education has on a woman's probability of unemployment; the impact that specific labor policies have on this gender differential, such as raising unemployment benefits or decreasing the minimum wage; and the impact that specific policies aimed to contribute to higher gender equality and to provide better conditions for women in the labor market, such as the impact of possible equal mandatory parental leave and the provision of better child care benefits for parents.

REFERENCES

Ahmad, N. (2014). State dependence in unemployment. *International Journal of Economics and Financial Issues*, 4(1), 93.

Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *ILR Review*, 30(2), 175-187.

Albanesi, S., & Şahin, A. (2013). The gender unemployment gap. Staff Report, Federal Reserve Bank of New York, No. 613.

Arrow, Kenneth (1973). The theory of discrimination, in: O. A. Ashenfelter and A. Rees, (Eds.) *Discrimination in labor markets*, Princeton, N.J.: Princeton University Press, pp. 3-33.

Azmat, G., Güell, M., & Manning, A. (2006). Gender gaps in unemployment rates in OECD countries. *Journal of Labor Economics*, 24(1), 1-37.

Arulampalam, W., Booth, A. L., & Taylor, M. P. (2000). Unemployment persistence. *Oxford Economic Papers*, 52(1), 24-50.

Arulampalam, W. (2001). Is unemployment really scarring? Effects of unemployment experiences on wages. *The Economic Journal*, 111(475), 585-606.

Arulampalam, W., Gregg, P., & Gregory, M. (2001). Unemployment scarring. *The Economic Journal*, 111(475), 577-584.

Becker, Gary S. (1971). *The economics of discrimination*, 2ªEd. Chicago: The University of Chicago Press. (Original edition, 1957.)

Bielby, W. T., & Baron, J. N. (1986). Men and women at work: Sex segregation and statistical discrimination. *American Journal of Sociology*, 91(4), 759-799.

Blau, F. D., & Kahn, L. M. (1997). Swimming upstream: Trends in the gender wage differential in the 1980s. *Journal of Labor Economics*, *15*(1, Part 1), 1-42.

Cejka, M. A., & Eagly, A. H. (1999). Gender-stereotypic images of occupations correspond to the sex segregation of employment. *Personality and Social Psychology Bulletin*, *25*(4), 413-423.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*(3), 573.

Elmeskov, J., & MacFarlan, M. (1993). Unemployment persistence. *OECD Economic Studies*, 57-57.

Gauchat, G., Kelly, M., & Wallace, M. (2012). Occupational gender segregation, globalization, and gender earnings inequality in US metropolitan areas. *Gender & Society*, *26*(5), 718-747.

Gottfredson, L. S. (1981). Circumscription and compromise: A developmental theory of occupational aspirations. *Journal of Counseling Psychology*, *28*(6), 545.

Jackman, R. (2002). Determinants of unemployment in western Europe and possible policy responses. *Paper presented at UNECE's 5ᵗʰSpring Seminar Geneva*, 1-38.

Kögel, T. (2004). Did the association between fertility and female employment within OECD countries really change its sign?. *Journal of Population Economics*, *17*(1), 45-65.

Manning, A., & Swaffield, J. (2005). The Gender Pay Gap in Early Career Wages Growth. *CEP Discussion Paper*, *679*.

Pereira, P. T., & Wemans, L. (2015). Portugal and the Global Financial Crisis: Short-Sighted Politics, Deteriorating Public Finances and the Bailout Imperative. *The Global Financial Crisis and Its Budget Impacts in OECD Nations: Fiscal Responses and Future Challenges*, 231-254.

Petrongolo, B. (2004). Gender segregation in employment contracts. *Journal of the European Economic Association*, *2*(2-3), 331-345.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American Economic Review*, *62*(4), 659-661.

Statistics Portugal (2011). Statistical Yearbook of Portugal 2010, 30th September 2011. Portugal: Statistics Portugal Available at: https://www.ine.pt

Statistics Portugal (2012). Statistical Yearbook of Portugal 2011, 30th September 2012. Portugal: Statistics Portugal Available at: https://www.ine.pt

Statistics Portugal (2013). Statistical Yearbook of Portugal 2012, 30th September 2013. Portugal: Statistics Portugal Available at: https://www.ine.pt

Statistics Portugal (2014). Statistical Yearbook of Portugal 2013, 30th September 2014. Portugal: Statistics Portugal Available at: https://www.ine.pt

Şahin, A., Song, J., & Hobijn, B. (2010). The unemployment gender gap during the 2007 recession. *Current Issues in Economics and Finance,* 16 (2), 1–7 (Federal Reserve Bank of New York).

