I Z A

IZA DP No. 5932

# The Trend over Time of the Gender Wage Gap in Italy

Chiara Mussida
Matteo Picchio

August 2011

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

# The Trend over Time of the Gender Wage Gap in Italy

**Chiara Mussida**
*Università Cattolica del Sacro Cuore*

**Matteo Picchio**
*Tilburg University, CentER, ReflecT
and IZA*

# ABSTRACT

# The Trend over Time of the Gender Wage Gap in Italy[*]

We analyse gender wage inequalities in Italy in the mid-1990s and in the mid-2000s. In this period important labour market developments occurred: institutional changes have loosened the use of flexible and atypical contracts; the female employment rates and educational levels have substantially increased. We identify the time trends of different components of the gender wage gap by estimating wage distributions in the presence of covariates and sample selection and by counterfactual microsimulations. We find that women swam against the tide: whilst the trend in female qualifications slightly reduced the gender wage gap, the gender relative trends in the wage structure significantly increased it.

Corresponding author:

Matteo Picchio
Department of Economics
Tilburg University
PO Box 90153
5000 LE Tilburg
The Netherlands
E-mail: m.picchio@uvt.nl

# 1 Introduction

During the decade that goes from the mid-1990s to the mid-2000s and in a period of quite stable and low economic growth,[1] the Italian labour market experienced some important developments mostly in terms of legislation, but also in terms of workforce participation and composition. Up to the second half of the 1990s, the Italian standard work arrangement had traditionally been full-time, open-ended, and characterized by one of the strictest employment protection legislation, mostly against dismissals, in the OECD area (Lazear, 1990; Kugler and Pica, 2008). Since 1997, Italy has undertaken major steps towards "flexibility" by way of labour market reforms that have liberalized the use of flexible and atypical contracts. Law No. 196/1997 ("Treu Package"),[2] Legislative Decree No. 368/2001, Law No. 30/2003 ("Biagi's Law"), and the subsequent Legislative Decree No. 276/2003 have indeed legitimized temporary work agencies and fixed-term contracts under general conditions and given rise to the expansion of flexible and atypical work arrangements, like part-time jobs, seasonal jobs, youth work-training, apprenticeship contracts, jobs on call, time-sharing, staff-leasing, and work on project.

At the same time, the Italian labour market experienced two further developments which might be partially associated with the aforementioned reforms: a falling gender gap in employment rates and a falling gender gap in the levels of education. With regard to the former, the female employment rates increased significantly from 1995 to 2005, both in absolute value and relatively to men. Considering the population in the age range 25–54, the female employment rate went from 46.6% in 1995 to 57.9% in 2005, whereas the male employment rate remained stable, going from 84.5% in 1995 to 86.6% in 2005. As regards converging levels of education, the fraction of women with a post secondary or tertiary education has increased more than the corresponding share of men. Among the population in the age range 25–64, the share of women with at least a post secondary education went from 33.6% in 1995 to 50.6% in 2005, whilst the corresponding share of men went from 39.1% in 1995 to 50.2%.[3]

These developments of the Italian labour market might have asymmetrically hit the remuneration of men and women. One explanation of women suffering wage penalties is that men and women have different preferences towards work: given the Italian male breadwinner system, women might be more willing than men to exchange wages for flexible features of the work arrangement in order to reconcile family responsibilities with employment. Women might therefore be more altruistic and less career-oriented than males. When the institutional setting becomes more oriented and favourable to flexible form of work arrangements like part-time jobs, on call jobs, and job sharing, the trade-off between flexibility and wages might be less costly and, thereby, a reduction in the gender wage gap might be observed. There is however a debate about whether atypical contracts spur the development

---

[1]In Italy the average GDP annual growth rate was around 1.5% in the period 1995–2005 (Eurostat).

[2]For an assessment of the Treu reform, see Sciulli (2006a,b) and Schindler (2009)

[3]These figures are available in Internet at http://epp.eurostat.ec.europa.eu.

of a secondary labour market in which workers get trapped in a cycle of unstable and low paid jobs. If it is so and women are more likely to get into atypical jobs, observed female wage penalties might increase. Moreover, if flexible forms of job contracts make it easier for women to reconcile family responsibilities with working career, the female labour force participation increases and women with low earnings potential, who would not have participated otherwise, get into the workforce and start featuring in the observed wage distribution. This generates an increase in the observed gender wage gap over time, a statistical artefact due to a time-variation in the composition of the sample of working women. With regard to the relative change in the level of education, when women become more and more educated with respect to men, we would expect a decrease in the observed gender wage gap. Nevertheless, if the rewards of education has decreased over time, the gender wage gap might be only slightly affected by women becoming relatively more educated. As pointed out by Blau and Khan (1997, 1999), the overall wage structure, i.e. the array of prices of labour market skills, can indeed have an important effect on gender differentials.

The main question addressed by this article is whether and how gender inequalities in wages have changed in Italy in the face of the important labour market developments of the last years. If one simply compares the raw gender wage gap in the mid-1990s to the one in the mid-2000s, it is noticed that the relative remuneration of men and women did not change. This preliminary evidence from raw data might however be just a statistical artefact, with the risk of generating misleading conclusions. The total gender wage gap can indeed be seen as made up of a component due to gender differences in the distribution of individual characteristics and the component due to gender differences in the rewards of the same characteristics. Understanding the contribution of each component to the evolution over time of the gender wage gap is policy relevant in order to design policies that are effective in pursuing the gender equality target.

The literature on gender inequalities in the labour market has recently paid attention to developing methodological tools and increasing thereby the available set of instruments to deepen the understanding of the gender wage gap issue (Blundell et al., 2007; Olivetti and Petrongolo, 2008; Albrecht et al., 2009; Picchio and Mussida, 2011). Despite the increased recognition of the importance of gender pay differentials, research on the gender wage gap in Italy has been scarce. Some effort has been nonetheless made in recent years by Rustichelli (2005), Addabbo et al. (2007), and Addabbo and Favaro (2007). However, to our knowledge, no country-specific study on the trends over time of the gender wage gap has been carried out for Italy and for other European countries with similar labour market trends. From the empirical point of view, this paper contributes to fill this gap in the literature.

In order to have a meaningful comparison over time of the evolution of the gender wage gap, we carry out an econometric analysis that is able to: i) net out the effect of changes in the distribution of individual characteristics, as they could erroneously point to a reduction in gender inequalities, whereas it is just an increase of the female endowments of characteristics that are highly rewarded; ii) take into account changes in the nonrandom selection

into the workforce, as they might wrongly point to an increase in wage inequality, whilst it is simply that women with low earnings potential are more likely to feature in the observed wage distribution than in the past.

The econometric analysis is based on the technique proposed by Picchio and Mussida (2011) to estimate wage distributions in the presence of covariates and sample selection and on simulation algorithms to derive counterfactual distributions. The contribution of this paper from the methodological viewpoint consists in extending the simulation procedure in Picchio and Mussida (2011) to decompose the evolution over time of the gender wage gap into the component due to gender trends in the distribution of individual characteristics and into the component due to gender trends in the rewards of characteristics.

The empirical analysis is carried out on the basis of two longitudinal datasets. The first one is the European Community Household Panel (ECHP), which covers the period before the labour market reforms, i.e. from 1994 until 1997. The second one is the European Union Statistics on Income and Living Conditions (SILC). It covers the period after the reforms, which goes from 2004 until 2007. We find that women have been swimming against the tide. Whilst the trend in women's relative qualifications slightly reduced the gender wage gap, the trends in wage structures of men and women substantially increased it.

The paper proceeds as follows. Section 2 describes the data and the samples. Section 3 presents the methodology to estimate wage distributions in the presence of covariates and sample selection. Section 4 describe how we decompose the gender wage gap in different components and analyse their time-changes across all quantiles of the distribution. We comment the findings and discuss policy implications. Section 5 concludes.

## 2   Data and Sample

The empirical analysis exploits data from two longitudinal surveys. The first one is the ECHP, which covers the period before the institutional changes, i.e. from 1994 to 1997. The second one is the SILC, which instead covers the period after the labour market reforms, i.e. from 2004 to 2007. We select data for Italy, where the surveys were conducted by the National Institute of Statistics (ISTAT) under the coordination of Eurostat.[4]

Since information on wages is retrospectively collected in both panels, we lose the last years (1997 and 2007). We exclude from our sample individuals younger than 25 years and older than 55 years to avoid to get mixed with education enrolment and early retirement issues. We drop individuals who were in the army, self-employed, inactive, or with missing values for some variables used in the econometric analysis. Finally, we exclude from the sample individuals lying in the first or last percentile of the wage or working hours distributions.

---

[4]More information about the ECHP is available in Internet at http://epp.eurostat.ec.europa.eu/portal/page/-portal/eurostat/home. See Eurostat (2004) or/and go to http://epp.eurostat.ec.europa.eu/portal/page/portal/-microdata/eu_silc for details about the SILC data.

Considering both the employed and the non-employed, 10,863 female observations and 8,064 male observations remain over the period 1994–1996. The corresponding numbers for the period 2004–2006 are 12,635 for women and 10,118 for men. The numbers of wage observations are 4,316 (6,722) for women and 6,663 (8,918) for men in 1994–1996 (2004–2006). The definitions of employment and non-employment are based on self-defined economic status and therefore do not match the ILO definition. The female employment rate rose from 39.7% in the mid-1990s to 53.2% in the mid-2000s, while the male employment rate went from 82.6% to 88.1%. Hence, the gender gap in employment rates decreased over time because of a substantial increase in the female employment rate.

The wage variable is the net hourly wage of employees' main occupation,[5] deflated to 2004 constant prices.[6] The raw statistics in Tables 1 and 2 report the gender wage gap in the mid-1990s and in the mid-2000s. The raw relative remuneration of men and women remained stable with a slight decline over time: men were earning on average 6% more than women in the mid-1990s and 5% more than women in the mid-2000. The pay gap is not constant over the wage distributions. It is well perceived at the bottom and in the middle of the wage support, but it disappears at the top of the distribution. In the mid-1990s (mid-2000s), it went from 20% (12%)[7] at the bottom of the distribution to 6% (5%) at the median and virtually nil at the top of the wage distribution. Figure 1 plots the kernel estimate of the wage density for men and women in both periods. The empirical probability density functions of the net hourly wage are positively skewed and display excess kurtosis for both periods and genders.[8]

Table 1: Summary Statistics of Net Hourly Wage by Gender in Italy, 1994–1996 and 2004–2006

| Net hourly wage (€) | Men | | Women | | Overall | |
|---|---|---|---|---|---|---|
| | Mean | Std.Dev. | Mean | Std.Dev. | Mean | Std.Dev. |
| 1994–1996 | 7.943 | 2.800 | 7.604 | 3.005 | 7.810 | 2.887 |
| 2004–2006 | 9.566 | 3.646 | 9.208 | 3.778 | 9.412 | 3.707 |

*Notes*: Wages are in constant prices (2004 prices). They are deflated by using the Consumer Price Index (CPI), gathered by ISTAT.

