LIETUVOS BANKAS
EUROSISTEMA

The Changing Nature of Gender Selection into Employment over the Great Recession

Working Papers Series

No. 58 / 2019

ISSN 2029-0446 (ONLINE)

Working Papers Series

No. 58 / 2019

Juan J. Dolado†

Cecilia Garcia-Peñalosa‡

Linas Tarasonis§

January 2019*

† European University Institute. Email: juan.dolado@eui.eu.
‡ Aix-Marseille University, EHESS, CNRS, Central Marseille & AMSE. Email: cecilia.garcia-penalosa@univ-amu.fr.
§ Vilnius University & Bank of Lithuania. Email: linas.tarasonis@evaf.vu.lt.

* We are grateful to G. Jolivet, C. Schluter, G. Spanos, H. Turon and participants in seminars at AMSE, the Banque de France, EUI, SAE Bilbao, Univ. of Bristol, Univ. of Strasbourg and the COSME 2017 Workshop for their helpful comments and suggestions. Financial support from the ADEMU project (EC-H2020, Grant no. 6649396) programme and the Spanish Ministerio de Economía y Competitividad and from the French National Research Agency (ANR-18-CE41-0003-01) is gratefully acknowledged. The views expressed are those of the authors and do not necessarily represent those of the Bank of Lithuania. All errors are our own.

Gedimino pr. 6, LT-01103 Vilnius

www. lb.lt

Working Papers describe research in progress by the author(s) and are published to stimulate discussion and critical comments.

The Series is managed by Applied Macroeconomic Research Division of Economics Department and Center for Excellence in Finance and Economic Research.

All Working Papers of the Series are refereed by internal and external experts.

Abstract

The Great Recession has strongly influenced employment patterns across skill and gender groups. This paper analyzes how the resulting changes in non-employment have affected selection into jobs and hence gender wage gaps. Using data for the European Union, we show that male selection into the labour market, traditionally disregarded, has become positive. This is particularly so in Southern Europe, where dramatic drops in male unskilled employment have taken place during the crisis. As regards female selection, traditionally positive, we document two distinct effects. An *added-worker effect* has increased female labour force participation and hence reduced selection in some countries. In others, selection has become even more positive as a result of adverse labour demand shifts in industries which are intensive in temporary work, a type of contract in which women are over-represented. Overall, our results indicate that selection has become more important among men and less so among women, thus changing traditional gender patterns and calling for a systematic consideration of male non-employment when studying gender wage gaps.

JEL code: J31.
Keywords: Sample selection, gender wage gaps, gender employment gaps.

1 Introduction

While there has been extensive discussion in the academic literature and in the media about the effects of the Great Recession on household income inequality, its impact on gender wage inequality has received much less attention.[1] This is somewhat surprising since industries exhibiting large differences in their use of male and female labour have experienced quite unequal cyclical fluctuations in employment and labour-force participation, both of which are expected to induce sizeable changes in wage differentials by gender. In particular, this has been the case in the European Union (EU), where the last recession in several of its member states has been longer and more severe than in the US and other high-income economies.[2] As a result, the EU constitutes an interesting case study in which to analyze how large shifts in labour demand and labour supply can impinge on gender wage gaps via changes in the way men and women have self-selected into different EU labour markets over the crisis period.

A number of recent reports, most notably OECD (2014), have documented that *raw* gender wage gaps (i.e., based on reported wages by employed workers and unadjusted for characteristics; termed in short as *RG* hereafter) have narrowed in several EU countries during the Great Recession.[3] Several explanations have been provided for this finding. For example, declining RG could be the outcome of women being over-represented in the public sector (where gender gaps are generally lower) and under-represented in other sectors where men tend to earn well and that have been subject to much higher job destruction. Likewise, it has been argued that the intensive use of early retirement policies in some EU countries – to alleviate social pressure against collective dismissals during the slump – could have reduced RG because men are a majority among elderly workers with long professional careers and higher wages. However, these potential rationalizations of lower RG since the onset of the crisis could be restrictive because they do not take into consideration the effects on gender gaps of relevant changes in patterns of non-random selection of male and female workers into employment, as a result of the major cyclical fluctuations that EU labour markets have gone through over this period.

In effect, when comparing wages across two groups, non-random selection into employment can imply that measured wage gaps differ considerably from the gaps that one would obtain if the two populations had experienced the same employment

[1] See, for example, Jenkins et al. (2012)

[2] This is so since the Great Recession in most of the EU not only covers the global financial crisis in 2008-09, but also the subsequent sovereign debt crisis in the Euro area from late 2009 to mid 2012.

[3] More precisely, the gender wage gap is defined in the sequel as the difference between male and female hourly wages in log points.

fluctuations. The observed RG can be larger or smaller than the *potential* wage gap (termed *PG* in the sequel), depending on the sign of selection. The literature usually assumes no selection of the majority group (white, natives, men, etc.) and considers both positive and negative selection of the minority. Thus, as demonstrated by a large body of literature on this topic, accounting for selection becomes essential to obtain a corrected measure of RG that better reflects PG.[4]

Our focus on selection issues is dictated by previous evidence indicating that they have played a key role in explaining EU cross-country differences in gender gaps prior to the last recession. For example, Olivetti and Petrongolo (2008) document that, from the mid 1990s to the early 2000s, gender gaps in southern Mediterranean countries (*Southern EU*, hereafter), based on imputed wage distributions for the working age population, were considerably higher than RG based on reported wages. In contrast, both gaps were fairly similar in central and northern EU countries (*Rest of EU*, henceforth) and in the US. The historically low female labour-force participation (LFP) in Southern EU economies implies positive selection among participating women, as they often have relatively high-wage characteristics. In the rest of the sample, both female and male LFP rates are uniformly high, implying no selectivity, so that observed medians of the female and male reported wage distributions accurately represent their population counterparts. Hence, lacking selection-bias corrections, Olivetti and Petrongolo (2008) convincingly argue that measured RG in Southern EU would appear as being much lower than in other high-income countries.

In view of these considerations, our aim in this paper is to explore whether Olivetti and Petrongolo (2008)'s assessment of gender sorting into EU labour markets before the slump has changed as a result of the large differences in the scale of employment and LFP adjustments that have taken place in different parts of the EU during the recession. To analyze this issue, we use the EU-SILC longitudinal dataset on wages which is available for 13 EU member states, and which covers periods before and after the global financial crisis.

Specifically, in contrast to the previous assessment, we conjecture that selection during the Great Recession may have become more important among men and less so among women, especially in Southern EU. We refer to this phenomenon as "the *changing nature* of selection by gender during the Great Recession".[5] The insight is that, following a massive job shedding in sectors intensive in low-skilled male workers (e.g., in the construction sector in some EU economies), the distribution of

[4] See, amongst others, Heckman (1979), Johnson et al. (2000), Neal (2004), Mulligan and Rubinstein (2008), Olivetti and Petrongolo (2008), and Arellano and Bonhomme (2017).

[5] To the best of our knowledge, Arellano and Bonhomme (2017) is the only paper that documents positive male selection into the labour market. Their focus is on the UK prior to the Great Recession.

observed male wages has become a censored version of the imputed distribution. In parallel, we argue that the observed rise in LFP among less-skilled female women could be due to the so-called "added-worker" effect, whereby inactive women start searching for jobs in order to help restore household income when male breadwinners became jobless.[6] In line with previous findings by Bentolila and Ichino (2008), Bredtmann et al. (2014, Table 2) have shown that this effect is particularly strong in southern Mediterranean countries, due to their less generous welfare states. Hence, combining male job destruction with a rise in female LFP (and possibly in female employment) would lead to a lower (resp. larger) difference between RG and PG among females (resp. males) during the crisis than prior to it.

Two empirical strategies are used to correct for non-random selection in measuring gender wage gaps in EU countries. Following Olivetti and Petrongolo (2008), we first apply the sample-selection correction methodology advocated by Johnson et al. (2000) and Neal (2004). This approach requires credible assumptions on the position of imputed wages for non-employed workers relative to the median (rather than the actual level of missing wages) according to different imputation rules. An advantage of this approach is that it avoids arguable exclusion restrictions often invoked in the standard Heckit approach to extrapolate the distribution below the reservation wage.[7] Yet, a potential caveat is that the reliability of the results using this method hinges on the plausibility of the imputation rules. Hence, to check how robust our findings are under a more conventional control-function approach, we also provide results based on Arellano and Bonhomme's (2017) estimation procedure of quantile wage regressions by gender subject to selectivity corrections. Note that the main reason for using quantile regressions is that our rationalization of changes in the gender wage gap relies on different behaviour of male and female workers at the bottom and at other parts of the wage distribution.

Our empirical findings broadly support the mechanisms outlined above. First, we document that the absence of male selection traditionally assumed in the literature may no longer be a valid description of the data in several EU countries after the Great Recession. Strong evidence of positive male selection is found for Southern EU, and to a lesser extent for Rest of EU. Second, we show that patterns of female selection are

[6] Bredtmann et al. (2014) – using the same database (EU-SILC; see Section 3) and a similar sample period as ours – find evidence of a high responsiveness of women's labor supply to their husband's loss of employment. Given that this evidence is based on the same panel dataset we use here and for a similar sample period (2004-13), in the sequel we take the "added-worker" effect as a given stylised fact for this set of countries.

[7] For example, this might be the case regarding number of children or being married (as proxies for household chores). Such variables are often assumed as only affecting labour-market participation via reservation wages. However, one could argue that they might as well affect effort at market-place work, and therefore productivity and wages.

mixed. The added-worker effect has resulted in increased female LFP at the bottom of the wage distribution, making female selection less favourable than before the Great Recession. However, in those countries where female unemployment rates have also surged (especially among the less skilled) in parallel with rising female LFP, female selection has become more positive than prior to the crisis. In particular, the relative weight of these two effects varies across Southern EU countries. We argue that female selection may have become even more positive in some of these countries due to the highly dual nature of their labour markets. Dualism is reflected by the intensive use of temporary jobs (subject to low or no termination costs), where women are over-represented and which have been massively destroyed during the crisis.

Related literature

This paper contributes to a vast literature on gender outcomes in developed (and developing) countries; cf. Blau et al. (2013) and Goldin (2014) for comprehensive overviews. While most of the research documents historical trends, our paper complements this approach by providing evidence on how sizeable changes in gender wage gaps are shaped at a particularly relevant business cycle phase, as in the case of the recent slump.

The issue of how hourly real wages vary over the business cycle, taking into account differences between observed and unobserved characteristics of workers moving in and out of the labor force over the business cycle, has been studied by Keane et al. (1988) for the US by means of the well-known Heckman (1979)'s self-selection correction techniques.[8] We differ from Keane et al. (1988) in four main respects. First, we focus on gender wage gaps rather than exclusively on male wages. Second, our evidence refers to a cross-country comparison of RG in EU countries, where the evolution and causes of gender gaps has been subject to much less research than in the US (see e.g., Blau et al., 2013). Third, we provide both some theory and new empirical evidence of how the Great Recession may have affected the traditional assessment of difference in selection by gender. Last, we also depart from these authors in choice of econometric approaches. Rather than applying Heckman's conventional control function approach, we make use of the (missing) wages imputation methodology advocated by Johnson et al. (2000) and Neal (2004), as well as the selectivity-correction techniques for quantile regressions proposed by citetarellano2017quantile.