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3), 355-374.

Wright, D. B., Eaton, A. A., & Skagerberg, E. (2015). Occupational segregation and psychological gender differences: How empathizing and systemizing help explain the distribution of men and women into (some) occupations. *Journal of Research in Personality*, *54*, 30-39.

Wooldridge, Jeffrey M. (2010), Econometric analysis of cross section and panel data, The MIT Press, Cambridge, England.

Wooldridge, Jeffrey M. (2005). Simple solutions to the initial conditions problem in

dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of applied econometrics*, *20*(1), 39-54.

APPENDIX

Table A1 – Variable definitions and expected effects

Variable	Description	Expected effect
Unemp	Unemployed at time of the interview (ILO definition)	-
Age	Age of the individual	Positive
Unemp Spouse	Equals 1 if the spouse is unemployed	Ambiguous
Female	Equals 1 if the individual is a woman	Positive
Number Kids	Number of kids of the individual who were also in the original database	Positive
Experience	Number of years in paid work	Negative
Higher Educ	Equals 1 if the individuals has a degree equivalent to high school or higher	Negative

Notes:

1. Pooled data for 4 waves of the ICOR (2010-2013)
2. Sample size = 3096
3. The information on the real number of kids of each individual was not available. This variable was computed, for individual i, as the sum of individuals in the original sample who had reported i to be his or her parent. Therefore, it is a proxy for the number of kids of the individual i.

Table A2 – Descriptive Statistics of the variables

Variable	Mean	Std. Deviation	Min	Max
Unemp	0,150	0,357	0	1
Age	42,375	10,474	17	66
Unemp Spouse	0,101	0,302	0	1
Female	0,500	0,500	0	1
Number Kids	0,461	0,713	0	5
Experience	23,610	12,335	0	54
Higher Educ	0,386	0,487	0	1

Notes:

1. Pooled data for 4 waves of the ICOR (2010-2013)
2. Sample size = 3096

Table A3 – Descriptive Statistics of the variables for women

Variable	Mean	Std. Deviation	Min	Max
Unemp	0,150	0,357	0	1
Age	41,753	10,015	17	66
Unemp Spouse	0,076	0,265	0	1
Number Kids	0,460	0,689	0	3
Experience	22,318	11,842	0	54
Higher Educ	0,456	0,498	0	1

Notes:

1. Pooled data for 4 waves of the ICOR (2010-2013)
2. Sample size = 1536

Table A4 – Employed population according to main occupation (ISCO-08) in

thousands (%Δ)

Occupations	2011*	2012	2013
Managers	287.7	288	304.8
	(NA)	(-%0.10)	(+%5.83)
Office clerks	375.5	336.2	310.2
	(NA)	(-%10.47)	(-%7.73)
Craft, industry and construction qualified workers	712.8	618.5	540.5
	(NA)	(-%13.23)	(-%12.61)
Plant and machine operators and assemblers	380.3	356.2	354.4
	(NA)	(-%6.34)	(-%0.51)
Unqualified workers	523.6	503.8	495.4
	(NA)	(-%3.78)	(-%1.67)

Notes:
1. Data from Statistics Portugal, Labor Force Survey
2. *Values of ISCO-08 not available for 2010

Table A5 – Transition probabilities (from unemployment or employment to

unemployment)

	2011	2012	2013
Male			
Unemployed $_{t-1} = 0$	0.0552	0.0708	0.0495
Unemployed $_{t-1} = 1$	0.6957	0.8431	0.8060
Total	0.1308	0.1718	0.1795
Female			
Unemployed $_{t-1} = 0$	0.0431	0.0517	0.0547
Unemployed $_{t-1} = 1$	0.6949	0.6909	0.7818
Total	0.1432	0.1432	0.1589

Notes:
1. The total is the proportion of individuals who were unemployed in the sample for the correspondent year
2. *Unemployed* $_{t-1} = 0$ has the proportion of the individuals who are unemployed in t given that they were employed in $t-1$
3. *Unemployed* $_{t-1} = 1$ has the proportion of the individuals who are unemployed in t given that they were unemployed in $t-1$