Table 3 displays summary statistics of the covariates used in the econometric analysis.

---

[5]The net hourly wage is computed starting from the employees' yearly net cash income, variable PI111 (PY010N), and using the number of months at work, obtained from the calendar activities variables PC001-PC013 (PL210A-PL210L), and the number of hours usually worked per week, PE005 (PL060) in the 1994–1996 (2004–2006) panel. Considering that on average there are 4.345 weeks in a month, the hourly wage in the mid-1990s (mid-2000s) is computed as follows: $w = \text{PI111}/(\# \text{ of months at work} \times \text{PE005} \times 4.345)$ ($w = \text{PY010N}/(\# \text{ of months at work} \times \text{PL060} \times 4.345)$).

[6]The deflator is the Consumer Price Index (CPI), gathered by ISTAT.

[7]$20\% = [\exp(0.182) - 1] * 100$ and $12\% = [\exp(0.113) - 1] * 100$.

[8]The skewness and the kurtosis of the wage distributions are, respectively, around 1.4 and 5.5 in both time periods and for both genders.

Table 2: Raw Gender Wage Gap in Italy, 1994–1996 and 2004–2006 (gap in log points)

|  | 1994–1996 | 2004–2006 |
|---|---|---|
| Proportion of women | 0.393 | 0.430 |
| Mean | 0.059 | 0.048 |
| 5th percentile | 0.182 | 0.113 |
| 10th percentile | 0.107 | 0.086 |
| 25th percentile | 0.081 | 0.070 |
| 50th percentile | 0.059 | 0.048 |
| 75th percentile | 0.036 | 0.031 |
| 90th percentile | -0.008 | 0.000 |
| 95th percentile | -0.011 | -0.002 |

*Notes*: The raw wage gap is measured as the difference between the log male hourly wage and the log female hourly wage.
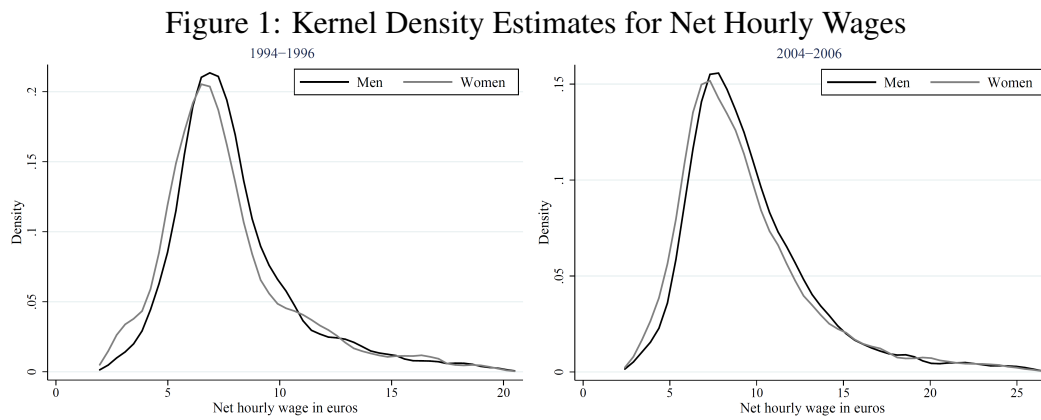
Figure 1: Kernel Density Estimates for Net Hourly Wages

## Table 3: Summary Statistics of the Covariates by Gender and Time Period

| | 1994–1996 | | | | 2004–2006 | | | |
| | Male | | Female | | Male | | Female | |
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
|---|---|---|---|---|---|---|---|---|
| | *All sample* | | | | | | | |
| Employed | .826 | .379 | .397 | .489 | .881 | .323 | .532 | .499 |
| Age (years) | 38.885 | 8.692 | 39.515 | 8.724 | 39.937 | 8.467 | 4.564 | 8.462 |
| Education§ | | | | | | | | |
|   None, elementary, or lower secon. | .492 | .500 | .566 | .496 | .436 | .496 | .437 | .496 |
|   Upper secondary | .404 | .491 | .354 | .478 | .389 | .487 | .363 | .481 |
|   Post secondary or tertiary | .108 | .310 | .092 | .289 | .175 | .380 | .200 | .400 |
| Area of residence | | | | | | | | |
|   North-West | .198 | .398 | .187 | .390 | .221 | .415 | .222 | .416 |
|   North-East | .166 | .372 | .164 | .370 | .243 | .429 | .222 | .416 |
|   Centre | .248 | .432 | .240 | .427 | .232 | .422 | .234 | .424 |
|   South | .382 | .486 | .403 | .491 | .304 | .460 | .321 | .467 |
| Bad health† | .300 | .458 | .357 | .479 | .251 | .434 | .301 | .459 |
| Married | .714 | .452 | .801 | .400 | .652 | .476 | .728 | .445 |
| # household members | 2.884 | 1.182 | 2.897 | 1.154 | 3.288 | 1.213 | 3.349 | 1.146 |
| Presence of child < 12 years | .247 | .431 | .253 | .435 | .345 | .476 | .350 | .477 |
| Potential Experience (years) | 17.512 | 10.761 | 12.325 | 11.950 | 21.686 | 10.077 | 17.549 | 12.161 |
| 1994 or 2004 | .328 | .470 | .330 | .470 | .176 | .381 | .174 | .379 |
| 1995 or 2005 | .346 | .476 | .344 | .475 | .339 | .473 | .339 | .474 |
| 1996 or 2006 | .325 | .469 | .326 | .469 | .485 | .500 | .487 | .500 |
| Observations | 8,064 | | 10,863 | | 10,118 | | 12,635 | |
| | *Employees* | | | | | | | |
| Age (years) | 39.596 | 8.362 | 38.379 | 8.159 | 4.161 | 8.339 | 39.967 | 8.227 |
| Education§ | | | | | | | | |
|   None, elementary, or lower secon. | .461 | .499 | .339 | .473 | .416 | .493 | .283 | .451 |
|   Upper secondary | .428 | .495 | .523 | .500 | .405 | .491 | .432 | .495 |
|   Post secondary or tertiary | .109 | .312 | .139 | .346 | .179 | .383 | .285 | .452 |
| Area of residence | | | | | | | | |
|   North-West | .222 | .416 | .265 | .441 | .230 | .421 | .267 | .442 |
|   North-East | .188 | .391 | .223 | .416 | .258 | .438 | .282 | .450 |
|   Centre | .261 | .439 | .250 | .433 | .239 | .426 | .250 | .433 |
|   South | .325 | .468 | .258 | .438 | .273 | .445 | .202 | .401 |
| Bad health† | .289 | .453 | .312 | .463 | .241 | .428 | .272 | .445 |
| Married | .778 | .416 | .723 | .448 | .688 | .463 | .643 | .479 |
| Potential Experience (years) | 18.772 | 9.881 | 16.269 | 9.255 | 22.228 | 9.593 | 20.755 | 9.496 |
| 1994 or 2004 | .337 | .473 | .335 | .472 | .174 | .379 | .173 | .378 |
| 1995 or 2005 | .344 | .475 | .345 | .475 | .341 | .474 | .336 | .472 |
| 1996 or 2006 | .318 | .466 | .320 | .466 | .485 | .500 | .491 | .500 |
| Part-time worker | .025 | .156 | .149 | .356 | .043 | .203 | .302 | .459 |
| Temporary contract | .049 | .215 | .056 | .230 | .072 | .259 | .116 | .321 |
| Occupation‡ | | | | | | | | |
|   White-collar high-skilled worker | .200 | .400 | .319 | .466 | .326 | .469 | .442 | .497 |
|   White-collar low-skilled worker | .309 | .462 | .426 | .495 | .205 | .403 | .342 | .474 |
|   Blue-collar worker | .492 | .500 | .255 | .436 | .470 | .499 | .216 | .411 |
| Observations | 6,663 | | 4,316 | | 8,918 | | 6,722 | |

§ Educational indicators refer to the highest and successfully completed educational attainment of a person. We used the ISCED-97 educational classification to build these indicators.

† "Bad health" is a dummy indicator based on self-perceived health. It is equal to one if the individual declares that her health is in a fair, rather bad, or bad conditions. It is equal to zero, if the answer is either good or rather good condition.

‡ The occupational indicators are built on the basis of the International Standard Classification of Occupation (ISCO-88). Taking the one-digit ISCO-88 categories, the white-collar high skilled indicator corresponds to categories 1 up to 3; the white-collar low skilled indicator corresponds to categories 4 and 5; the blue-collar indicator corresponds to categories from 6 to 9.

They are calculated by gender and time period on the overall sample (top panel) and on the subsample of employees (bottom panel). Men and women are on average 39 years old in the mid-1990s and somewhat older (40 years of age) in the the mid-2000s. Educational dummies are defined according to UNESCO's International Standard Classification of Education (ISCED) and distinguish between education completed at three different stages: lower secondary, upper secondary, and post-secondary or tertiary education. In both periods working women are more educated than working men. This gender difference becomes even stronger in the mid-2000s, when 28.5% of working women have a post-secondary education attainment against 17.9% of working men. This is reflected in the proportion of people with a lower educational attainment: 28.3% of working women have no more than lower secondary levels against 41.6% of working men. Gender gaps in education are however much lower if one considers both employed and non-employed people: in the mid-2000s, 20% of women have a post-secondary or tertiary educational attainment, against 17.5% of men. This shows that women are more educated than men and that women with low education have a lower propensity to be at work than low educated men.

Four indicators capture the geographical area of residence by splitting Italy in North-West, North-East, Centre, and South. Around one half of the employees live in the North of Italy, while the others are equally distributed in the Centre and South. The indicator of self-perceived health captures the effect of health status (perceived or subjective) on wages and propensity to work and less than one third of the sample declares not to be in good health. Marital status and household variables like the number of household members and the presence of child (younger than 12) are also considered. The percentage of married people decreased over time going from 71% (80%) in the mid-1990s to 65% (73%) in the mid-2000s for (wo)men, while the average number of household components increased. A measure of potential experience, computed as difference between age and the age when the first regular job started, is also introduced as likely to affect earnings and individual job opportunities.