The rest of the paper is organized as follows. Section 2 provides some theoretical underpinning of the main mechanisms at play and derives testable hypotheses in terms of both changes in selection biases and employment by gender. Section 3

[8] See also Bowlus (1995) and Gayle and Golan (2012) for further examples in the gender-gap literature accounting for the dynamics of employment selection over the cycle.

describes the EU-SILC longitudinal dataset used throughout the paper. Section 4 explains our first empirical approach (imputation rules around the median) to compute the potential wage distributions, and discusses its results. Section 5, in turn, presents our second empirical approach (quantile selection models) and its corresponding results. Section 6 interprets the main findings of the paper in the light of the various mechanisms explored in Section 2. Finally, Section 7 concludes. An Appendix provides further details on the construction of hourly wages, while an Online Appendix reports additional results on the quality of alternative imputation procedures, quantile regressions, and further descriptive statistics for the 13 European countries considered in our sample.

2 A Simple Theoretical Framework

2.1 The model

To provide some theoretical underpinning for the mechanisms at play, we start by briefly reviewing the basic effects of selection on the measurement of gender wage gaps. As in Mulligan and Rubinstein (2008), we consider the following equation for the (logged) hourly *potential wage*:

$$w_{it} = \mu_t^w + g_i\gamma_t + \varepsilon_{it} \tag{1}$$

where w_{it} denotes individual i's potential hourly wage in year t, g_i is a gender indicator variable (males have $g = 0$, females have $g = 1$), μ_t^w represents the determinants of wages that are common to *all* workers, and γ_t captures those determinants of female wages common to all women but not applicable to men (including discriminatory practices by employers). Finally, ε_{it} is an error term normalized to have a unit variance (for both males and females) such that $m(\varepsilon_{it}|\, \mu_t^w, g_i) = 0$, where $m(\cdot)$ denotes the (conditional) *median* function.[9]

If we were able to measure potential wages for all men and women, then the *potential* median gender wage gap at year t, PG_t, would be:

$$PG_t \equiv m(w_{it}|g_i = 0) - m(w_{it}|g_i = 1) = -\gamma_t. \tag{2}$$

where we expect $PG_t > 0$ (i.e., $\gamma_t < 0$) on historical grounds (see Olivetti and Petrongolo, 2016).

[9]Consistent with the empirical section, our focus in this section is on median rather than mean gender gaps. This choice is without loss of generality since the results can be rewritten in terms of mean gaps and selection biases. As is well known, in this case the latter become functions of the inverse Mill's ratio, as in Mulligan and Rubinstein (2008).

However, to the extent that selection into employment is not a random outcome of the male and female populations, the observed gender gap in median $(RG)_t$, where the wage comparison by gender is restricted exclusively to *employed* individuals, will differ from the PG_t,namely:[10]

$$\begin{aligned} RG_t &\equiv m(w_{it}|g_i = 0, L_{it} = 1) - m(w_{it}|g_i = 1, L_{it} = 1) \\ &= -\gamma_t + m(\varepsilon_{it}|g_i = 0, L_{it} = 1) - m(\varepsilon_{it}|g_i = 1, L_{it} = 1) \\ &= PG_t + \underbrace{b_t^m - b_t^f}_{\text{selection bias differential}} \end{aligned} \tag{3}$$

where L_{it} is an indicator for whether individual i is employed in year t, and $b_t^m = m(\varepsilon_{it}|g_i = 0, L_{it} = 1)$ and $b_t^f = m(\varepsilon_{it}|g_i = 1, L_{it} = 1)$ are the (median) selection biases of males and females, respectively. These two terms differ from zero to the extent that non-employed males and females have different potential wages than employed ones. As discussed above, Olivetti and Petrongolo (2008) argue that, prior to the Great Recession, $b_t^m < b_t^f$ held in Southern EU countries, so that $RG_t < PG_t$, whereas $b_t^m \simeq b_t^f$ held in Rest of EU countries, and hence $RG_t \simeq PG_t$.

Using (3), the change in the observed RG over time becomes:

$$\Delta RG_t = \Delta PG_t + \Delta b_t^m - \Delta b_t^f. \tag{4}$$

Equation (4) has three terms. The first one ($\Delta PG_t = -\Delta\gamma_t$) is the change in the gender-specific component of wages, which may exist due to changes in gender wage discrimination / relative market valuation of skills / relative human capital accumulation when considering *all* men and women. In addition, the second and third terms in (4) capture the changes in the selection biases of males and females, respectively, which constitute the main focus of this paper.[11]

Traditionally, this setup has been used to predict which females are employed, using both a potential market wage equation determining w_{it}, as in (1), plus an additional equation determining the reservation wage, r_{it}. We extend this conventional setup by including an extra equation determining productivity, x_{it}, to capture labour-demand constraints that can affect both men and women. This leads to the following

[10]The discussion below echoes the well-known arguments on selection biases in the seminal work by Gronau (1974) and Heckman (1979), albeit based on gaps in median wages rather than on average wages as these authors do.

[11]Note that, had we allowed for changes in the variance in the error term ε_{it},an additional term would appear in (4), namely $(b_t^m - b_t^f)\Delta\sigma_t^\varepsilon$, where σ_t^ε is its time-varying standard deviation . This term captures changes in the dispersion of wages which has been shown to play an important role in explaining female selection in the US (see Mulligan and Rubinstein, 2008). Yet, we ignore these changes in the sequel because, as shown in Figure 6 in Appendix A, where wage dispersion is measured by the logarithm of the ratio between wages at the 90th and 10th percentiles, no major trends seem to present over 2004-2012, with the possible exceptions of Greece and Portugal.

three-equation model (where equation (1) is repeated below in (5) for convenience):

$$w_{it} = \mu_t^w + g_i \gamma_t + \varepsilon_{it} \tag{5}$$

$$x_{it} = \mu_t^x + (1+\rho)\varepsilon_{it} \tag{6}$$

$$r_{it} = g_i \mu_t^r + g_i v_{it}, \tag{7}$$

where μ_t^x captures the determinants of the average productivity of a worker, μ_t^r is the female reservation wage (male reservation wage is normalized to zero), $(1+\rho)\varepsilon_{it}$ is a productivity shock, and v_{it} is a reservation-wage shock.[12] We assume that $\rho > 0$ to capture the fact that, when individual productivity changes by $(1+\rho)\varepsilon_{it}$, wages only adjust by ε_{it}, reflecting they are not totally flexible.[13] As mentioned above, note that the productivity equation captures labour demand constraints, namely the fact that some individuals who sort themselves into the labour market may not be able to find a job when wages are partially rigid. Finally, whereas ε_{it} has a continuous support, for the sake of simplicity, it is assumed that the shock v_{it} in the reservation wage equation is equal to zero for males and only takes two values for females: a high one, $\overline{v}$, with probability $p \in (0,1)$ and a low one, $\underline{v}$, with probability $1-p$.

Accordingly, individual i works at time t if her/his reservation wage is higher than her/his potential market wage (labour supply condition), $w_{it} > r_{it}$, and her/his productivity is greater than the wage, leaving a positive surplus for the firm (labour demand condition), $x_{it} - w_{it} > 0$. To simplify the analysis, we set the male reservation wage equal to zero, so that men participate if their productivity shock ε_{it} exceeds a labour supply (LS) cut-off value given by:

$$a_t^{LS}(g_i = 0) = -\mu_t^w, \tag{8}$$

which for a sufficiently large value of $\mu_t^w > 0$ implies that the LS constraint does not bind for males, so that (almost) all men would participate in the labour market, in line with their high LFP.

For women, the labour supply (LS) condition, $w_{it} > r_{it}$, is satisfied if and only if ε_{it} exceeds the following LS thresholds:

$$a_t^{LS}(g_i = 1, v_{it} = \overline{v}) \equiv \overline{a}_t = \mu_t^r + \overline{v} - \mu_t^w - \gamma_t, \tag{9}$$

$$a_t^{LS}(g_i = 1, v_{it} = \underline{v}) \equiv \underline{a}_t = \mu_t^r + \underline{v} - \mu_t^w - \gamma_t. \tag{10}$$

[12] Note that, for convenience, we normalize to unity the coefficient on ε_{it} in the wage equation (5), rather than in the productivity equation (6).

[13] This is particularly the case in most European countries, where unions play a more important role in wage setting than in the US. Our model implies symmetry of positive and negative shocks, although it could be easily generalized to the case in which wages respond more to positive than to negative shocks.

The labour demand (LD) condition, $w_{it} < x_{it}$, holds if and only if ε_{it} exceeds the following labour demand threshold:

$$a_t^{LD}(g_i) \equiv \frac{\mu_t^w + g_i\gamma_t - \mu_t^x}{\rho}. \tag{11}$$

for $g_i = 1, 0$.

The conditions above yield gender-specific lower bounds for ε_{it} implying that *only one* of the two constraints above binds. Since the LS constraint for males has been assumed not to be binding, then the LD threshold $a_t^{LD}(g_i = 0)$ is the only one that binds. By contrast, both constraints may be binding for female workers. For example, in the case of women with a high reservation-wage shock, the LD constraint would be binding if and only if: $a_t^{LS}(g_i = 1, v_{it} = \overline{v}) < a_t^{LD}(g_i = 1)$ or:

$$\frac{\mu_t^x - (\mu_t^w + \gamma_t)}{\overline{a}_t} < \rho. \tag{12}$$

whereas for women with low reservation wage shock, the corresponding LD condition becomes:[14]

$$\frac{\mu_t^x - (\mu_t^w + \gamma_t)}{\underline{a}_t} < \rho. \tag{13}$$

Intuitively, equations (12) and (13) hold when: (i) the potential female wage is high relative to productivity, i.e. when the numerator $\mu_t^x - (\mu_t^w + \gamma_t)$ is small; (ii) the reservation wage is low relative to potential wage, i.e., when $\underline{a}_t$ and $\overline{a}_t$ are high; (iii) the surplus is high, i.e., when ρ is much larger than zero. By contrast, when $\mu_t^x - (\mu_t^w + \gamma_t)$ is high, $\underline{a}_t$ and $\overline{a}_t$ are low and ρ is close to unity, it is likely that $a_t^{LD} < a_t^{LS}$, so that the LS constraint would be the binding one. For example, in more traditional societies (such as those in Southern EU), where the average female reservation wage is high due to cultural norms, and the surplus is low, reflecting lower productivity in these countries, the LS condition will be binding. Conversely, in more modern societies (such as in the Rest of the EU), where the average female reservation wage is low and the surplus is high, the LD condition is the binding one. Moreover, the LS constraint is more likely to affect lower-educated women in all countries. This is so since it is often thought that, relative to their distribution of offered wages, they have a higher reservation wage than higher-educated ones because they are more subject to traditional social norms on the distribution of household tasks.