Table A6 – Panel data models for the probability of unemployment

	Random Effects Probit (1)	Pooled Probit (2)	Random Effects Probit (3)	Pooled Probit (4)
Unemp at t-1	0.603 ***	0.994***	0.614***	1.012***
	(0.158)	(0.101)	(0.158)	(0.099)
Unemp at t-1×Female	-0.288	-0.218	-0.315	-0.253*
	(0.213)	(0.148)	(0.211)	(0.145)
Age	0.196 ***	0.162***	0.184***	0.149***
	(0.047)	(0.036)	(0.045)	(0.034)
Age×Female	-0.061**	-0.045**	-0.040***	-0.023**
	(0.028)	(0.018)	(0.014)	(0.009)
Unemp Spouse	0.403*	0.392**	0.401*	0.383**
	(0.235)	(0.180)	(0.235)	(0.180)
Unemp Spouse×Female	0.375	0.290	0.380	0.300
	(0.303)	(0.248)	(0.304)	(0.245)*
Female	1.620**	1.070**	1.250**	0.720*
	(0.738)	(0.106)	(0.558)	(0.378)
Number Kids	-0.267	-0.145	-0.266	-0.142
	(0.224)	(0.156)	(0.224)	(0.154)
NumberKids×Female	0.434**	0.277**	0.427**	0.263*
	(0.210)	(0.136)	(0.208)	(0.138)
Experience	-0.074 ***	-0.059***	-0.063***	-0.046***
	(0.018)	(0.012)	(0.012)	(0.008)
Experience×Female	0.022	0.022	-	-
	(0.024)	(0.016)		
Higher Educ	-0.506 **	-0.436***	-0.434***	0.344***
	(0.215)	(0.163)	(0.148)	(0.103)
Higher Educ×Female	0.109	0.140	-	-
	(0.291)	(0.209)		
Constant	-3.612	-2.704	-3.415	-2.492
	(0.530)	(0.314)	(0.474)	(0.286)
Sigma_u (rho)	0.956 (0.478)	-	0.964 (0.482)	-
Log likelihood	-695.376	-729.060	-695.777	-730.092
Wald Statistic	303.82	820.31	300.47	812.29
p-value	0.0000	0.0000	0.0000	0.0000
Pseudo-Rsquared	-	0.4426	-	0.4418
Sample size	3095	3095	3095	3095

Notes:
1. Standard errors are in brackets. For the Pooled Probit model we computed cluster-robust standard errors.
2. All models contain year dummies for 2011, 2012 and 2013 and, additionally, controls as specified by the Mundlak device in equation (6).
3. Significance levels: *10%, **5%, ***1%.
4. Age squared and experience squared were not included because they were highly statistically insignificant

Table A7 – Restricted panel data models for the probability of unemployment

	Random Effects Probit (1)	Pooled Probit (2)
Unemp at t-1	0.627***	1.012***
	(0.157)	(0.099)
Unemp at t-1×Female	-0.312	-0.251*
	(0.210)	(0.145)
Age	0.067***	0.049***
	(0.015)	(0.010)
Age×Female	-0.038***	-0.022**
	(0.014)	(0.009)
Unemp Spouse	0.717***	0.530***
	(0.151)	(0.121)
Female	1.221**	0.720*
	(0.552)	(0.376)
Number Kids	-0.312**	-0.208**
	(0.144)	(0.095)
NumberKids×Female	0.432	0.267*
	(0.206)	(0.137)
Experience	-0.062***	-0.047***
	(0.012)	(0.008)
Higher Educ	-0.444***	0.350***
	(0.146)	(0.103)
Constant	-3.555***	-2.6481***
	(0.469)	(0.288)
Sigma_u (rho)	0.957 (0.478)	-
Log likelihood	-697.334	-731.224
Wald Statistic	306.47	810.59
p-value	0.0000	0.0000
Pseudo-Rsquared	-	0.4409
Sample size	3095	3095

Notes:

1. Standard errors are in brackets. For the Pooled Probit model we computed cluster-robust standard errors;
2. Both models contain year dummies for 2011, 2012 and 2013.
3. Significance levels: *10%, **5%, ***1%.
4. Wald p-value is based on a $\chi^2(14)$.
5. The t-ratio of the coefficient associated with Unemp at t-1×Female has a p-value of 13,7%