Finally, we consider a set of variables aimed at capturing job heterogeneity, such as indicators for part-time jobs, for temporary jobs, and for type of occupation. In both periods working women are more likely to hold temporary or part-time jobs than working men. This evidence is stronger in the mid-2000s: 30.2% of female employees have a part-time job, against 4.3% of male employees; 11.6% of working women hold a temporary contract, against 7.2% of working men. Using the International Standard Classification of Occupations (ISCO-88), we define indicators aimed at capturing tasks heterogeneity. Men are more likely to be blue collar workers in both periods. In 2004–2006 almost 50% of working men are blue collar workers against 21.6% of working women. The share of female blue collars decreased over time, while the share of high skilled white collar women increased from 31.9% in the mid-1990s to 44.2% in the mid-2000s.

# 3 Estimation of Wage Distributions in the Presence of Covariates and Sample Selection

As said in the Introduction, this study is aimed at understanding the evolution of the gender wage gap over time and disentangle the contribution given by the time trend of the different components of the gender wage gap. In other words, in order to have a meaningful comparison over time of the evolution of the gender wage gap, we require an econometric approach that is able to net out the effect of changes over time in the distribution of individual characteristics and take into account changes over time in the nonrandom selection into the workforce. Moreover, as labour markets could be internally very differentiated, we are not just interested in looking at the evolution of the average gender wage gap. We are rather interested in its evolution across all the wage distribution. The starting point is therefore to design an econometric model that allows us to flexibly estimate wage distributions in the presence of covariates and sample selection.

Different non/semi-parametric estimators have been proposed to model wage distributions in the presence of covariates. DiNardo et al. (1996) adapted kernel density estimator to the case in which sample weights are attached to each observation. Counterfactual densities are estimated by "re-weighting" functions, which depend on different covariates distribution between populations. Fortin and Lemieux (1998) divided the wage support in small intervals and estimated the probability of being in each wage interval using an ordered response (probit) model. Once the model is estimated, counterfactual distributions can be predicted by playing with covariates and/or estimated coefficients of the ordered response model. Donald et al. (2000) considered wage distributions as if they were duration distributions and applied hazard function (HF) based estimation techniques. Picchio and Mussida (2011) adapted the HF estimator proposed by Donald et al. (2000) to a panel data framework with sample selection correction. Finally, probably the most common method to estimate distributions in the presence of covariates is quantile regression (e.g. Buchinsky, 1994), followed by Machado and Mata's (2005) (MM) procedure to carry out counterfactual comparisons of densities. Albrecht et al. (2009) adapted the MM decomposition to account for sample selection.

The approach used here for modelling wage distributions in the presence of covariates and sample selection follows Picchio and Mussida (2011). It takes advantage of panel data: multiple observations per individual are exploited to identify the individual unobservables jointly determining workforce participation and wages, without relying exclusively on exclusion restrictions or strict parametric restrictions on the unobserved heterogeneity distribution.[9] This method allows us to recover the sample selection corrected wage structure – the set of parameters determining the wage distribution – by gender and by time period. In Section 4, we will exploit the estimated wage structures and the empirical distributions of individual characteristics of men and women in the mid-1990s and mid-2000s to understand

---

[9]See Picchio and Mussida (2011) for a discussion about this and other methods to estimate wage distributions in the presence of covariates and sample selection.

the evolution over time of different components of the Italian gender wage gap by way of microsimulations and counterfactual analysis.

## 3.1 Model Specification

In what follows, consider a panel data setting where $t$ is the time indicator.[10] We assume that selection into employment is described by the following random effects logit model

$$y_t = \mathbb{1}[z_t'\delta + \varepsilon + u_t > 0], \quad t = 1, \cdots, T \tag{1}$$

where $\mathbb{1}(\cdot)$ is the indicator function, $y_t$ is the scalar indicator variable denoting employment status at time $t$, $z_t$ is the $K$-vector of regressors explaining the employment probability, $\delta \in \Re^K$ is a parameter vector, $\varepsilon$ is unobserved heterogeneity, and $u_t$ is the idiosyncratic error term with logistic distribution $\Lambda$.

With regard to wages, we assume that all individual differences in the wage distribution can be characterized by observed characteristics $x_t$ and individual heterogeneity $v$. For a covariate vector $x_t$ and individual heterogeneity $v$ there is a positive random variable $W_t$ with associated probability and cumulative density functions $f(w_t|x_t, v)$ and $F(w_t|x_t, v)$. If $y_t = 1$, the wage hazard function is defined as

$$\theta(w_t|x_t, v) = \frac{f(w_t|x_t, v)}{1 - F(w_t|x_t, v)} \equiv \frac{f(w_t|x_t, v)}{S(w_t|x_t, v)}, \tag{2}$$

where $S(w_t|x_t, v)$ is the survivor function, i.e. the probability of being paid at least a wage equal to $w_t$. We specify the hazard function as

$$\theta(w_t|x_t, v) = h_t(w_t|x_t)v, \tag{3}$$

where $h_t$ is the structural wage hazard function at time $t$.[11]

While in the duration literature the hazard function is usually assumed to have a mixed proportional hazard form, here we allow the covariate effect to be different over the wage support, at different percentiles of the unconditional wage distribution. We partition the wage support into $P$ intervals and we allow the covariate effect to vary over these intervals. The variation is with respect to a baseline wage hazard function $h^0(w_t)$ that, in order to avoid too strict parametric assumptions, is assumed to be piecewise constant.[12] Formally,

---

[10]We suppress the individual indicator for the sake of clarity.

[11]In the empirical analysis there are some regressors that enter the selection equation but do not enter the wage hazard function. These exclusion restrictions are in line with those often used in the labour supply literature (e.g. Mroz, 1987): indicators for the presence of children and the number of members in the household.

[12]A piecewise constant function is constant within predefined intervals. We divided the wage support into $J = 72$ intervals $I_j = [w_{j-1}, w_j)$, where $j = 1, \ldots, J$, $w_0 < w_1 < \ldots < w_J$, $w_0 = 0$, and $w_J = \infty$. $w_1$ and $w_{J-1}$ correspond to the first and last percentiles of the wage distribution in each time period. We chose the

denote $\Omega_p = [w_L^p, w_H^p)$ the $p$th wage interval over which the covariate effect is assumed to be constant (i.e. proportional to the baseline wage hazard function), such that $\Omega_p \cap \Omega_q = 0$ for all $p \neq q$ and $\bigcup_{p=1}^{P} \Omega_p = [0, \infty)$. The structural hazard function is specified as follows

$$h_t(w_t | x_t) = h^0(w_t) \exp\left[ x_t' \beta(w_t) \right] = h^0(w_t) \exp\left[ \sum_{p=1}^{P} \mathbb{1}(w_t \in \Omega_p) x_t' \beta^p \right]. \qquad (4)$$

In order to reduce the risk of over-fitting bias and to have a reasonable trade-off between model parsimony and fit to the data at hand, we chose $P$ by minimizing the Akaike Information Criterion (AIC). It turned out that $P = 8$ for men in 1994–1996 and $P = 10$ for men in 2004–2006 and women in both time periods minimize the AIC, with dividing points approximatively at the $[100\frac{1}{P}, 100\frac{2}{P}, \dots, 100\frac{P-1}{P}]$ percentiles of the unconditional wage distribution.

If the unobserved determinants of the wage distribution and of the employment equation ($v$ and $\varepsilon$) are independent, inference on the wage distribution can be based solely on the distribution of $w | x$. However, if $v$ and $\varepsilon$ are dependent then inference has to be based on the joint distribution of $w, y | x$.[13] The joint distribution of $v$ and $\varepsilon$ is approximated by a bivariate discrete distribution (Heckman and Singer, 1984) with a fixed number of support points, which have unknown locations and probability masses. By doing so, we avoid too strict parametric assumptions. We assume that $(v, \varepsilon)$ has four probability points with probability masses defined as follows:

$$
\begin{aligned}
p_1 &\equiv \Pr(v = v_1, \varepsilon = \varepsilon_1) & p_2 &\equiv \Pr(v = v_2, \varepsilon = \varepsilon_1) \\
p_3 &\equiv \Pr(v = v_1, \varepsilon = \varepsilon_2) & p_4 &\equiv \Pr(v = v_2, \varepsilon = \varepsilon_2) = 1 - p_1 - p_2 - p_3.
\end{aligned}
$$

In this case, we need to estimate four points of support and three probability masses.[14] The probabilities associated to the mass points are specified as logistic transforms:

$$p_m = \frac{\exp(\lambda_m)}{\sum_{r=1}^{4} \exp(\lambda_r)} \quad \text{with} \quad \lambda_4 = 0.$$

width of the other 70 wage baseline segments by dividing the wage support between $w_1$ and $w_{J-1}$ in 70 equally spaced intervals. Our choice of the number of the baseline segments is somewhat arbitrary. Nevertheless, it returns a narrow segment width (0.22€ and 0.28€ for the mid-1990s and the mid-2000s respectively) and it is thereby suitable for flexibly approximating all possible wage hazard functions.

[13]In this case, $y$ is indeed related to $v$ and knowledge that $y = 1$ is informative on $v$. The dependence between $v$ and $\varepsilon$ will alter $E(v | W \geq w, x)$ and therefore the wage distribution conditional on $x$. Consider for example that, if the distribution of $v | y = 1$ is stochastically dominated by the population distribution of $v$ and this is ignored, the conditional wage hazard function $\theta(w | x, v)$ would be underestimated, i.e. every quantile of the conditional wage distribution would be overestimated.

[14]$v$ and $\varepsilon$ are independent if and only if $p_1 p_4 = p_2 p_3$ (Van den Berg et al., 1994; Van den Berg and Lindeboom, 1998). This makes easy to test for sample selection.