[14]Note that, since $\underline{a}_t < \overline{a}_t$, the LD condition is more likely to be the binding one for women with a high reservation-wage shock than for women with a low reservation-wage shock.

2.1.1 Male Participation

In order to examine male LFP, for illustrative purposes we make use of the following result concerning the median of a (standardized) normal distribution which is truncated from below (see Johnson et al., 1994). Assuming $\varepsilon_{it} \sim \mathcal{N}[0,1]$ and denoting the c.d.f. of the standardized normal distribution by $\Phi(\cdot)$, then the median, $\underline{m}(a)$, of the truncated from below distribution of ε_{it}, such that $a < \varepsilon_{it}$, is given by:

$$\underline{m}(a) = \Phi^{-1}\left[\frac{1}{2}(1+\Phi(a))\right].$$

Using this result, the observed male wage, for which the LD constraint binds, $a_t^{LS}(g=0) < a_t^{LD}(g=0)$, has a closed-form solution:

$$\begin{aligned} w_t^m &\equiv m(w_{it}|g_i = 0, L_{it} = 1) = m(w_{it}|g_i = 0, a_t^{LD}(g=0) < \varepsilon_{it}) \\ &= \mu_t^w + \underline{m}(a_t^{LD}(g=0)). \end{aligned}$$

Given the properties of $\Phi(\cdot)$, it holds that the $\underline{m}\ (\cdot)$ term is a non-negative increasing function of $a_t^{LD}(g=0)$ which measures the strength of the selection bias, $b_t^m = m(\varepsilon_{it}|g_i = 0, L_{it} = 1) = \underline{m}(a_t^{LD}(g=0))$.

Then, the response of w_t^m with respect to a change in μ_t^x is given by:

$$\frac{dw_t^m}{d\mu_t^x} = \frac{\partial \underline{m}}{\partial a_t^{LD}(g=0)} \times \frac{\partial a_t^{LD}(g=0)}{\partial \mu_t^x} < 0, \tag{14}$$

since $a_t^{LD}(g=0)$ is decreasing in μ_t^x. Hence, if we identify the Great Recession with a drop in productivity, $\Delta\mu_t^x < 0$, then the median of the observed male wage distribution increases, due to a stronger positive selection of males into employment, $\Delta b_t^m > 0$.[15] In other words, less-skilled male workers with lower wages will not show up in the observed wage distribution because they become unemployed, and so the median wage for men will rise.

2.1.2 Female Participation

Under our assumption on the reservation-wage shocks v_{it}, it is easy to check that in the case of women the corresponding median, $\underline{m}(a(v))$, of the truncated-from-below distribution of ε_{it}, such that $a(v) < \varepsilon_{it}$, is given by:

$$\underline{m}(a(v)) = \Phi^{-1}\left[\frac{1}{2}(1+p\Phi(\bar{a}) + (1-p)\Phi(\underline{a}))\right].$$

[15] Note that the converse argument could be used to model the effects of a rise in early retirement. Because older male workers have longer experience and this typically leads to higher wages, early retirement would imply stronger negative selection, $\Delta b_t^m < 0$.

Mutatis mutandis, the female wage among the employed workers is given by:

$$
\begin{aligned}
w_t^f &\equiv m(w_{it}|g_i = 1, L_{it} = 1) = m(w_{it}|g_i = 1, a_t^f(v) < \varepsilon_{it}) \\
&= \mu_t^w + \gamma_t + \underline{m}(a_t^f(v))
\end{aligned}
$$

$$
a_t^f(v) \equiv \begin{cases} a_t^{LS}(g=1; v) & : a_t^{LS}(g=1;v) > a_t^{LD}(g=1) \\ a_t^{LD}(g=1) & : a_t^{LS}(g=1;v) < a_t^{LD}(g=1) \end{cases}
$$

Thus, the observed female wage will depend on which of the LS and LD constraints is binding. Again, the strength of the selection bias for females is measured by the $\underline{m}(\cdot)$ term, that is, $b_t^f = m(\varepsilon_{it}|g_i = f, L_{it} = 1) = \underline{m}(a_t^f(v))$. If the binding constraint is LD, i.e., $a_t^{LS}(g=1;v) < a_t^{LD}(g=1)$, then a reduction in labour productivity ($d\mu_t^x < 0$) during the Great Recession will have the same impact on observed female wages as the one discussed before for male wages, namely:

$$
\frac{dw_t^f}{d\mu_t^x} = \frac{\partial \underline{m}(a_t^f(v))}{\partial a_t^{LD}(g=1)} \times \frac{\partial a_t^{LD}(g=1)}{\partial \mu_t^x} < 0. \tag{15}
$$

That is, observed female median wages will increase due to an even stronger positive selection of women into employment when productivity goes down, since those at the bottom of the wage distribution are the ones losing their jobs.

However, if the LS constraint is the binding one, $a_t^{LS}(g=1;v) > a_t^{LD}(g=1)$, then:

$$
\frac{dw_t^f}{d\mu_t^r} = \frac{\partial \underline{m}(a_t^f(v))}{\partial a_t^{LS}(g=1;v)} \times \frac{\partial a_t^{LS}(g=1;v)}{\partial \mu_t^r} > 0. \tag{16}
$$

Hence, insofar as the downturn has generated an added-worker effect among previous female non-participants in the less-skilled segment of the labour market, this would translate into a reduction in the reservation wage, $\Delta\mu_t^r < 0$. This results in a reduction of the observed female wage due to a less positive selection, $\Delta b_t^f < 0$, since less-skilled women enter the labour market and are able to find a job.

In sum, depending upon which of the two opposite forces (LD and LS constraints) dominates, the observed female wage may increase or fall as a result of the recession.

2.2 Gender-gap scenarios over the Great Recession

The implications of the analysis above results in a number of hypotheses that can emerge (individually or jointly) and which lead to our main testable implications:

- **Hypothesis I**: *Higher job destruction rate among low-skilled workers.*
  - **Hypothesis I$_m$**: If the Great Recession has largely resulted in the shedding of unskilled low-paid jobs in *male* labour-intensive industries, then one

would expect a positive male selection bias ($\Delta b_t^m > 0$). From (4), this implies that $\Delta RG_t > \Delta PG_t$. Furthermore, denote employment rates at time t by E_t^{ij}, where $i = f, m$ denotes gender and $j = u, s$ whether the individual is unskilled or skilled. Then, the employment-rate patterns consistent with this hypothesis would be a decline in the employment rate of unskilled male workers, i.e. $\Delta E_t^{mu} < 0$, and no relevant changes in either skilled male or overall female employment rate, i.e., $\Delta E_t^{ms} = \Delta E_t^f = 0$ respectively.

  - **Hypothesis I$_f$** : The implications are the same as in Hypothesis II$_m$, except that now the focus is on employment reductions only among less-skilled *female* workers. This may be more pronounced in countries with dual labour markets where temporary jobs (in which females are over-represented) can be easily terminated at low cost. It then holds that $\Delta E_t^{fu} < 0$, while $\Delta E_t^m = \Delta E_t^{fs} = 0$ and the female selection bias becomes more positive ($\Delta b_t^f > 0$), so that equation (4) implies $\Delta RG_t < \Delta PG_t$.

- **Hypothesis II**: *Added-worker effect*.

  As documented by Bredtmann et al. (2014), the Great Recession has led to a rise in less-skilled female LFP (particularly in Southern EU labour markets), as a response to a decline in the employment rate of less-skilled men ($\Delta E_t^{mu} < 0$). Two cases can be distinguished, depending on whether higher female LFP translates into higher female employment (the binding constraint is LS) or does not (the binding constraint is LD).

  - **Hypothesis II$_{fe}$.** When the LS constraint binds, then new less-skilled female entrants in the labour market become employed. As a result, female selection becomes *less* positive ($\Delta b_t^f < 0$), while male selection (previously absent) becomes positive ($\Delta b_t^m > 0$) because of the job losses among the less-skilled men. Hence, $\Delta RG_t \gg \Delta PG_t$ since in this case both selection terms in (4) would push the selection gap upwards. In terms of employment rates, this hypothesis would be consistent with $\Delta E_t^{fu} > 0$ and $\Delta E_t^{mu} < 0$.

  - **Hypothesis II$_{fu}$.** If LD is the binding constraint, the increase in female LFP does not translate into a net increase in jobs for less-skilled women, so that their unemployment rate rises. Then, while male selection remains positive ($\Delta b_t^m > 0$), female selection becomes even *more* positive ($\Delta b_t^f > 0$), and therefore ΔRG_t could be larger or smaller than ΔPG_t. This hypothesis would be consistent with $\Delta E_t^{fu} < 0$ and $\Delta E_t^{mu} < 0$.

3 Data

In order to measure both RG and PG, we use the European Statistics on Income and Living Conditions (EU-SILC) data set.[16] This is an unbalanced household-based panel survey which has replaced the European Community Household Panel Survey (ECHPS) as the standard data source for many gender wage gap studies in Europe, including the aforementioned Olivetti and Petrongolo (2008). It collects comparable multidimensional annual micro-data on a few thousand households per country, starting in 2004. Our sample covers the period 2004-20012; that is, it includes observations for years both before and after the Great Recession.

The countries in our sample are classified into two groups: (i) "Southern EU": Greece, Italy, Portugal and Spain, and (ii) "Rest of EU": Austria, Belgium, Denmark, Finland, France, Ireland, The Netherlands, UK, and Norway. It is noteworthy that some large countries in the latter group, notably Germany, are not included in our sample due to lack of longitudinal information on several key variables affecting wages.

We restrict our sample to individuals aged 25-54 as of the survey date, and we use self-defined labour market status to exclude those in self-employment, full-time education, and military service.