## 3.2 The Likelihood Function

The contribution to the likelihood function of a wage $w_t$ lying in the baseline interval $[w_{j-1}, w_j)$ is

$$
\begin{aligned}
\Pr(w_{j-1} \leq W_t < w_j | x_t, v) &= S(w_{j-1}|x_t, v) - S(w_j|x_t, v) \qquad (5)\\
&= L_{it}^w(w_t|x_t, v; \Theta^w),
\end{aligned}
$$

where $\Theta^w$ is the set of parameters to be estimated, and

$$
S(w_j|x_t, v) \equiv \exp\left[ - \int_0^{w_j} \theta(s|x_t, v) ds \right]
$$

is the wage survivor function, i.e. the probability of observing a wage at least as large as the upper limit of the $j^{\text{th}}$ segment. If the wage is larger than the $99^{\text{th}}$ percentile of the wage distribution, it is treated as a right censored observation. The corresponding contribution to the likelihood function is the probability that the wage is larger than the $99^{\text{th}}$ percentile,

$$
L_{it}^w(w_t|x_t, v; \Theta^w) = S(w_{J-1=99^{\text{th}}}|x_t, v). \qquad (6)
$$

Considering the nonrandom selection process into the workforce and the discrete distribution assumption on the unobserved heterogeneity distribution, the contribution to the likelihood function of individual $i$ is given by

$$
\mathscr{L}_i = \sum_{m=1}^{4} p_m \prod_{t=1}^{T} [\Lambda(z_{it}'\delta + \varepsilon_{im}) L_{it}^w(w_{it}|x_{it}, v_{im}; \Theta^w)]^{y_{it}} [1 - \Lambda(z_{it}'\delta + \varepsilon_{im})]^{1-y_{it}}.
$$

The log-likelihood function sums the logarithm of this expression over all the individuals (employed and nonemployed) and is estimated by maximum likelihood. The estimations are separately run for men and women and for the periods 1994–1996 and 2004–2006.

## 3.3 Estimation Results of the Wage Structures

This subsection reports and comments on the impact of the regressors on the male and female wage hazard functions and, thereby, on the wage distribution functions. The estimation results of the employment selection equation are in line with the expectations. They are reported in Table 8 and commented in Appendix A. As mentioned in Subsection 3.1, the impact of covariates on the shape of the wage density functions is flexibly modelled, so that it can be different over the wage support. In other words, analogously to quantile regression, the covariates can have different effects at different percentiles of the wage distribution. Hence, in Tables 4 and 5 we document the effects of the covariates at selected percentiles for men and women, respectively, both in the 1990s and in the 2000s.

11

Two points are worthy of mention. First, Tables 4 and 5 display selection-corrected estimation results.[15] The Log-likelihood Ratio (LR) tests reported at the end of these tables indicate that the null hypothesis of no sample selection can be rejected quite confidently for women, both in the 1990s and in the 2000s. For men, the null hypothesis of no sample selection is rejected in the 1990s, but not rejected in the 2000s. Second, the estimated coefficients inform us about the covariate impact on the wage hazard function: individual characteristics that have a negative effect on the wage hazard rate reduce the probability of getting a low wage. Individuals holding these characteristics are therefore more likely to get a higher wage with respect to the reference group.

Getting older increases the probability of getting high wages in both periods. Education and occupation play significant roles in explaining the male and female wage distributions. Higher educational attainments (such as post secondary or tertiary) and high-skilled and white-collar occupations are associated with higher wages. This is true across the overall wage distribution for both genders and time periods, with the exception of women at the top of the distribution (since the $75^{th}$ percentile) in the mid-2000s.

We find geographical differentials in the distribution of wages, characterized by relevant pay disadvantages in the Centre and especially in the South of Italy in the mid-1990s. The geographical pay gaps maintain their relevance at the bottom of the wage distribution (up to the $25^{th}$ percentile) for both genders in the mid-2000s and seem to disappear at the top of the wage distributions of both men and women.

Having a temporary job is associated with high wages especially for men in the middle and at the top of the wage distribution in the mid-1990s. Ten years later this evidence has changed. Male temporary workers are indeed paid lower wages than male permanent workers across all the wage distribution. Only at the top of the distribution, the negative effect of temporary contracts becomes smaller and not significantly different from zero. This in line with Barbieri and Cutuli (2009), who, using similar data and quantile regression, find that temporary workers earn less than permanent workers with a gap tending to zero at the top of the wage distribution. The change over time of the association between the type of contract and wages might be due to the labour market reforms. Before the reforms, temporary contracts could be used in very limited circumstances (e.g. to replace a worker in sick leave) and workers willing to take a job for a limited period of time were maybe compensated for the instability and idiosyncrasy of the close-ended contract. The labour market reforms have made temporary contracts the norm and have created a dual system in which permanent workers have kept the usual level of employment protection. The presence of a dual labour market might therefore explain the increased gap in wages between temporary and permanent workers before and after the labour market reforms.

With regard to the remaining covariates, subjective (or perceived) health status does not affect earnings so much. Marital status affects male wage distribution in both periods and married men are better paid, especially in the mid-2000s. Part-time jobs are found to be

---

[15]Parameter estimates without sample selection are not reported in the paper but available upon request.

Table 4: Coefficient Estimates of the Covariates with Correction for Selectivity at Selected Percentiles of the Wage Distribution for Men, 1994–1996 and 2004–2006

| Variables | 1994–1996 | | | | | | 2004–2006 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Percentiles | 25th | | 50th | | 75th | | 25th | | 50th | | 75th | |
| | Coef | SE | Coef | SE | Coef | SE | Coef | SE | Coef | SE | Coef | SE |
| Age | -.650 *** | .082 | -.348 *** | .089 | -.336 *** | .091 | -.752 *** | .088 | -.621 *** | .114 | -.488 *** | .096 |
| Potential experience | .122 * | .070 | -.029 | .075 | -.007 | .078 | .116 | .077 | .052 | .103 | .161 * | .085 |
| Education – Reference: None, elementary, or lower secondary | | | | | | | | | | | | |
| Upper secondary | -.660 *** | .079 | -.607 *** | .085 | -.424 *** | .075 | -.649 *** | .072 | -.408 *** | .093 | -.174 ** | .081 |
| Post secondary or tertiary | -1.051 *** | .166 | -1.604 *** | .152 | -1.520 *** | .141 | -1.148 *** | .132 | -.934 *** | .140 | -.656 *** | .115 |
| Occupation – Reference: Blue collar workers | | | | | | | | | | | | |
| High skilled & white collar | -1.151 *** | .125 | -1.293 *** | .117 | -.793 *** | .100 | -1.312 *** | .093 | -1.373 *** | .102 | -.683 *** | .087 |
| Low skilled & white collar | -.878 *** | .084 | -.799 *** | .090 | -.129 * | .075 | -.667 *** | .080 | -.885 *** | .106 | .026 | .083 |
| Area of residence – Reference: North-West | | | | | | | | | | | | |
| North-East | -.097 | .098 | .129 | .112 | -.072 | .090 | .152 * | .090 | .063 | .106 | .011 | .090 |
| Centre | .175 * | .096 | .449 *** | .102 | .282 | .079 | .662 *** | .090 | .066 | .115 | .072 | .091 |
| South | .594 *** | .093 | .555 *** | .102 | .217 ** | .087 | .839 *** | .096 | .294 ** | .115 | -.003 | .095 |
| Bad health | .044 | .077 | .027 | .084 | .082 | .083 | -.037 | .078 | .202 ** | .092 | .058 | .080 |
| Married | -.685 *** | .082 | -.283 *** | .106 | -.008 | .095 | -.770 *** | .070 | -.550 *** | .094 | -.298 *** | .078 |
| Temporary contract | .361 ** | .161 | -.465 * | .271 | -.760 *** | .283 | .802 *** | .116 | .553 *** | .186 | .220 | .181 |
| Part-time | -1.027 *** | .338 | -1.584 *** | .355 | -1.797 *** | .273 | -1.392 *** | .204 | -1.470 *** | .247 | -1.993 *** | .201 |
| Time dummies – Reference: 1996 or 2006 | | | | | | | | | | | | |
| 1994 or 2004 | -.188 ** | .091 | .013 | .091 | -.032 | .100 | -.074 | .096 | .271 ** | .112 | .176 * | .105 |
| 1995 or 2005 | -.016 | .093 | .093 | .097 | .144 | .101 | .046 | .078 | .180 * | .097 | -.006 | .091 |
| Log-likelihood | 24,964.0 | | | | | | 33,505.3 | | | | | |
| # of individuals | 3,348 | | | | | | 5,884 | | | | | |
| # of observations | 8,064 | | | | | | 10,118 | | | | | |
| LR test for sample selection | $\chi^2(1)=1.7$, $p$-value$=.001$ | | | | | | $\chi^2(1)=.3$, $p$-value$=.584$ | | | | | |

Notes: * Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Table 5: Coefficient Estimates of the Covariates with Correction for Selectivity at Selected Percentiles of the Wage Distribution for Women, 1994-1996 and 2004–2006

| Percentiles | 1994–1996 | | | | | | 2004–2006 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 25th | | 50th | | 75th | | 25th | | 50th | | 75th | |
| Variables | Coef | SE | Coef | SE | Coef | SE | Coef | SE | Coef | SE | Coef | SE |
| Age | -.511*** | .091 | -.664*** | .124 | -.228* | .118 | -.744*** | .088 | -.763*** | .102 | -.737*** | .110 |
| Potential experience | -.095 | .083 | .099 | .109 | -.186* | .104 | .040 | .080 | .189** | .091 | .327*** | .092 |
| Education – Reference: None, elementary, or lower secondary | | | | | | | | | | | | |
| Upper secondary | -1.199*** | .115 | -.807*** | .138 | -.760*** | .178 | -.996*** | .088 | -1.089*** | .102 | -.276** | .130 |
| Post secondary or tertiary | -1.593*** | .285 | -1.494*** | .213 | -1.743*** | .224 | -1.544*** | .116 | -1.521*** | .125 | -.825*** | .151 |
| Occupation – Reference: Blue collar workers | | | | | | | | | | | | |
| High skilled & white collar | -2.722*** | .193 | -2.041*** | .197 | -1.541*** | .258 | -1.628*** | .102 | -1.245*** | .122 | -.031 | .164 |
| Low skilled & white collar | -1.420*** | .116 | -.820*** | .171 | -.707*** | .239 | -.704*** | .088 | -.274** | .112 | .412** | .161 |
| Area of residence – Reference: North-West | | | | | | | | | | | | |
| North-East | .534*** | .128 | .346*** | .128 | .136 | .140 | .136 | .087 | .114 | .099 | .300*** | .099 |
| Centre | .706*** | .126 | .271** | .133 | .386*** | .144 | .545*** | .094 | .228** | .102 | .119 | .109 |
| South | .559*** | .157 | .513*** | .137 | .367** | .151 | .349*** | .110 | -.065 | .128 | -.057 | .123 |
| Bad health | .055 | .111 | .180* | .109 | -.078 | .127 | -.047 | .087 | -.010 | .091 | .034 | .099 |
| Married | -.497*** | .100 | -.225* | .117 | -.057 | .128 | -.129* | .072 | .189** | .085 | .046 | .086 |
| Temporary contract | -.115 | .256 | -.170 | .249 | -.169 | .355 | .351*** | .104 | -.012 | .146 | -.209 | .169 |
| Part-time | -1.286*** | .183 | -1.420*** | .168 | -1.685*** | .180 | -.640*** | .080 | -.928*** | .094 | -1.172*** | .094 |
| Time dummies – Reference: 1996 or 2006 | | | | | | | | | | | | |
| 1994 or 2004 | -.233* | .138 | .016 | .131 | .111 | .165 | .128 | .111 | .205* | .113 | .071 | .128 |
| 1995 or 2005 | .043 | .138 | .166 | .131 | -.086 | .171 | .075 | .090 | .096 | .097 | .245** | .097 |
| Log-likelihood | 17,718.2 | | | | | | 33,505.3 | | | | | |
| # of individuals | 4,340 | | | | | | 7,334 | | | | | |
| # of observations | 10,863 | | | | | | 12,635 | | | | | |
| LR test for sample selection | $\chi^2(1)$=1.8, $p$-value=.001 | | | | | | $\chi^2(1)$=16.7, $p$-value=.000 | | | | | |

*Notes*: * Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

better paid on hourly basis, for both genders and time periods. Finally, the time indicators have always a small and not significant impact on the wage hazard rate, both for men and women. It suggests that the shape of the wage distributions did not change significantly within each time window.