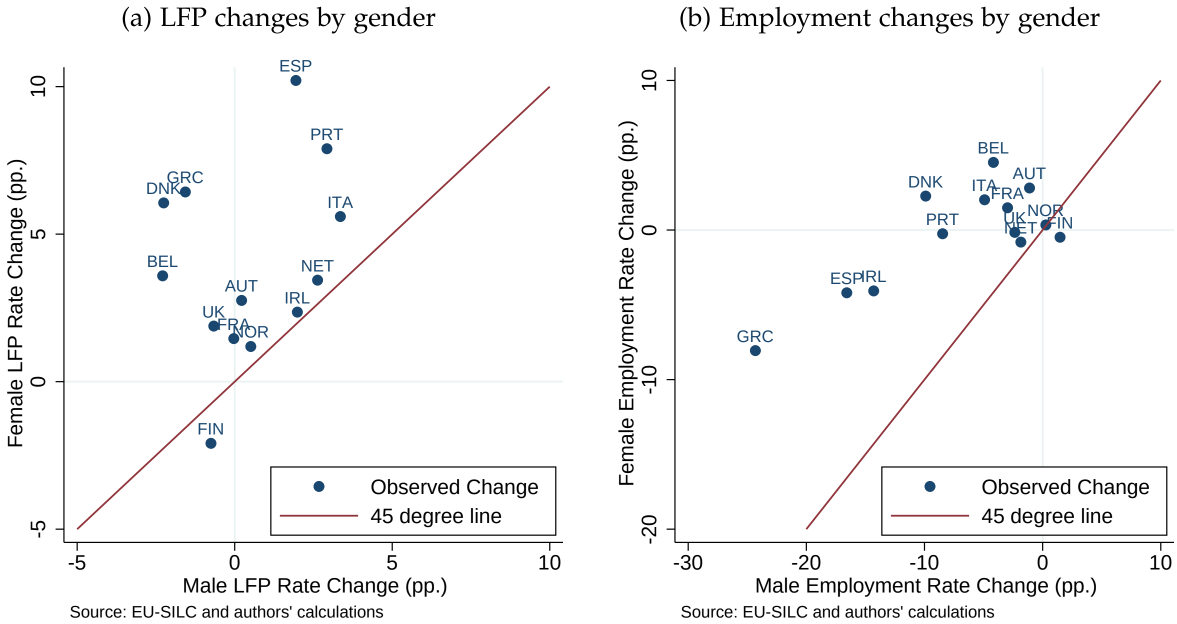
One of the shortcomings of the EU-SILC data is that income information is only available for the income reference period while labour market status and additional variables are recorded at the moment of the interview during the survey year, which for most countries does not cover the same period. In effect, the income reference period corresponds to the previous calendar year for all countries except the UK (where the income reference period is the current year) and Ireland (where the income reference period is the 12 months preceding the interview). We follow a methodology similar to Engel and Schaffner (2012) to derive hourly wages. A detailed account of this procedure is provided in the Appendix A.

The educational attainment categories (no college and college) correspond to ISCED 0-4 and 5-7, respectively. Spouse income is calculated as annual labor income for spouses of respondents. Descriptive statistics are reported in the Online Appendix A. Finally, throughout our empirical analysis, observations are weighted using population weights when available.[17]

[16]Existing literature using EU-SILC data for international comparisons of gender gaps includes Christofides et al. (2013), who use OLS and quantile regressions to document the differences in the gender gap across the wage distribution in a number of countries.

[17]Specifically, we use personal base weights, PB050. For Denmark, Finland, Sweden and The Netherlands, income data is only available for selected respondents. We use personal base weights for selected respondents, PB080, for these countries. Personal weights are not available for Norway and Ireland.

Figure 1: Labour market attachment by gender, 2007-2012.



Before proceeding to the analysis, let us consider gender differences in the participation and employment responses to the recession. As shown in Figure 1a, where changes in female LFP rates (in pp., vertical axis) during the crisis are plotted against changes in male LFP rates (in pp., horizontal axis), most EU countries (Finland is the exception) exhibit a much larger rise in female LFP since 2007 (i.e., at the beginning of the recession) than before. Nonetheless, as pointed out earlier, it should be borne in mind that higher LFP by women does not necessarily translate into female employment gains, if jobs are not available. In effect, according to Figure 1b, where changes in female employment rates (in pp.,vertical axis) are displayed against the corresponding changes in male employment rates (in pp., horizontal axis), both turn out to be negative in almost half of the countries under consideration.[18] For example, Greece, Ireland, Portugal and Spain exhibit much larger drops in male, as compared to female, employment rates (points above the 45^o line), capturing large job destruction in male-intensive industries. However, even within Southern EU countries, one can find some diverging patterns. For example, there are much more muted employment changes in Italy than in the other three southern countries. Economies in Rest of EU (excepting Denmark and Ireland, which also experienced housing bubbles) have in turn followed rather different employment patterns than the southern ones, with much lower male and female job losses.

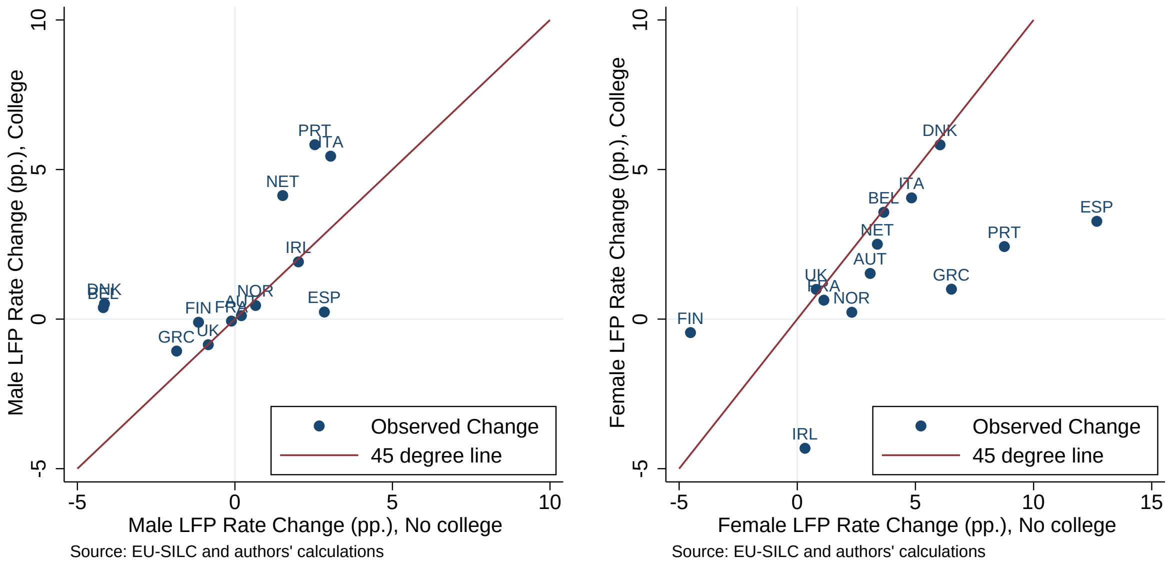
When LFP and employment changes are analysed by workers' educational attain-

[18]Employment rates are defined as the ratios between employment and the labour force.

ment (for males in Figure 2a and 3a and for females in Figure 2b and 3b), it becomes clear that the fall in employment has been more pronounced among less-educated (no-college) male workers. This has been particularly the case in Ireland and Spain, as a result of the bursting of their respective housing bubbles, as well as in Greece, due to the sovereign debt crisis. Likewise, regarding LFP, it can be seen that most of the gains in participation in Southern EU countries are due to females with lower educational attainments. Overall, we take this preliminary evidence as yielding some support for the mechanisms underlying the set of hypotheses outlined above.

Figure 2: Cross-country changes in LFP by gender and skill, 2007-2012.

(a) Males (b) Females

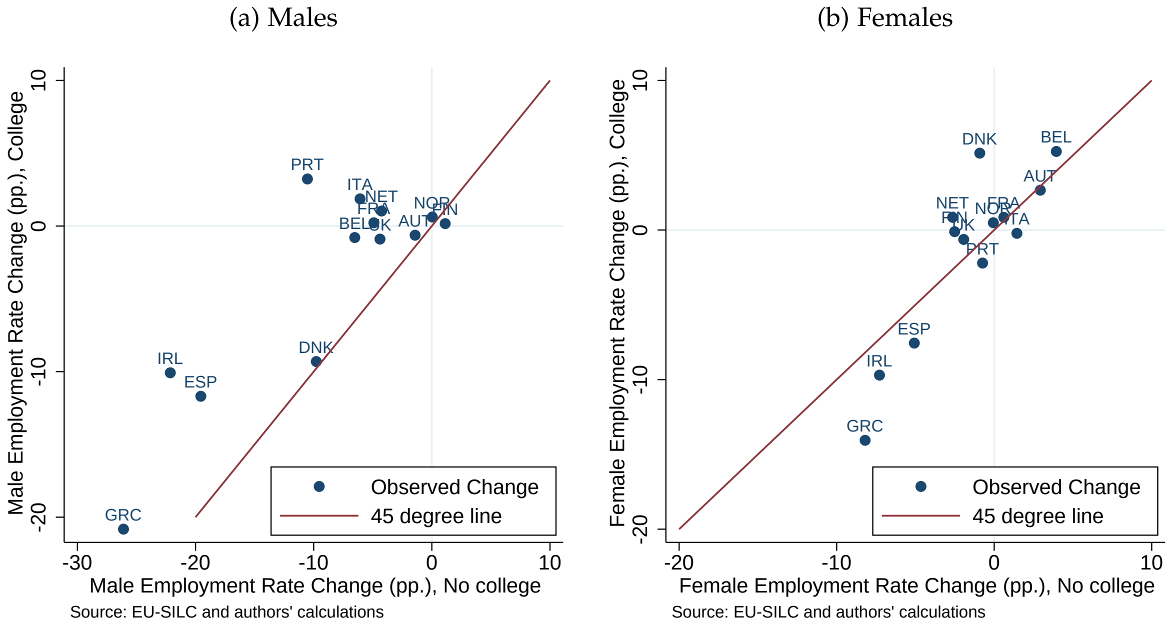


4 Evidence from Imputation Rules around the Median

As mentioned above, the use of median (or mean) wage regressions to estimate coefficients involved in the components μ_t^w and γ_t in equation (1) is bound to yield biased estimates of gender gaps since wages w_{it} are only observed for the employed and are missing for the rest of the sample. To the extent that employed males and females have have different potential wage distributions than non-employed ones, this will result in selection biases.

As discussed in Olivetti and Petrongolo (2008), the median estimator using a transformed dependent variable which equals w_{it} for those who are employed at time t, $L_{it} = 1$, and some arbitrary low or high imputed value, $\underline{w}_t$ and $\overline{w}_t$ respectively, for those in the non-employment, $L_{it} = 0$, will yield an unbiased estimator of the me-

Figure 3: Cross-country changes in employment rates by gender and skill, 2007-2012.



dian gap in potential wages insofar as the missing wage observations are imputed on the correct side of the median. To understand this procedure, let us consider the following illustrative linear wage equation:

$$\omega = \beta_0 + \beta_1 g + \epsilon, \tag{17}$$

where ω is the (logged) potential wage of an (atomistic) agent in a very large (continuous) sample of individuals, β_0 is an intercept, β_1 is the parameter capturing the pay gap, g is a gender dummy, and ϵ is a disturbance term with support $(-\infty, +\infty)$ and c.d.f. $F(\cdot)$, such that $m(\epsilon|g) = 0$. Let $\widehat{\beta} = (\widehat{\beta}_0, \widehat{\beta}_1)'$ be the hypothetical least absolute deviations (LAD) regression estimators based on potential wages, namely $\widehat{\beta} \equiv$ arg min $\int_{-\infty}^{\infty} |\omega - \beta_0 - \beta_1 g| dF(\epsilon)$. Suppose now that wages are only observed for the employed, while the missing wages for the non-employed fall completely below the median regression line, i.e., $\omega < \widehat{\omega} \equiv \widehat{\beta}_0 + \widehat{\beta}_1 g$, that is, $F(m|g, L = 0) = 1$.[19] Then, defining a transformed dependent variable y such that it equals the observed wage w for $L = 1$ and an arbitrarily low value $\underline{w}$ (with $\underline{w} < \widehat{\omega}$) for $L = 0$, the LAD estimator of the median of y, denoted as $\widehat{y}$, verifies:[20]

$$\widehat{y} = \arg\min_{\widehat{y}} \left[\int_{-\infty}^{\underline{w}} |\underline{w} - \widehat{\omega}| dF(\epsilon) + \int_{\underline{w}}^{\widehat{y}} |w - \widehat{\omega}| dF(\epsilon_i) + \int_{\widehat{y}}^{\infty} |w - \widehat{\omega}| dF(\epsilon) \right].$$

[19]Similar arguments as below would apply if all the missing observations happen to be above the median regression lines, with y being defined as $\overline{w}$ when $L = 0$.

[20]See Bloomfield and Steiger (2012)

Differentiating this object function w.r.t. $\widehat{y}$ yields the following f.o.c.:

$$[F(\underline{w}) + F(\widehat{y}) - F(\underline{w})] - [1 - F(\widehat{y})] = 0,$$

that is, $F(\widehat{y}) = 0.5$, whereas the f.o.c. for the LAD estimator of the median of potential wages verifies $F(\widehat{\omega}) = 0.5$. Hence, it follows that $\widehat{y} = \widehat{\omega}$.

Our first estimation approach relies on this methodology.[21] Thus, we compute median gender gaps as well as the effects of selection into non-employment on the basis of wage imputations that require only assumptions on the position of the imputed wage with respect to the median of the gender-specific wage distribution.

We use a small number of observable characteristics, X_i to make assumptions about the position of the imputed wage with respect to the median of the gender-specific wage distribution. We define a threshold for X_i below which non-employed workers would earn wages below the gender-specific median, and another threshold above which individuals would earn above-median wages.

Specifically, our core specification is based on standard human capital theory, and uses both observed *educational* attainment and labour market *experience* ("Imputation on EE") to predict the position of the missing wages. As explained above, the imputed dependent variable is set to equal a low value, $\underline{w}_t$, if an individual has low education and limited labour market experience, and a high value, $\overline{w}_t$, when an individual is highly educated and has extensive labour market experience. In addition, to take into account non-employed individuals with low (high) education and long (limited) experience, we follow Olivetti and Petrongolo (2008) in fitting a probit model for the probability that the wage of employed individual lies above the gender specific median, based on education, experience and its square. In this way, predicted probabilities for the non-employed are obtained. An imputed sample using all individuals in the sample is then constructed using these predicted probabilities as sample weights. The reference wage is calculated for the base sample with wage observations from adjacent waves.

In a second specification, we exploit the hypothesis of assortative mating, which implies a positive correlation between *spousal incomes* within the household ("Imputation on SI"). Further details on the precise rules of imputation that we use are provided below.

These methods of imputation of missing wages follow an educated guess. Two procedures are used to assess the goodness of alternative guesses. Following Olivetti and Petrongolo (2008), the first procedure (Goodness Method 1) uses wage informa-

[21] As noted, this approach is closely related to Johnson et al. (2000) and Neal (2004).

tion for non-employed individuals from other waves in the panel in which individuals report having received a wage. In this way, it is possible to check whether the relative position as regard the median of imputed wages using information of the aforementioned demographics corresponds to the actual one when the wage is observed. The second method (Goodness Method 2) takes all employed workers and computes the fraction of those with the relevant personal characteristics and wage observations on the correct side of the median as predicted by the imputation rule.

As an alternative imputation method which does not rely on using arbitrary assumptions based on observable characteristics, as above, Olivetti and Petrongolo (2008) exploit the panel nature of the data so that, for all those not employed in year t, they recover their wages from the nearest wave, t'. As argued by Olivetti and Petrongolo (2008), the identifying assumption is that the wage position with respect to the median when an individual is not employed can be proxied by the observed wage in the nearest wave. While this procedure, denoted as Imputation on Wages from Other Waves ("WOW") relies exclusively on wages and therefore has the advantage of incorporating selection on time-invariant unobservables, it has the disadvantage of not providing any wage information on individuals who never worked during the sample period. Thus, this method will be relatively conservative in assessing the effects of positive selection in the countries with a relatively low labour market attachment of females. Moreover, since the panel dimension of our data set is relatively short, this procedure yields less satisfactory results in terms of goodness of fit.[22] Consequently, we relegate this analysis to the Online Appendix.

4.1 Results

4.1.1 Imputation based on education and experience

Table 1 presents results for our preferred Imputation EE method, based on education and experience. As discussed above, two education categories are defined: those individuals with upper secondary education or less are considered to be "less-education", while those with some tertiary education are defined as "high-educated". Similarly, we define as "low (high) experienced individuals" those with less than (at least) 15 years of work experience.

The upper and lower panels of Table 1 present the results for the four Southern EU and the nine Rest of EU countries considered in our sample. We report the RG and PG in levels, the selection biases and employment rates by gender in 2007 (at the onset of the recession), as well as the corresponding changes between 2007 and

[22]The longitudinal component of EU-SILC allows to follow each household for four years, with the exception of France, where each household is followed for eight consecutive years.

Table 1: Median Wage Gaps under Imputation on Education and Experience

	Levels in 2007						Changes over 2007-2012					
	Raw Wage Gap	Potential Wage Gap	Selection Bias		Employment Rate		Raw Wage Gap	Potential Wage Gap	Selection Bias		Employment Rate	
			M	F	M	F			M	F	M	F
Southern Europe:												
Greece	.182	.450	.016	.283	.853	.542	-.076	-.036	.069	.109	-.257	-.111
Italy	.035	.266	.029	.260	.849	.558	.053	.017	.008	-.028	-.057	.002
Spain	.132	.254	.012	.134	.889	.674	-.027	-.021	.087	.094	-.167	-.078
Portugal	.172	.229	.030	.087	.838	.708	-.038	-.067	.011	-.018	-.084	-.014
Mean	**.130**	**.300**	**.022**	**.191**	**.857**	**.620**	**-.022**	**-.027**	**.044**	**.039**	**-.141**	**-.050**
Rest of Europe:												
Austria	.192	.299	.009	.117	.879	.711	.012	-.021	.000	-.033	.003	.011
Belgium	.074	.142	.021	.089	.866	.742	-.019	-.063	.004	-.040	-.034	.031
Ireland	.170	.296	.029	.155	.851	.668	-.040	-.064	.002	-.022	-.139	-.076
United Kingdom	.247	.302	.009	.063	.942	.806	-.065	-.049	.010	.026	-.035	-.025
Netherlands	.158	.190	.003	.034	.933	.802	-.054	-.043	-.001	.010	-.031	-.018
France	.114	.159	.006	.051	.917	.816	.005	-.015	.008	-.012	-.034	.000
Finland	.203	.209	.013	.019	.897	.864	-.072	-.072	.003	.003	-.020	-.038
Denmark	.116	.121	.001	.006	.985	.941	-.072	-.064	-.001	.007	-.126	-.045
Norway	.154	.161	.002	.009	.975	.913	.027	.014	-.003	-.016	-.015	.004
Mean	**.158**	**.209**	**.010**	**.060**	**.916**	**.807**	**-.031**	**-.042**	**.003**	**-.008**	**-.048**	**-.017**

Source: EU-SILC and authors' calculations. Note: Selection bias = an increase in observed wage due to selection. Wage imputation rule: Impute wage $<$ median when non-employed and education $\leq$ upper secondary and experience $<$ 15 years; impute wage $>$ median when non-employed and education $\geq$ higher education and experience $\geq$ 15 years.

2012. Selection biases are measured as a percentage decrease in the median wage once wages of non-employed individuals are imputed. In line with the results of Olivetti and Petrongolo (2008), Southern EU countries exhibit a greater employment gap and a much stronger female selection bias than in Rest of EU. For instance, the average female selection bias in the former group amounts to 19 pp. out of the 30 pp. yielded by PG (i.e., 60%), whereas it reaches only 6.0 pp. out of 21 pp. (i.e., 27%) in the latter. In general, female selection biases are fairly small in Rest of EU (bottom panel). The exceptions are Austria, Belgium and Ireland, which have the lowest female employment rates (between 65% and 75%) in this set of countries. In spite of having similar average selection biases (2.2 pp. against 1.0 pp.), male biases are also higher in southern Mediterranean countries, a finding which is again consistent with their lower aggregate employment rates.

Regarding changing patterns in selection biases since 2007, two findings are noteworthy. The first is that the female selection bias has increased on average by 3.9 pp. in Southern EU while it has hardly changed in Rest of EU (-0.8 pp.). However, patterns differ among Southern EU countries in interesting ways. On the one hand, female selection biases experience substantial reductions in Italy and Portugal, where the fall in female employment rates is fairly small. Given the strong reduction in male employment rates (-5.7 pp. and -8.4 pp.), this finding is not only consistent with the added-worker hypothesis, but it is also clearly indicative that increases in female

LFP in these two countries have been matched by similar increases in female employment. Conversely, female employment has fallen sharply in Greece and Spain, implying that a downward shift in male and female labour demand is the dominant force in these countries. Hence, selection biases in both countries have become stronger (more positive).

As reported in the Online Appendix (see Table A.2 in section A), female LFP rates have increased in the four Southern EU economies and, in general, these changes have been stronger among less-educated female workers. It is worth noting, however, that the largest drops in female selection in our sample of Rest of EU countries have taken place in Austria and Belgium (bottom panel), since these are the two economies where female employment rates have risen the most. In both cases, the increase in female employment is associated with a similar decline in male employment which has improved the employability of both high- and less-educated women.

Table 1 also indicates that the male selection bias in the Southern EU has increased on average by much more (4.4 pp.) than in Rest of EU (0.3 pp.). Among the former set of countries, the rise in male selection is largest in Greece and Spain (in line with large drops in less-skilled male employment of 27.6 pp. and 19.2 pp., respectively), whereas in Portugal wage flexibility imposed by the memorandum of understanding with the "Troika" and out-migration have reduced job shedding of less-skilled men.[23] Overall, the degree of post-GR selection bias in considerable, reaching around 4 pp. in both Italy and Portugal, 8.5 pp. in Greece, and finally 9.9 pp. in Spain.

When we focus on changes in wage gaps over the Great Recession, we see that, in line with the evidence that gender gaps have fallen over this period, RG have fallen by 2.2 pp. and 3.1 pp. in Southern EU and Rest of EU, respectively, and that accounting for selection accentuates the decline by about 0.5 pp. and 1.0 pp., respectively. Note, however, that while Rest of EU countries share similar patterns in RG, there are substantial variations among Southern EU countries. For example, selection has become more positive for men and less positive for women in Portugal, while somewhat similar results hold for Italy. Hence, in both instances $\Delta RG > \Delta PG$, and these two southern economies provide good illustrations of labour markets where LS is the binding constraint. Conversely, accounting for selection either induces $\Delta RG < \Delta PG$ in Greece, and $\Delta RG \simeq \Delta PG$ in Spain. This reflects that female selection bias has increased similarly or even more than male selection, due to adverse labour demand shifts. As a result, these two countries provide the best examples of labour markets where the binding constraints is LD, rather than LS.