# 4 Counterfactual Wage Distributions and Decompositions of Gender Wage Gaps

## 4.1 The Components of the Gender Wage Gap in the Mid-1990s and in the Mid-2000s

In the previous section we estimated the male and female wage structures, i.e. the set of parameters determining the male and female hazard functions that fully characterize the conditional wage distributions. We can exploit these estimated wage structures to simulate wage distributions by way of the following algorithm, which is run once for men and once for women:

1. Draw a vector of parameter estimates assuming normality around the point estimates $\widehat{\Theta}$ with a variance-covariance matrix equal to the estimated one.

2. Generate $N$ individuals by drawing $N$ times: i) a vector $x$ of covariates from the empirical distribution of covariates among all women (men);[16] ii) unobserved characteristics from the estimated distribution of unobserved heterogeneity $\widehat{G}$.

3. For each baseline wage interval $[w_{j-1}, w_j)$ with $j = 1, \ldots, 71$, we compute the predicted (wage histogram) cumulative distribution function from the following theoretical cumulative distribution function:

$$F(w_j|x_{it}, v_i) \equiv \sum_{r=0}^{j} \left[ S(w_{r-1}|x_{it}, v_i) - S(w_r|x_{it}, v_i) \right].$$

For $i = 1, \ldots, N$, we draw $\kappa_i$ from a standard uniform distribution. If $\widehat{F}(w_{j-1}|x_{it}, \widehat{v}_i) \leq \kappa_i < \widehat{F}(w_j|x_{it}, \widehat{v}_i)$, individual $i$ is assigned a random wage in $[w_{j-1}, w_j)$.[17]

4. Repeat steps 1 to 3 $R = 999$ times to get $R$ independent realizations and build Monte Carlo confidence intervals.

We exploit this simulation algorithm and a slightly modified version to decompose the gender wage gap into the component due to gender differences in the distribution of individ-

---

[16]Work specific covariates are drawn from the empirical distribution conditional on work participation.

[17]We can indeed only simulate an interval in which the wage is located. A point value is assigned by drawing from a uniform distribution between the lower and upper bound of the interval, without losing much in terms of precision given the small size of the intervals.

ual characteristics and the component due to gender differences in the remunerations of the same characteristics. The idea of the decomposition of the gender pay gap is based on recovering the counterfactual distribution that would prevail if men had the same distribution of characteristics as women, that is

$$\int_{x \in \chi} \int_{\Re_+} f_M(w|x, v; \Theta_M) dG(v) dF_F(x), \qquad (7)$$

where the subscripts $F$ and $M$ stand for female and male, respectively. The counterfactual marginal density in (7) is obtained by repeating the simulation algorithm for men with step 2 revised so that the covariate vectors are drawn from the empirical covariate distribution of women.

At this point, by denoting by $Q(\cdot)$ the log-quantile function, we can decompose the gender wage gap at each quantile $q \in [0, 100]$ as follows

$$Q(q|\widehat{\Theta}_M, x_M) - Q(q|\widehat{\Theta}_F, x_F) =$$
$$\left[ Q(q|\widehat{\Theta}_M, x_M) - Q(q|\widehat{\Theta}_M, x_F) \right] + \left[ Q(q|\widehat{\Theta}_M, x_F) - Q(q|\widehat{\Theta}_F, x_F) \right], \quad (8)$$

where, for $G \in \{F, M\}$, $\widehat{\Theta}_G$ is the set of estimated coefficients of the wage structure and $x_G$ is the set of individual characteristics. On the right-hand side of (8), the first term in brackets is the gender wage gap (difference in log points) at quantile $q$ if men and women were equally paid for their own characteristics. In other words, it is the component of the gender wage gap explained by gender differences in the distribution of individual characteristics. The second term in brackets is the gender wage gap at quantile $q$ if men had the same characteristics as women, i.e. the part explained by different coefficients.

Figure 2 reports the gender wage gaps and their decompositions in the mid-1990s (top panel) and in the mid-2000s (bottom panel). The central curves plot the total wage gaps, which are computed by taking the difference between the log male wages and the log female wages at each quantile of the corresponding wage distribution, i.e. the left-hand side of equation (8). The curves at the bottom are instead the wage gaps once gender differences in the remuneration are netted out, i.e. the first term in brackets on the right-hand side of equation (8). Finally, the curves at the top plot the wage gap once gender differences in observed characteristics are taken into account, i.e. the second term in brackets on the right-hand side of equation (8). The corresponding point estimates for selected quantiles are reported in Table 6.

We saw in Section 2 that the raw median wage penalty for women was about 6% in the mid-1990s and 5% in the mid-2000s. When we net out the component due to gender differences in the distribution of individual characteristics and we control for sample selection,

16

the median pay gap increases to 17.4%[18] in the mid-1990s and to 24.6%[19] in the mid-2000s. The importance of controlling for gender differences in the distribution of individual characteristics is even more striking at the top of the distribution. While the raw gender wage gap at the $90^{\text{th}}$ percentile was nil both in the mid-1990s and in the mid-2000s, it rises up to 30% in the mid-1990s and 31.5% in the mid-2000s when gender differences in the distribution of individual characteristics are controlled for. This suggests that in Italy men and women had different levels of labour market qualifications both in the mid-1990s and in the mid-2000s, with women relatively more endowned of qualifications that ensure large rewards, especially at the top of the wage distribution.

Figure 2: Sample Selection Corrected Decomposition of the Gender Wage Gaps (gap in log points)
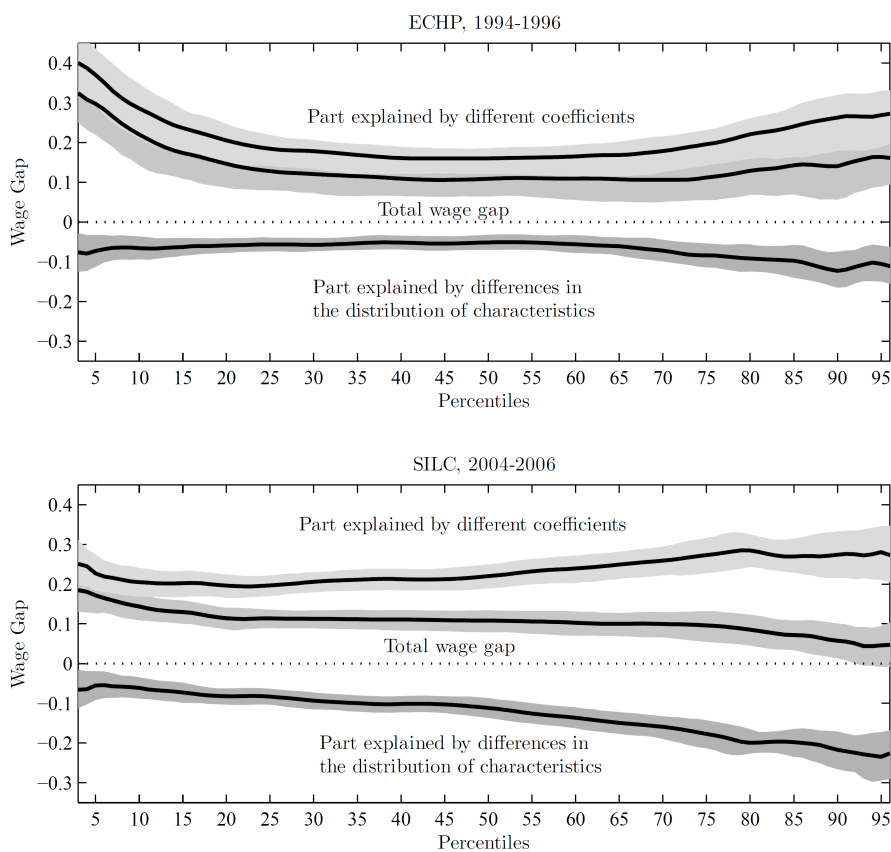


*Notes*: The grey areas are Monte Carlo 95% confidence intervals, computed by 999 replications.