[23]Note that amongst the core countries, only the UK exhibits a sizeable increase; see Arellano and Bonhomme, 2017 for a discussion of the UK case.

Table 2: Rate and Goodness of Imputation on Education and Experience

	2007						2012					
	Imputation Rate		Goodness Method 1		Goodness Method 2		Imputation Rate		Goodness Method 1		Goodness Method 2	
	M	F	M	F	M	F	M	F	M	F	M	F
Southern Europe:												
Greece	.43	.71	.93	.86	.84	.85	.45	.63	.72	.80	.83	.82
Italy	.54	.74	.81	.74	.70	.69	.51	.70	.85	.75	.72	.74
Spain	.41	.66	.79	.71	.75	.80	.73	.73	.70	.69	.73	.77
Portugal	.39	.54	.63	.56	.71	.77	.29	.40	.68	.60	.74	.80
Mean	**.44**	**.66**	**.79**	**.72**	**.75**	**.78**	**.50**	**.61**	**.74**	**.71**	**.76**	**.78**
Rest of Europe:												
Austria	.34	.57	.89	.70	.76	.80	.33	.54	.80	.70	.83	.80
Belgium	.39	.58	.81	.88	.79	.80	.47	.64	.82	.78	.77	.81
Ireland	.41	.54	.92	.87	.83	.81	.40	.45	.73	.65	.73	.78
United Kingdom	.42	.50	.36	.62	.74	.74	.41	.55	.94	.61	.76	.70
Netherlands	.39	.64	.55	.94	.81	.75	.50	.59	.92	.91	.82	.77
France	.44	.64	.85	.79	.80	.79	.44	.70	.68	.67	.79	.80
Finland	.58	.47	.95	.85	.76	.78	.54	.45	.74	.70	.78	.73
Denmark	.21	.43	.63	.75	.66	.76	.23	.57	.13	1.00	.72	.77
Norway	.40	.40	.79	.71	.75	.80	.33	.45	.70	.69	.73	.77
Mean	**.40**	**.53**	**.75**	**.79**	**.77**	**.78**	**.41**	**.55**	**.72**	**.75**	**.77**	**.77**

Source: EU-SILC and authors' calculations. Note: Wage imputation rule: Impute wage $<$ median when non-employed and education $\leq$ upper secondary and experience $<$ 15 years; impute wage $>$ median when non-employed and education $\geq$ higher education and experience $\geq$ 15 years. Imputation Rate = proportion of imputed wage observations in total non-employment. Goodness Method 1 = proportion of imputed wage observations on the same side of the median as wage observations from other waves in the panel. Goodness Method 2 = proportion of employed workers on the same side of the median as predicted by the imputation rule.

Finally, Table 2 reports results on our two measures of goodness of fit, computed for men and women separately, for the years 2007 and 2012. We report both the imputation rates for each year and the share of imputations that place the individual on the correct side of the median.[24] Both measures indicate a satisfactory goodness of fit for about 75% of the individuals of either gender in our sample. Furthermore, there is no indication that we do a better job in imputing female than male missing wages.[25]

4.1.2 An illustration of the mechanisms at play: Portugal *vs.* Spain

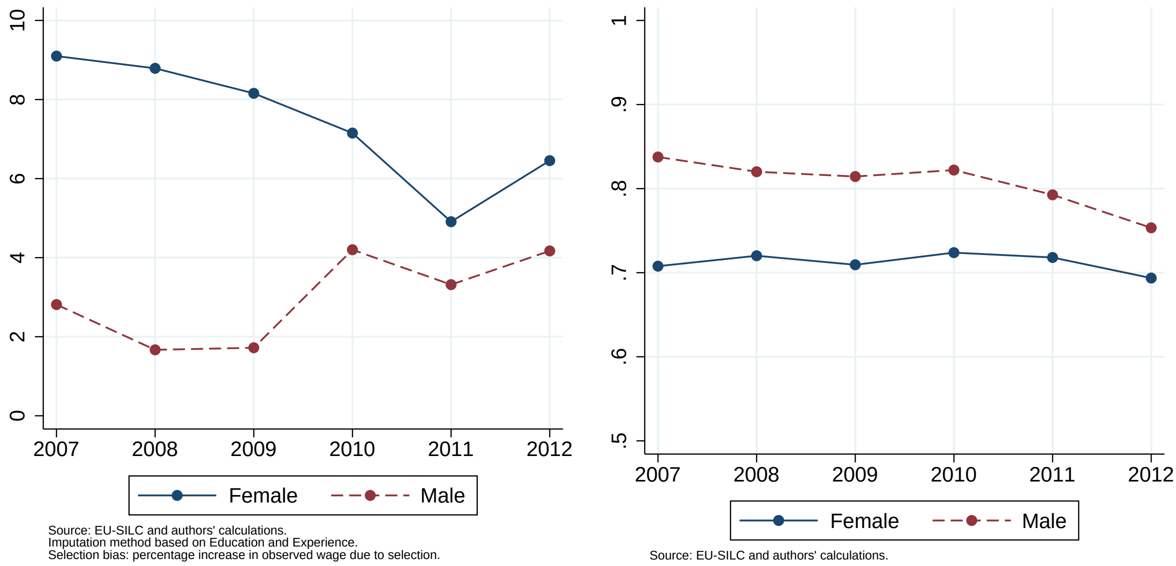
To understand how the contrasting LS and LD constraints operate in practice, we focus on the case of Portugal and Spain, two neighbouring Mediterranean countries badly hit by negative shocks during the recession. As can be observed in Figures

[24] Recall that Method 1 compares our imputation with the positioning implied by looking at the wage observed for the individual in other waves, while Method 2 computes the proportion of employed workers which are on the same side of the median as would be implied if we applied our imputation rule to them.

[25] In order to check the robustness of our imputation method, the Online Appendix B reports estimates based on a probit model. The results are qualitatively similar to our findings in Table 1.

Figure 4: Selection bias and employment rates by gender, Portugal, 2007-2012.

(a) Selection Bias (b) Employment Rate



2 and 3 above, while less-skilled male workers suffered massive job losses during the slump, less-educated women who were previously inactive (especially married ones) increasingly searched for jobs in both countries. However, in parallel with a rise in female LFP, it is well known that many unskilled jobs were destroyed during the slump, particularly in Spain, where the unemployment rate rose from 9% in 2008 to 26% in 2012. This implies that, while only adverse LD shifts (i.e., higher job destruction) apply to males, both LD and LS considerations are likely to have been relevant for women.

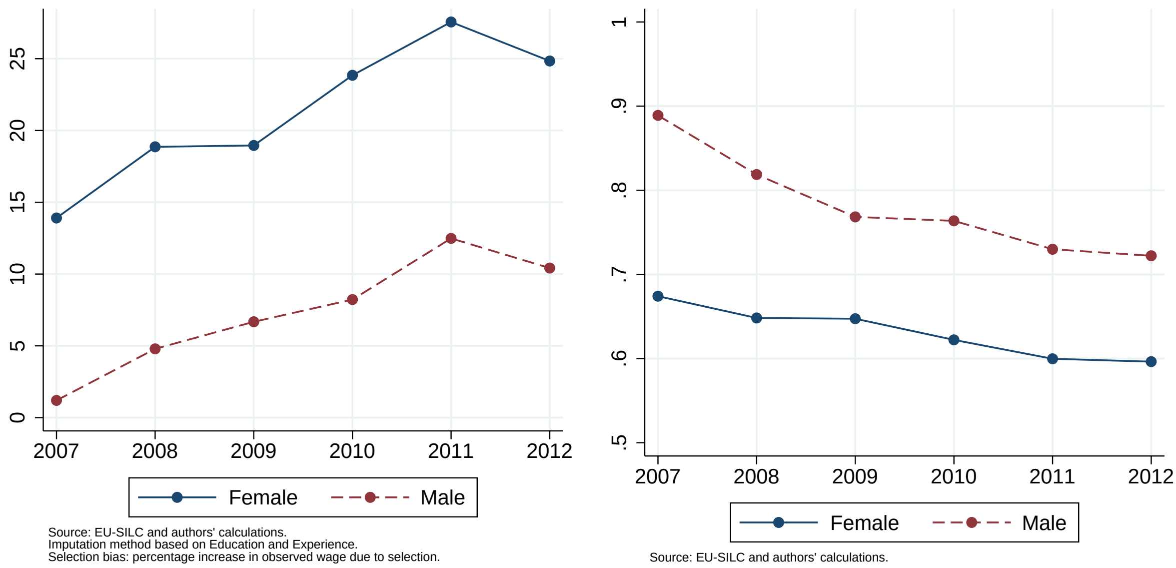
This is illustrated by the left-hand-side (LHS) panels in Figures 4 and 5, which display selection biases for males and females in both countries from 2007 to 2012, computed using imputation based on employment and experience. For comparison, the RHS panels of these two figures display employment rates (shares of occupied in the population of working age). As can be seen, male selection (dashed line) increases drastically in both countries. Yet, while female selection declines in Portugal, it rises in Spain.

The different behaviour of female selection biases between these two countries is likely due to the fact that, while both female and male employment rates collapsed in Spain, only male employment declined in Portugal. The better performance of the Portuguese labour market can be attributed to its larger wage flexibility prior to 2012, as well as to its less dualized nature (see Dolado, 2016). At any rate, given that employment adjustment in Spain was mainly borne by the termination of temporary

Figure 5: Selection bias and employment rates by gender, Spain, 2007-2012.

(a) Selection Bias

(b) Employment Rate



contracts with low or no dismissal cost, where women are over-represented, this evidence seemingly indicates that the rise of female LFP (a positive LS shift) has been more than offset by an even larger reduction in female employment (a negative LD shift). Hence, to the extent that those women who retained their jobs were favourably selected, an increasing, rather than decreasing, female selection bias arises in this country.[26]

4.1.3 Imputation based on spousal income

As mentioned above, a number of alternative imputation methods can be used. This subsection reports the analysis on spousal income, while the Online Appendix provides a further robustness case using wages from other waves.

Under the assumption of assortative matching in marriages, spousal income could be a good proxy for an individual's earning capacity. Hence, we impute a wage below (above) the median to those who are non-employed and whose spouses have earnings that are in the bottom (top) quartile of the gender and year specific earnings distribution. Table 3 presents the results of this imputation method. The main findings of Imputation on SI echo those based on Imputation on EE, although they tend to be less strong, probably due to a weaker performance of Imputation SI in terms of goodness of fit, as can be seen from Table A.5 in the Online Appendix B, which

[26] As will be discussed further in Section 4.1, similar patterns hold in Greece, a country whose cumulated collapse in GDP of more than 25% during the GR meant even more dramatic employment losses than in Spain.