---

[18]$17.4\% = [\exp(0.160) - 1] * 100.$
[19]$24.6\% = [\exp(0.220) - 1] * 100.$

Table 6: Sample Selection Corrected Decomposition of the Gender Wage Gaps (gap in log points) at Selected Quantiles

| Quantiles $q$ | ECHP, 1994–1996 | | | SILC, 2004–2006 | | |
|---|---|---|---|---|---|---|
| | Wage gap | 95% confidence interval[§] | | Wage gap | 95% confidence interval[§] | |
| | | | | | | |
| *Total wage gap: $Q(q|\widehat{\Theta}_M, x_M) - Q(q|\widehat{\Theta}_F, x_F)$* | | | | | | |
| 5 | 0.298* | 0.220 | 0.366 | 0.172* | 0.127 | 0.224 |
| 10 | 0.220* | 0.142 | 0.279 | 0.143* | 0.109 | 0.176 |
| 25 | 0.128* | 0.080 | 0.162 | 0.114* | 0.090 | 0.136 |
| 50 | 0.109* | 0.064 | 0.133 | 0.108* | 0.079 | 0.135 |
| 75 | 0.112* | 0.055 | 0.144 | 0.096* | 0.055 | 0.132 |
| 90 | 0.140* | 0.057 | 0.194 | 0.057* | 0.013 | 0.101 |
| 95 | 0.164* | 0.090 | 0.215 | 0.046 | -0.008 | 0.095 |
| | | | | | | |
| *Part explained by different coefficients: $Q(q|\widehat{\Theta}_M, x_F) - Q(q|\widehat{\Theta}_F, x_F)$* | | | | | | |
| 5 | 0.370* | 0.293 | 0.437 | 0.227* | 0.183 | 0.274 |
| 10 | 0.286* | 0.209 | 0.348 | 0.205* | 0.168 | 0.240 |
| 25 | 0.184* | 0.141 | 0.219 | 0.197* | 0.170 | 0.222 |
| 50 | 0.160* | 0.121 | 0.186 | 0.220* | 0.189 | 0.249 |
| 75 | 0.196* | 0.143 | 0.232 | 0.273* | 0.228 | 0.317 |
| 90 | 0.263* | 0.179 | 0.320 | 0.274* | 0.216 | 0.333 |
| 95 | 0.269* | 0.188 | 0.330 | 0.280* | 0.210 | 0.346 |
| | | | | | | |
| *Part explained by different characteristics: $Q(q|\widehat{\Theta}_M, x_M) - Q(q|\widehat{\Theta}_M, x_F)$* | | | | | | |
| 5 | -0.073* | -0.113 | -0.033 | -0.055* | -0.092 | -0.019 |
| 10 | -0.066* | -0.094 | -0.036 | -0.061* | -0.088 | -0.035 |
| 25 | -0.056* | -0.074 | -0.039 | -0.083* | -0.104 | -0.064 |
| 50 | -0.051* | -0.070 | -0.032 | -0.112* | -0.137 | -0.087 |
| 75 | -0.084* | -0.114 | -0.054 | -0.177* | -0.213 | -0.143 |
| 90 | -0.123* | -0.165 | -0.075 | -0.217* | -0.266 | -0.171 |
| 95 | -0.105* | -0.150 | -0.059 | -0.235* | -0.294 | -0.173 |

*Notes*: * Significant at the 5% level.
[§] Monte Carlo confidence intervals computed by 999 replications.

Estimations of wage distributions and gender wage gap decomposition are corrected for sample selection, so that we take into account that, due to nonrandom selection into employment, non-working men and women have different earnings potential or give different value to earning money.[20] As a matter of fact, social norms according to which men have to be the sole breadwinner might repress female participation in the workforce. As a consequence, only women with very high earnings potential are able to compensate the cost of deviating from the social norms and enter the workforce. If it is so and nonrandom selection into employment is not controlled for, the gender wage gap would be underestimated. As a matter of fact, like in Picchio and Mussida (2011), if we do not correct for sample selection we find lower pay penalties for women.[21] In Italy there is therefore evidence of positive selection of women into the workforce, i.e. women with low earnings potential are less likely to feature in the observed wage distribution than comparable men.

Although we net out the effect of gender differences in the selection into employment, we cannot conclude that the sizeable gender wage gap, persistent over time and even larger after controlling for gender differences in the distribution of individual characteristics, is evidence of discrimination.[22] There might be a discriminatory component, like disparity in the remuneration of similar characteristics or sectoral segregation such that women are excluded from "male" and better rewarded jobs. Nevertheless, there might be non-discriminatory forces at work. For instance, the gap might be due to female deficits in unmeasured skills with high rewards. Alternatively, women might follow social expectations and choose careers which allow them to fulfil family commitments. Lastly, experimental studies suggest that women are more altruistic and less career-oriented than men. For example, Niederle and Vesterlund (2007) find that, given the same level of ability and skills, women are more likely than men to choose pay schemes which are less rewarding and less competitive. This might reflect greater risk aversion or also that women dislike negotiation.

## 4.2 The Trend over Time of the Components of the Gender Wage Gap

If one looks at the raw descriptive statistics in Table 2, it seems that wages of women have caught up, although not fully, with the wages of men in the period that goes from the mid-1990s to the mid-2000s. Although the median and the mean gender wage gap decreased just from 6% to 5%, sizeable changes are instead displayed at the bottom of the wage distribution: the gender pay gap at the 5th percentile went from almost 20% in the mid-1990s to

---

[20]As stressed by the job satisfaction literature (see e.g. Chevalier, 2007), men and women might have different tastes and preferences, leading to different choices in the labour market.

[21]The estimation results and the gender wage gap decompositions without sample selection are available upon request from the authors.

[22]The unexplained component of the gender pay gap is often viewed as discrimination by the literature (Chevalier, 2007). Discriminatory behaviours are nonetheless difficult to observe. Only very detailed information might help finding evidence of discrimination for the process under investigation, e.g. employment discrimination in hiring (Goldin and Rouse, 2000).

12% in the mid-2000s. The sample selection corrected figures reported in the top panel of Table 6 seem to confirm such a time evolution. Once we control for gender differences in the nonrandom selection into employment, the gender wage gap decreases from almost 35% in the mid-1990s to 18.8% in the mid-2000s at the 5$^{\text{th}}$ percentile of the wage distribution, it is rather stable in the centre of the wage distribution, and it goes from about 15% to about 6% at the top of the distribution.

However, once we net out the component due to gender differences in individual characteristics, a different picture pops up. The top graph of Figure 3 and the top panel of Table 7 display the change over time of the gender wage gap if men and women had the same characteristics. They tell us that, from the mid-1990s to the mid-2000s, the gender wage gap if men and women had the same characteristics significantly decreased by 10.6 (6.2) percentage points at the 5$^{\text{th}}$ (10$^{\text{th}}$) percentile, but it significantly increased in the centre of the distribution, especially between the median and the 75$^{\text{th}}$ percentile (4.9 and 6.1 percentage points respectively).

How can we explain rising gender wage gaps in the middle of the distribution and decreasing gender wage gaps at the bottom? Is it that a change in skill prices benefited women relative to men at bottom jobs and men relative to women at better jobs? Or is it that women at better jobs became relatively richer in skills that are paid more and more differently between genders, with time increasing penalties for women?

In order to answer these questions, we parcel out the sources of the time variation of the gender wage gap into a component due to the gender relative changing prices in the labour market and a component due to the time change in female-specific characteristics. Formally, the change over time of the gender wage gap if men and women had the same characteristics, i.e. the second term in brackets of equation (8), is decomposed as follows $\forall q \in [0, 100]$:

$$\left[ Q(q|\widehat{\Theta}_M^{00s}, x_F^{00s}) - Q(q|\widehat{\Theta}_F^{00s}, x_F^{00s}) \right] - \left[ Q(q|\widehat{\Theta}_M^{90s}, x_F^{90s}) - Q(q|\widehat{\Theta}_F^{90s}, x_F^{90s}) \right] = \qquad (9)$$

$$\left\{ \left[ Q(q|\widehat{\Theta}_M^{00s}, x_F^{00s}) - Q(q|\widehat{\Theta}_F^{00s}, x_F^{00s}) \right] - \left[ Q(q|\widehat{\Theta}_M^{90s}, x_F^{00s}) - Q(q|\widehat{\Theta}_F^{90s}, x_F^{00s}) \right] \right\} + \quad (10)$$

$$\left\{ \left[ Q(q|\widehat{\Theta}_M^{90s}, x_F^{00s}) - Q(q|\widehat{\Theta}_F^{90s}, x_F^{00s}) \right] - \left[ Q(q|\widehat{\Theta}_M^{90s}, x_F^{90s}) - Q(q|\widehat{\Theta}_F^{90s}, x_F^{90s}) \right] \right\}. \quad (11)$$

This decomposition illustrates that the change over time of the gender wage gap if men and women had the same characteristics can be explained by the the trend in the wage structure, equation (10), and by the trend in female characteristics, equation (11). The counterfactual quantile functions which appear in equations (9)-(11) are predicted by running simulation algorithms similar to the one described in the previous subsection, but modified so as to draw estimated parameters and individual characteristics by the corresponding time period/gender combination.[23] For example, the values taken by the quantile function $Q(q|\widehat{\Theta}_M^{90s}, x_F^{00s})$, which appears in equations (10) and (11) are simulated by running the fol-

---

[23]As the reliability of these simulations depends on the capacity of our model to predict the realized wage distributions, we compute goodness-of-fit checks of the estimated model and report them in Appendix B.

lowing algorithm:

1. Draw a vector of parameter estimates assuming normality around the point estimates $\widehat{\Theta}_M^{90s}$ with a variance-covariance matrix equal to the estimated one.

2. Generate $N$ women by drawing $N$ times: i) a vector $x$ of covariates from the empirical distribution of covariates among all women in 2004–2006; ii) the unobserved component from the estimated distribution of unobserved heterogeneity $\widehat{G}_M^{90s}$.

3. For each baseline wage interval $[w_{j-1}, w_j)$ with $j = 1, \ldots, 71$, we compute the predicted (wage histogram) cumulative distribution function from the following theoretical cumulative distribution function:

$$F(w_j | x_{it}, v_i) \equiv \sum_{r=0}^{j} \big[ S(w_{r-1} | x_{it}, v_i) - S(w_r | x_{it}, v_i) \big].$$

For $i = 1, \ldots, N$, we draw $\kappa_i$ from a standard uniform distribution. If $\widehat{F}(w_{j-1} | x_{it}, \widehat{v}_i) \leq \kappa_i < \widehat{F}(w_j | x_{it}, \widehat{v}_i)$, individual $i$ is assigned a random wage in $[w_{j-1}, w_j)$.

4. Repeat steps 1 to 3 $R = 999$ times to get $R$ independent realizations.

The central parts of Figure 3 and Table 7 display the component of the time variation of the gender wage gap due to the gender relative price change in the labour market. The gender relative changes in prices of observed and unobserved characteristics significantly increased the gender wage gap by 7.9 and 8.9 percentage points at the median and at the 75$^{\text{th}}$ percentile of the wage distribution, respectively. Women at the bottom were instead favoured, with a 6.1 percentage points decrease in the gender wage gap, however not significant at 5%. The bottom panels of Figure 3 and Table 7 report the component due to the time change in female-specific characteristics. The impact, if any, was negative: women became richer (poorer) in skills which are characterized less (more) by female penalties in rewards. The trend in the female distribution of individual characteristics gave however a small and not significant contribution to the reduction of the gender wage gap over time.

## 4.3 Discussion

Hence, as found by Blau and Khan (1997) in the US, in recent years Italian women have been mostly swimming upstream, i.e. the wage structure has been increasingly unfavourable, especially for women at centre-top jobs. What are the possible reasons for which women at the centre-top of the wage distribution swam against the tide in the decade under analysis? Why did women at the bottom of the wage distribution experience instead a slightly favourable trend in the gender relative wage structure? We claim that the institutional change determined by the labour market reforms might have generated both favourable and unfavourable trends in the female wage structure, with different relative sizes across the wage distribution.

21

# Figure 3: Decomposition of the Change over Time of the Gender Wage Gap
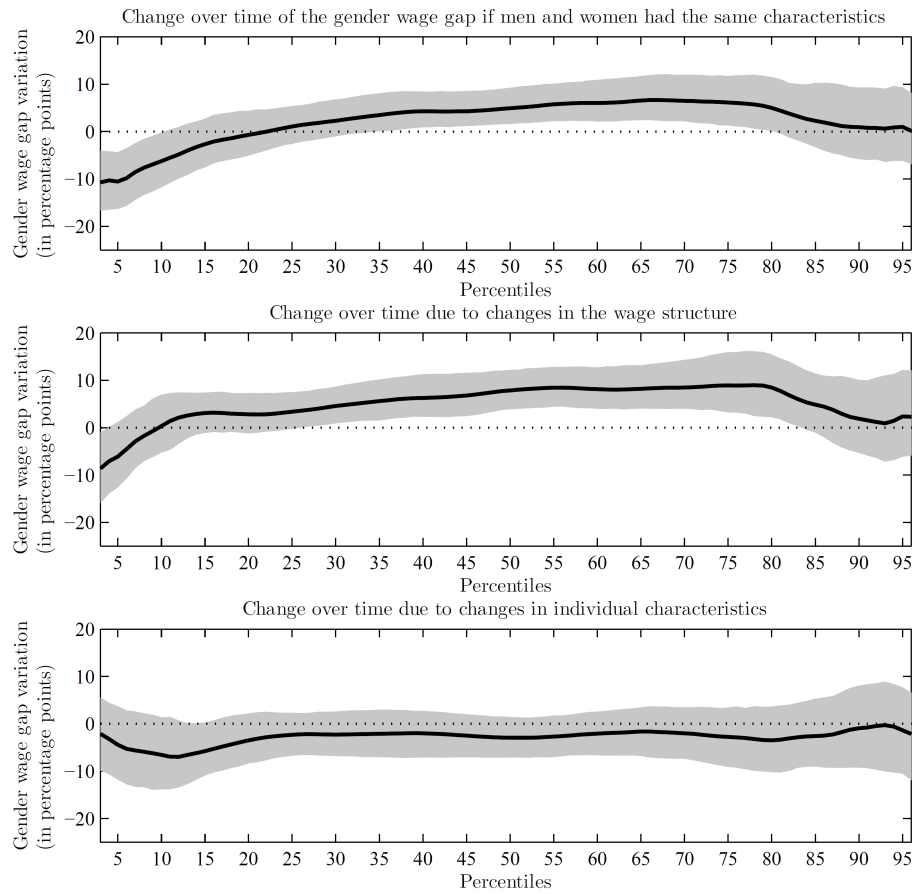


*Notes*: The grey areas are Monte Carlo 95% confidence intervals, computed by 999 replications.

Table 7: Decomposition of the Change over Time of the Gender Wage Gap at Selected Quantiles

| Quantiles $q$ | Gender wage gap variation in percentage points | 95% confidence interval[§] | |
|---|---|---|---|
| *Change over time of the gender wage gap if men and* | | | |
| *women had the same characteristics: equation* (9) | | | |
| 5 | -10.6* | -16.4 | -4.4 |
| 10 | -6.2* | -11.8 | -0.2 |
| 25 | 1.1 | -2.6 | 5.4 |
| 50 | 4.9* | 1.6 | 9.2 |
| 75 | 6.1* | 1.3 | 11.7 |
| 90 | 0.9 | -5.9 | 9.2 |
| 95 | 1.0 | -6.2 | 9.3 |
| *Change over time due to changes in the wage structure: equation* (10) | | | |
| 5 | -6.1 | -12.6 | 1.2 |
| 10 | 0.3 | -5.2 | 7.0 |
| 25 | 3.4 | -0.3 | 7.5 |
| 50 | 7.9* | 3.9 | 12.2 |
| 75 | 8.9* | 3.2 | 16.0 |
| 90 | 1.9 | -5.3 | 10.1 |
| 95 | 2.3 | -6.1 | 12.2 |
| *Change over time due to changes in individual characteristics: equation* (11) | | | |
| 5 | -4.4 | -11.8 | 3.7 |
| 10 | -6.6 | -13.9 | 1.4 |
| 25 | -2.3 | -6.8 | 2.3 |
| 50 | -2.9 | -7.1 | 1.4 |
| 75 | -2.8 | -9.1 | 3.4 |
| 90 | -1.0 | -9.3 | 8.1 |
| 95 | -1.3 | -10.5 | 7.7 |

*Notes*: * Significant at the 5% level.
[§] Monte Carlo confidence intervals computed by 999 replications.

As mentioned in the Introduction, from the mid-1990s until the mid-2000s, the Italian labour market went through important reforms, which essentially made it easier to create flexible and atypical job contracts, while leaving unchanged the legislation and the protection of employees with open-ended contracts. This shaped a two tier regime (Boeri and Garibaldi, 2007; OECD, 2004, chapter 2) with, on one side, highly protected workers (mostly male unionized workers) and, on the other side, highly flexible jobs (part-time jobs, seasonal jobs, youth work-training, on call jobs, and work on project) especially for the marginal and weakest labour market segments.[24] Therefore, after the labour market reforms women have been more likely to be at work in more flexible, less unionized,[25] and less protected segments of the labour market, where special treatments or discriminatory practices on gender ground might arise or become more marked.

Nevertheless, the introduction of flexible work arrangements might have generated a favourable trend in the female wage structure by increasing the likelihood that women have better opportunities to combine work and family life: flexible work arrangements provide women not only with ports of entry into employment in the form of flexible contracts, but also with springboards towards jobs with better employment conditions. As a matter of fact, Barbieri and Sestito (2008) find that in Italy temporary work increases future chances of having a "satisfactory employment" especially for women.

However, the stepping stone effect of atypical contracts might have been heterogeneous across the sample. The introduction of flexible and atypical contracts can especially increase the chances of the most disadvantaged groups of being at work with better employment conditions.[26] This might explain why we observe a slight improvement in the wage structure only for women at the bottom jobs.

Finally, these explanations need to be qualified as there might be other forces at work. For example, as pointed out by Mulligan and Rubinstein (2008), if men and women are not perfect substitutes in production and the female workforce grows (as it did in the time window under analysis in Italy), women's relative wages are reduced, other things constant. Hence, even if our findings about the wage structure becoming increasingly unfavourable for women are correct, we cannot exclude that there might be other factors explaining this trend.

---

[24]Empirical studies have shown that the introduction of new atypical work arrangements segmented the Italian labour market between the insider core labour force and a secondary segment of unstable and flexible jobs (Barbieri and Scherer, 2008).

[25]Union density rates among atypical workers can be lower for different reasons (Goslinga and Sverke, 2003). For instance, employees with atypical contracts are more difficult to organize for unions (Gallagher and Sverke, 2005). Moreover, unions are reluctant to accept temporary contracts or part-time contracts as unions often oppose the growth of forms of employment which can be exploited by employers as a mere buffer against the business cycle, or which go towards an overall reduction of working time (Delsen, 1995; Barling and Gallagher, 1996).

[26]For example, Cockx and Picchio (2011) find that for Belgian youth short-lived jobs are stepping stones to long-lasting jobs especially for more disadvantaged individuals, e.g. the lower educated and those living in areas where the unemployment rate is higher.

# 5 Conclusions

The Italian gender wage gap from raw data, either in mean or at several quantiles of the wage distribution, decreased in the time window that goes from the mid-1990s until the mid-2000s. The closing raw gender wage gap might induce to think that gender inequality in the Italian labour market diminished and that gender discriminatory practices subsided. The decreased raw gender wage gap was however coincident with relevant developments in the labour market, mostly in terms of legislation, but also in terms of workforce participation and composition. Since the late 1990s Italy has undertaken steps towards "flexibility" by way of labour market reforms that have liberalized the use of flexible and atypical contracts. The female employment rates and the female educational levels have experienced a sizeable increase, both in absolute terms and relatively to men. These labour market developments might have partially contributed to a falling gender gap, but there are reasons to believe that these important changes might have asymmetrically hit different components of the gender wage gap. If it is so, the raw evidence of decreasing gender wage gaps might simply be a statistical artefact, with the risk of generating misleading conclusions.

As a matter of fact, the gender wage gap can be decomposed into a component due to gender differences in the distribution of individual characteristics and a component due to gender differences in the reward of the same characteristics. These two components might have had different time trends. Understanding the evolution over time of each component and the contribution of each component to the evolution of the gender wage gap can be very policy relevant to design interventions that are effective in pursuing the gender equality target.

The empirical analysis exploited the estimator of probability distribution functions in the presence of covariates and sample selection proposed by Picchio and Mussida (2011) and microsimulations to decompose the gender wage gap and its variation over time. We decomposed the gender wage gap at each quantile of the wage distribution into a part due to different coefficients determining the wage structure and a part due to different individual characteristics. The empirical analysis was based on 1994–1996 ECHP data and on 2004–2006 SILC data.

When we netted out the component due to gender differences in the distribution of individual characteristics and we controlled for sample selection, the median pay gap for women increased from 6% to 17% in the mid-1990s, and from 5% to the 25% in the mid-2000s. Hence, if men and women had had the same distribution of individual characteristics fixed at the corresponding time period, the gender wage gap at the median of the wage distribution would have increased over the decade under analysis from 17% to 25%.

When we parcelled out the source of the time variation, we found that women in Italy have been mostly swimming upstream, i.e. the wage structure has been increasingly unfavourable, especially for women at centre-top jobs. The trend in the female distribution of individual characteristics contributed instead to the reduction of the gender wage gap over time. However, this negative impact on the gender wage gap was not significant and too

small in size to compensate the positive impact of the relative change in the wage structure.