Table 3: Median Wage Gaps under Imputation on Spousal Income

	Levels in 2007				Changes over 2007-2012			
	Raw Wage Gap	Potential Wage Gap	Selection Bias M	Selection Bias F	Raw Wage Gap	Potential Wage Gap	Selection Bias M	Selection Bias F
Southern Europe:								
Greece	.182	.321	.016	.154	-.076	-.039	.049	.086
Italy	.035	.107	.013	.085	.053	.032	.011	-.010
Spain	.132	.179	.007	.054	-.027	-.057	.033	.003
Portugal	.172	.205	.026	.059	-.038	-.073	.021	-.014
Mean	**.130**	**.203**	**.015**	**.088**	**-.022**	**-.034**	**.028**	**.016**
Rest of Europe:								
Austria	.192	.221	.012	.041	.012	.013	-.001	.000
Belgium	.074	.093	.013	.032	-.019	-.036	.004	-.013
Ireland	.170	.235	.031	.096	-.040	-.071	.074	.044
United Kingdom	.247	.268	.014	.035	-.065	-.052	.009	.023
Netherlands	.158	.151	.003	-.003	-.054	-.052	.005	.006
France	.114	.127	.004	.018	.005	-.007	.003	-.009
Finland	.203	.202	.004	.003	-.072	-.062	.003	.013
Denmark	.116	.115	.001	.000	-.072	-.071	.006	.007
Norway	.154	.155	.000	.001	.027	.025	.007	.005
Mean	**.158**	**.174**	**.009**	**.025**	**-.031**	**-.035**	**.012**	**.008**

Source: EU-SILC and authors' calculations. Note: Selection bias = an increase in observed wage due to selection. Wage imputation rule: Impute wage < median when non-employed and spouse income in bottom quartile; impute wage > median when non-employed and spouse income in top quartile.

indicates both a lower imputation rate and worse fit. As before, in 2007 we observe a larger selection in Southern EU countries than in Rest of EU, and that this selection is particularly strong for women. The changes that have occurred during the recession are similar for men, as male selection has increased in all of them, but not regarding women: female selection has increased in Greece and, very slightly, in Spain, while it has declined in Italy and Portugal. For the EU core, we find again little change in female selection, while the average increase in male selection is 1.2 pp., mainly driven by its large rise in Ireland.

5 Evidence from Quantile Selection Models

Since our estimates of selection biases so far have exclusively relied on the credibility of the chosen criteria to impute missing wages in a context of short panels and large fraction of people who never worked throughout the panel, it seems convenient to compare these results to those obtained from a more conventional control-function approach, taking advantage of the longitudinal structure of the data.

Recall that the key ingredients of our theoretical argument are that male job destruction has mostly affected less-skilled workers (i.e., those in the lower part of the wage distribution) and that changes in female LFP and employment have also taken place in that segment of the working-age population - possibly as a result of an added-worker effect. Therefore, it is natural to implement selection correction in a *quantile* regression framework rather than in a standard Heckit setup. The insight is that, if our interpretation is correct, we should observe positive, rather than zero, selection bias at the lower quantiles of the observed male wage distribution. By the same token, selection bias should be more positive in the female distribution if the adverse shifts in LD dominate the favourable shifts in LS (due to the added-worker effect) or, conversely, less positive when LS happens to be the binding constraint. To do so, we apply the methodology recently advocated by Arellano and Bonhomme (2017; AB hereafter).

In AB's (2017) quantile model, sample selection is modeled via the bivariate cumulative distribution function, or *copula*, of the errors in the wage and the selection equations. In particular, the following selection model is considered for the latent wage of each individual of gender g ($g = m, f$), labeled as w^{*g}, and their decision to accept a job:

$$
\begin{aligned}
w^{*g} &= X^{g'}\beta^{g}(U), && (18)\\
D^{g} &= \mathbf{1}\{V \leq p(Z^{g})\}, && (19)\\
w^{g} &= w^{*g} \text{ if } D^{g} = 1, && (20)
\end{aligned}
$$

where $\beta^{g}(U)$ in (18) is increasing in U which is a random variable uniformly distributed on the unit interval, independent of X^{g}, i.e., the set of covariates determining wages, such that $Q(\tau, X^{g}) = X^{g'}\beta^{g}(\tau)$ is the τ-th conditional quantile of w^{*g} given X^{g}. Moreover, (19) is the selection equation with $Z^{g} = (X^{g},\ B^{g})$, where B^{g} are those extra covariates which appear in the participation equation but not in the wage equation, and V is the rank of the error term in this equation, which is also uniformly distributed on $(0, 1)$. Assuming that (U, V) is jointly statistically independent of Z^{g} given X^{g}, denoting by $C(u, v)$ the c.d.f. of (U, V), and finally assuming that $p(Z^{g}) = \Pr(D^{g} = 1|\ Z^{g}) > 0$, the presence of dependence between U and V is the source of sample selection bias. In particular, this dependence is captured by $G(\tau, p; \rho^{g}) = C(\tau, p; \rho^{g})/p$ which is the the conditional copula of U given V, defined on $(0, 1) \times (0, 1)$. Then, AB (2017) show that

$$\beta^g(\tau) = \arg\min_{b(\tau)} E\left[\left(D^g(G_{\tau Z^g}(w^g - X^{g'}b^g(\tau))^+ + (1 - G_{\tau Z^g})(w^g - X^{g'}b^g(\tau))^-\right)\right],$$

where $a^+ = \max(a, 0)$, $a^- = \max(-a, 0)$, and $G_{\tau Z^g} = G(\tau, F^{-1}(z^{g\prime}\gamma^g); \rho^g)$ denotes the rank of $X^{g'}\beta^g(\tau)$ in the selected sample $D^g = 1$, conditional on $Z^g = z^g$. Since the above optimization problem is a linear program, given γ^g and ρ^g, the parameters $\beta^g(\tau)$ can be estimated in a τ-by-τ fashion by solving linear programs, as with a conventional check function (see Koenker and Bassett, 1978) in standard quantile regressions. The only difference is that, in the latter regressions, τ replaces $G_{\tau Z^g}$; in other words, correcting for selection in quantile regressions implies that one needs to rotate the check function depending on Z^g. AB (2017) suggest two previous steps in order to compute $\beta^g(\tau)$: estimation of the propensity score $p(Z^g)$ in (19) (e.g., using a probit model) and estimation by means of a grid-search GMM of the degree of selection (i.e., the copula parameter ρ) using a Frank copula, though they also cover more general cases.

Using the method described above, we estimate wage quantile regressions separately for male and female wages, allowing for sample selection using EU-SILC unbalanced panel data for 2007-2012. The dependent variable is the log-hourly wage, covariates X^g contain potential experience (age minus years of schooling) and its square, marital status, our two education indicators discussed earlier, year indicators, and a set of dummies for region of residence (NUTS). As for B^g(determinants of participation that do not affect wages directly), we take the number of children in 6 age brackets and their interaction with marital status, plus a dummy variable of whether the corresponding spouse lost his/her job in the previous year interacted with marital status. Note that, if the added-worker effect holds, we would expect a positive effect of this last variable on the probability of participation.

Table 4: Quantile Regression Estimates Corrected for Selection

	Males			**Females**		
Quantile	20%	50%	80%	20%	50%	80%
			Portugal			
Change in Selection Bias	.021	.013	.009	-.056	-.040	.016
(Participation eqn: SPUnem)	.012	-.005	-.011	.027	.013	.006
	(.017)	(.016)	(.019)	(.011)	(.015)	(.017)
(Copula)		-.432			.323	
		[-.528,-.341]			[.165,.427]	
No. Observations		4480			5030	
			Spain			
Change in Selection Bias	.072	.051	.021	.134	.076	-.021
(Participation eqn: SPUnem)	-.017	-.013	-.013	.048	.022	.013
	(.021)	(.018)	(.017)	(.009)	(.016)	(.021)
(Copula)		-.674			-.462	
		[-.733,-.539]			[-.563,-.265]	
No. Observations		11304			12823	
			France			
Change in Selection Bias	.013	.006	-.005	-.042	-.022	-.007
(Participation eqn: SPUnem)	-.023	-.016	-.003	.026	.016	.009
	(.021)	(.018)	(.017)	(.009)	(.012)	(.016)
(Copula)		-.272			.163	
		[-.312,-.177]			[-.017,.212]	
No. Observations		9060			9629	

Source: EU-SILC and authors' calculations. Note: Robust standard errors in parenthesis, confidence intervals in square brackets. Other covariates in the Participation eqn. are described in the main text. Confidence intervals of the estimate of the Frank copula parameter based on subsampling (subsample size 1000, 500 replications). Matlab code at: `https://drive.google.com/file/d/0B13ohL0_ULTDaDE2N0d1ZnEzZ1U/view`

5.1 Results

For brevity, we present evidence below only for Portugal and Spain, as representatives of Southern EU, and for France as a representative of Rest of EU. Table 4 presents a summary of the main results obtained for these countries for three relevant quantiles at the bottom, centre and upper part of the wage distribution: $\tau = 0.2$, 0.5, and 0.8. For each country, gender, and chosen quantile, we show three results concerning: (a) the estimated coefficient on the dummy variable of whether the spouse lost the job in the previous period (denoted as SPunem), (b) the estimated Frank copula, and (c) the estimated change in the selection bias (RG-PG) between 2007 and 2012.

As can be observed, the estimated coefficient on SPunem for men is often negative and statistically non-significant in all three countries for all the chosen quantiles. By contrast, the corresponding coefficient for women is often positive and highly significant for the 20% quantile, particularly in Spain and Portugal, but insignificant for the other two quantiles. In line with the evidence presented by Bredtmann et al. (2014), this is consistent with the conjectured added-worker effect for less-educated married women in these two countries. Regarding the estimated Frank copula, it is negative for males in all countries, and larger in magnitude for Spain. This means that individuals with higher wages (higher U) tend to participate more (lower V); in other words, there is positive selection due to the increasing proportion of male individuals in the lower part of the distribution who become non-employed. Contrariwise, the copula for women is positive in Portugal and France, though much lower in the latter; thus, this is an indication of negative female selection because less-skilled women entering the labour market found jobs and female employment increased. However, this is not the case in Spain, where the estimated copula for women is negative (as for men) pointing out that, in contrast to Portugal, the added-worker effect was not matched by new jobs but by a collapse in labour demand for both genders.

These findings explain the results reported in the last row of each panel where the selection bias for men increases over the sample period in all three countries, mainly at the bottom of the wage distribution. As for women, the selection bias decreases in Portugal and to a lesser extent in France, but it increases in Spain, in line with the results presented in subsection 2.2.3.

Estimates for Greece and Italy are reported in Table A.8 in the Online Appendix D. In a nutshell, the results show that, on the one hand, Greece behaves similarly to Spain (LD is binding for women) and, on the other, that Italy fares like Portugal (female LS binds).

Hence, overall we take these results as being fairly supportive of the previous evidence based on median imputation methods.

6 Interpreting the findings

In view of the previous empirical evidence using both selection-correction methods , we complete our analysis by providing an overview of our findings compare to the theoretical scenarios laid out in Section 2.2 about the main potential drivers of gender wage gaps in the EU during the Great Recession. Relying on the results in Tables 1 and 4 and Figures 2 and 3, we summarize our interpretation of the evidence in Table

5.

Table 5: Summary of Findings over the Great Recession

	Consistent Hypotheses			
	I_m	I_f	II_{fe}	II_{fu}
Southern Europe:				
Greece	✓			✓
Italy	✓		✓	
Spain	✓			✓
Portugal	✓		✓	
Rest of Europe:				
Austria			✓	
Belgium			✓	
Ireland	✓		✓	
United Kingdom	✓			✓
Netherlands	✓			✓
France	✓		✓	
Finland	✓			✓
Denmark	✓			✓
Norway	✓		✓	

Notes: Hypothesis I_m (I_f): higher job destruction rate among low-skilled *male* (*female*) workers. Hypothesis II_{fe} : added-worker effect with female employment *gains*. Hypothesis II_{fu} : added-worker effect with female employment *losses*.

The first conclusion is that neither the male (Hypothesis $\mathbf{I_m}$) nor the female version (Hypothesis $\mathbf{I_f}$) of Hypothesis I (destruction of low-skilled jobs) hold *per se* for any country in our sample. This is because our evidence points to sizeable changes in both male and female selection simultaneously. Hence, one can infer that the estimated selection biases and the observed employment changes in EU countries should be rationalized by a combination of some of the hypotheses listed in Section 2.2.

Within Southern EU, the patterns for Italy and Portugal fit the implications of the combination of Hypotheses $\mathbf{I_m} + \mathrm{II}_{fe}$ (added-worker effect with large male employment losses and female employment gains or small losses), which jointly leads a substantial reduction (resp. increase) in female (resp. male) selection, so that $\Delta RG > \Delta PG$. By contrast, the Greek and Spanish patterns seem to be better explained by Hypotheses $\mathbf{I_m}+\mathbf{II_{fu}}$, with a collapse in both male and female (unskilled) employment rates despite higher female LFP, and a simultaneous rise in selection biases for both genders.

Among Rest of EU countries, where employment losses have been much more muted than in Southern EU, except in Denmark and Ireland–, we find two distinct patterns. On the one hand, Austria, Belgium, France and Norway represent good illustrations of Hypothesis II_{fe} on its own (a significant decline in female selection and a mild increase in male selection, together female employment gains). On the other hand, the findings for Finland, The Netherlands, UK, and particularly Denmark and Ireland, are better rationalized by Hypothesis $\mathrm{I}_{\mathbf{m}}+\mathrm{II}_{\mathbf{fu}}$, since job destruction has affected both gender groups at the bottom of the skill distribution.

Overall, the increase in male selection emerges as a robust finding in most countries, although it has been much more pronounced in Southern EU than in Rest of EU. As for women, depending on whether LD or LS shifts dominate, we find cases where these changes have led to a larger or smaller reduction in PG than in RG. Furthermore, among the Southern EU countries most badly hit by the crisis, it seems that in those economies where female LFP was higher before the recession (Italy and Portugal), the female selection bias has declined, while the opposite has happened in those where female LFP was lower and had more dualized labour markets (Greece and Spain).

7 Conclusions

This paper has analyzed whether conventional patterns of selection of workers into EU labour markets have changed as a result of the large shifts in labour demand and labour supply produced by the Great Recession. Based on a large body of empirical evidence, it has been traditionally assumed that, because of their high LFP rates, there were no relevant differences between the observed and potential male wage distributions prior to the crisis. In contrast, due to their lower participation rates (particularly in southern Europe), women were favourably selected. Our working hypothesis is that, if the large job losses experienced during the crisis have mainly affected unskilled male-dominated sectors, then male selection may have become positive. Moreover, if non-participating women have increased their participation rates due to an added-worker effect, then female selection may have become less positive. The overall impact on the female bias is a priori ambiguous, as adverse labour demand shifts could offset the rise in female labour supply, in which case female selection changes would have been more muted or even become more positive.

Using alternative imputation techniques for the wages of non-participating individuals in EU-SILC datasets for a large group of EU countries, as well as quantile wage regressions corrected for selection biases, our findings support the conjecture that male selection has become positive. This has been especially the case in some Southern EU economies, where large male job losses have taken place in response to the bursting of real estate bubbles and the decline of other low-productivity industries. With regard to female selection, we find mixed results: while there are cases where, in line with the added-worker effect, female selection has weakened significantly (Austria, Belgium, France, Italy and Portugal), in other instances (most notably Greece and Spain) it has become stronger because widespread job destruction has also led to substantial reductions in female employment rates.

Our results highlight the importance of correcting for male selection in computing gender wage gaps. For example, in Spain, our measure of the PG implies a decline of 2.1 pp. Had we ignored male selection and only corrected for female selection, as is traditionally done, the measured PG would have *increased* by 6.7 pp. Hence, future work measuring gender gaps might require corrections for the two gender groups.

We conjecture that, once the Great Recession is over and employment growth picks up, it is likely that the increase in male selection will remain relevant. This is so since those less-skilled men who lost their jobs during the crisis (mostly concentrated in construction and other low-value-added industries) are likely to become long-term unemployed and hence non-employable. Likewise, the decrease in female selection is likely to be long-lasting since increasing female LFP seems to be a persistent trend at both ends of skills distribution, in line with the job polarization phenomenon documented by Autor and Dorn (2013) for the US and Goos et al. (2009) for some EU countries. That is, the selection patterns that we have identified during the last recession may well result in permanent changes to wage gaps.

References

Arellano, Manuel and Stéphane Bonhomme, "Quantile selection models with an application to understanding changes in wage inequality," *Econometrica*, 2017, *85* (1), 1–28.

Autor, David and David Dorn, "The growth of low-skill service jobs and the polarization of the US labor market," *The American Economic Review*, 2013, *103* (5), 1553–1597.

Bentolila, Samuel and Andrea Ichino, "Unemployment and consumption near and far away from the Mediterranean," *Journal of Population Economics*, 2008, *21* (2), 255–280.

Blau, Francine D, Marianne A Ferber, and Anne E Winkler, *The economics of women, men, and work*, Pearson; 7 edition, 2013.

Bloomfield, Peter and William Steiger, *Least absolute deviations: Theory, applications and algorithms*, Vol. 6, Springer Science & Business Media, 2012.

Bowlus, Audra J, "Matching workers and jobs: Cyclical fluctuations in match quality," *Journal of Labor Economics*, 1995, pp. 335–350.

Bredtmann, Julia, Sebastian Otten, and Christian Rulff, "Husband's Unemployment and Wife's Labor Supply–The Added Worker Effect across Europe," *Ruhr Economic Paper*, 2014, (484).

Christofides, Louis N, Alexandros Polycarpou, and Konstantinos Vrachimis, "Gender wage gaps, 'sticky floors' and 'glass ceilings' in Europe," *Labour Economics*, 2013, *21*, 86–102.

Dolado, Juan J., "EU Dual Labour Markets: Consequences and Potential Reforms," in R. Blundell et al., ed., *Economics without Borders. Economic Research for European Policy Challenges*, Cambridge: Cambridge University Press, 2016.

Engel, Melissa and Sandra Schaffner, "How to Use the EU-SILC Panel to Analyse Monthly and Hourly Wages," *Ruhr Economic Paper*, 2012, (390).

Gayle, George-Levi and Limor Golan, "Estimating a Dynamic Adverse-Selection Model: Labour-Force Experience and the Changing Gender Earnings Gap 1968–1997," *The Review of Economic Studies*, 2012, *79* (1), 227–267.

Goldin, Claudia, "A grand gender convergence: Its last chapter," *The American Economic Review*, 2014, *104* (4), 1091–1119.

Goos, Maarten, Alan Manning, and Anna Salomons, "Job polarization in Europe," *The American Economic Review*, 2009, pp. 58–63.

Gronau, Reuben, "Wage Comparisons–A Selectivity Bias," *The Journal of Political Economy*, 1974, pp. 1119–1143.

Heckman, James J, "Sample Selection Bias as a Specification Error," *Econometrica*, 1979, *47* (1), 153–161.

Jenkins, S. P., A. Brandolini, J. Micklewright, and B. (Eds.) Nolan, *The great recession and the distribution of household income*, Oxford University Press, 2012.

Johnson, Norman L, Samuel Kotz, and Narayanaswamy Balakrishnan, *Continuous univariate distributions, vol. 1-2*, New York: John Wiley & Sons, 1994.

Johnson, William, Yuichi Kitamura, and Derek Neal, "Evaluating a simple method for estimating black-white gaps in median wages," *American Economic Review*, 2000, pp. 339–343.

Keane, Michael, Robert Moffitt, and David Runkle, "Real wages over the business cycle: Estimating the impact of heterogeneity with micro data," *The Journal of Political Economy*, 1988, pp. 1232–1266.

Mulligan, Casey B and Yona Rubinstein, "Selection, investment, and women's relative wages over time," *The Quarterly Journal of Economics*, 2008, *123* (3), 1061–1110.

Neal, Derek, "The Measured Black-White Wage Gap among Women Is Too Small," *Journal of Political Economy*, 2004, *112* (S1), S1–S28.

OECD, "Employment Outlook 2014," *Paris: Organisation for Economic Co-operation and Development*, 2014.

Olivetti, Claudia and Barbara Petrongolo, "Unequal pay or unequal employment? A cross-country analysis of gender gaps," *Journal of Labor Economics*, 2008, *26* (4), 621–654.

__ **and** __ , "The Evolution of the Gender Gap in Industrialized Countries," *Annual Review of Economics*, 2016, *8*, 405–434.

A Deriving Hourly Wages

The main challenge in deriving hourly wages is to combine annual income (PY010) and monthly economic status information (PL210A-PL210L up to 2009 and PL211A-PL211L onwards) for the previous calendar year with the number of hours usually worked per week (PL060) at the date of the interview.

To do this we combine the longitudinal files from the period 2005-2013 and use the imputed annual hours of work

$$hours_{annual} = months_{annual} \times 4.345 \times hours_{week}$$

to calculate hourly wages. The following set of rules are used sequentially to impute missing annual hours of work during the previous calendar year:

1. For those workers who have only one employment spell (with no changes in full-time/part-time status), we use the number of months of this spell and the number of hours from the previous survey.

2. For those workers who have only one employment spell (with no changes in full-time/part-time status), we use the number of months of this spell and the number of hours declared at the date of the interview if the person hasn't changed job since last year (PL160).

In the case of United Kingdom, we only use the number of hours at the date of the interview since the income reference period coincides with the year of the interview.

3. For those workers who have only one employment spell (with no changes in full-time/part-time status), we use the number of months of this spell and approximate the number of hours by the year- gender- full-time/part-time status- specific mean.

4. For those workers who have multiple employment spells, we use the number of months of each spell and the number of hours for each spell approximated by the year- gender- full-time/part-time status- specific mean.