We claimed that the labour market reforms might explain why the trends in the wage structure were unfavourable to women, especially at the centre-top of the wage distribution. The institutional change shaped a two tier regime: on one side, highly protected workers (mostly male unionized workers) and, on the other side, highly flexible jobs especially for the marginal and weakest labour market segments. Therefore, after the labour market reforms women have been more likely to be at work in more flexible, less unionized, and less protected segments of the labour market, where special treatments, prejudices, or discriminatory practices on gender ground might arise or become more marked. Nevertheless, the labour market reforms might have provided women, especially at the bottom jobs, not only with ports of entry into employment in the form of flexible contracts, but also with springboards towards jobs with better employment conditions, explaining why we found a slight improvement in the wage structure for women at the bottom of the wage distribution.

The conclusion to be drawn from this research is that more effective policies in pursuing the gender equality target are needed in Italy. Policy interventions should aim at bringing about a less segregated labour market, as in less protected segments discriminatory practices are more likely to arise. Furthermore, flexible contracts, which can enhance gender equality by ensuring that women have a better access to good quality jobs and better opportunities to reconcile family and work, might not be enough as long as women are trapped into traditional roles by social norms. As a matter of fact, flexible jobs might be a barrier hindering gender equality, if they only ensure that women continue to bear on their shoulders most of the family responsibilities and to be confined to part-time or marginal jobs. Childcare and family policies can therefore be the leverage to make flexibility the way through which women can fully commit themselves to work.

# Appendix

## A   Non Random Selection into Employment

Table 8 reports the estimation results of the discrete mixture logit model for the probability of employment by gender and by time periods. The results are in line with the expectations. The employment probability decreases with age but increases with potential experience. Higher educated people are more likely to be at work. While in the mid-1990s the employment probability was the highest in the North-East, followed by the North-West and the Centre, and the lowest in the South, in the mid-2000s the North and the Centre shared the same employment probability. Family structure and married status have opposite effects on work participation between men and women, strongly consistent with the male breadwinner system: married (wo)men have a (lower) higher probability of being employed; work participation is increasing in the number of children for men, whereas it is decreasing in the number of children for women. Finally, the number of household members and a bad health

condition reduce the employment probability both for men and women.

Table 8: Estimation Results of the Employment Equation by Gender and Time Period

| | 1994–1996 | | | | | | 2004–2006 | | | | | |
| | Men | | | Women | | | Men | | | Women | | |
| | Coef | | SE | Coef | | SE | Coef | | SE | Coef | | SE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age (years) | -.696 | *** | .106 | -.926 | *** | .074 | -1.341 | *** | .147 | -1.767 | *** | .113 |
| Potential exper. (years) | 1.423 | *** | .095 | 2.720 | *** | .077 | 1.572 | *** | .126 | 2.436 | *** | .095 |
| Education – Reference: None, elementary, or lower secondary | | | | | | | | | | | | |
|   Upper secondary | 1.173 | *** | .138 | 2.250 | *** | .122 | 1.799 | *** | .167 | 2.508 | *** | .161 |
|   Post secon. or tertiary | 1.142 | *** | .203 | 2.479 | *** | .190 | 1.881 | *** | .219 | 4.274 | *** | .200 |
| Area of residence – Reference: North-West | | | | | | | | | | | | |
|   North-East | 1.489 | *** | .292 | -.370 | * | .208 | .118 | | .249 | .454 | ** | .209 |
|   Centre | -1.118 | *** | .206 | -.741 | *** | .176 | -.330 | | .251 | .013 | | .203 |
|   South | -2.512 | *** | .208 | -1.336 | *** | .159 | -2.424 | *** | .261 | -1.524 | *** | .197 |
| Bad health | -.460 | *** | .138 | -.138 | | .142 | -1.136 | *** | .170 | -.367 | ** | .147 |
| Married | 2.466 | *** | .188 | -.957 | *** | .143 | 2.635 | *** | .237 | -1.594 | *** | .165 |
| # household members | -2.538 | *** | .474 | -2.169 | *** | .462 | -2.316 | *** | .596 | -2.674 | *** | .615 |
| Presence kids<12 years | -.113 | | .201 | -.121 | | .178 | 1.075 | *** | .189 | -.629 | *** | .153 |
| Time dummies – Reference: 1996 or 2006 | | | | | | | | | | | | |
|   1994 or 2004 | .747 | *** | .174 | .440 | ** | .195 | .005 | | .179 | .624 | *** | .170 |
|   1995 or 2005 | .267 | | .185 | .202 | | .203 | .288 | | .177 | .158 | | .172 |
| Unobserved heterogeneity support points and probability masses | | | | | | | | | | | | |
| $\varepsilon_1$ | 2.285 | *** | .280 | .994 | *** | .247 | 2.552 | *** | .330 | 2.307 | *** | .285 |
| $\varepsilon_2$ | -4.464 | *** | .349 | -9.632 | *** | .353 | -4.217 | *** | .371 | -5.985 | *** | .333 |
| $\lambda_1$ | -.066 | | .054 | -.100 | | .067 | -.052 | | .045 | .244 | *** | .053 |
| $\lambda_2$ | -1.115 | *** | .109 | .427 | *** | .097 | -1.811 | *** | .184 | .070 | | .099 |
| $\lambda_3$ | -2.267 | *** | .275 | -2.146 | * | 1.104 | -1.715 | *** | .165 | -1.003 | *** | .260 |

*Notes*: * Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

## B   Goodness-of-Fit

In this appendix we check the goodness-of-fit of the model by contrasting empirical aspects of the data with those predicted by model simulations. Given the mixture of parametric and non-parametric assumptions on our econometric model and that the counterfactual exercises in Section 4 are based on simulations, it is indeed important to assess the ability of the model to provide quantitative predictions of the statistics of primary interest. Predictions are computed by implementing the following simulation algorithm for men and women, using ECHP data for the mid-1990s and SILC data for the mid-2000s:

1. Draw a vector of parameter estimates assuming normality around the point estimates $\widehat{\Theta}$ with a variance-covariance matrix equal to the estimated one.

2. Generate $N$ individuals by drawing $N$ times: i) $z$ and $x$ from the empirical distribution of covariates; ii) unobserved characteristics from the estimated distribution of unobserved heterogeneity $\widehat{G}$.

3. Simulate work participation by a lottery based on the predicted probability of work participation. More in detail, for $i = 1, \ldots, N$ we draw $\eta_i$ from a standard uniform

distribution. If $\Lambda(z'_{it}\widehat{\delta}_F + \widehat{\varepsilon}_i) > \eta_i$, individual $i$ participates. Otherwise, she does not participate and her simulation is halted.

4. Simulate the wage distribution over the subsample of participating individuals. For each baseline wage interval $[w_{j-1}, w_j)$ with $j = 1, \ldots, 71$, we compute the predicted (wage histogram) cumulative distribution function from the following theoretical cumulative distribution function:

$$F(w_j|x_{it}, v_i) \equiv = \sum_{r=0}^{j} \left[ S(w_{r-1}|x_{it}, v_i) - S(w_r|x_{it}, v_i) \right].$$

For $i = 1, \ldots, N$, we draw $\kappa_i$ from a standard uniform distribution. Individual $i$ is assigned a wage randomly drawn from a uniform distribution in $[w_{j-1}, w_j)$ if $\widehat{F}(w_{j-1}|x_{it}, \widehat{v}_i) \leq \kappa_i < \widehat{F}(w_j|x_{it}, \widehat{v}_i)$.

5. Repeat steps 1 to 4 $R = 999$ times to get $R$ independent realizations and build Monte Carlo confidence intervals.

Once we run this simulation algorithm for men and women in both time periods, we derive, per each time period, the predicted gender wage gap by contrasting the male simulated wage distribution with the female one. The goodness-of-fit can be verified by checking whether the actual gender pay gap lies within the confidence intervals of the simulated one.

The top panel of Table 9 and the top graph of Figure 4 report the goodness-of-fit of the model in predicting the gender wage gap in the mid-1990s using ECHP data. The panel in the middle of Table 9 and the graph in the middle of Figure 4 focus on the goodness-of-fit of the model in predicting the gender wage gap in the mid-2000s using SILC data. Finally, at the bottom of Table 9 and Figure 4, we contrast the actual variation and the predicted variation over time of the gender wage gap. We check thereby the goodness of the model in fitting the change over time in the gender wage gap, i.e. the statistic of primary interest in this study.

The model perfectly fits the gender wage gaps observed in the mid-2000. Nevertheless, it shows some problems in fitting the gender wage gaps in the mid-1990s: the fit is fine until the $80^{\text{th}}$ percentile of the wage distribution; thereafter the model somewhat systematically overpredicts the gender wage gap, especially at the $95^{\text{th}}$ percentile. However, the poor ability of the model in fitting gender wage gaps in the mid-1990s is limited to a small segment of the wage support. Finally, the bottom panel of Table 9 and the bottom graph of Figure 4 show that the model predicts very well the change over time in the gender wage gap. Only at the very top of the wage distribution the actual change lies outside the confidence interval of the simulated one: as a consequence of the overprediction of the gender wage gap at the top of the distribution in the mid-1990s, the variation in the gender wage gap is underpredicted in that region.

Table 9: Goodness-of-Fit: The Gender Wage Gap and Its Variation over Time at Selected Quantiles

| Quantile | Actual[†] | Simulated | 95% confidence interval[§] | |
|---|---|---|---|---|
| | | *Gender wage gap in 1994–1996 (log points)* | | |
| 5 | **0.182** | 0.204 | 0.147 | 0.267 |
| 10 | **0.107** | 0.124 | 0.090 | 0.162 |
| 25 | **0.081** | 0.085 | 0.064 | 0.105 |
| 50 | **0.059** | 0.064 | 0.045 | 0.080 |
| 75 | **0.036** | 0.058 | 0.031 | 0.082 |
| 90 | -0.008 | 0.049 | 0.008 | 0.089 |
| 95 | -0.011 | 0.084 | 0.049 | 0.117 |
| | | *Gender wage gap in 2004–2006 (log points)* | | |
| 5 | **0.113** | 0.119 | 0.078 | 0.157 |
| 10 | **0.086** | 0.092 | 0.069 | 0.117 |
| 25 | **0.070** | 0.075 | 0.057 | 0.092 |
| 50 | **0.048** | 0.060 | 0.039 | 0.081 |
| 75 | **0.031** | 0.042 | 0.015 | 0.068 |
| 90 | **0.000** | 0.020 | -0.015 | 0.057 |
| 95 | **-0.002** | 0.015 | -0.032 | 0.063 |
| | | *Change over time in the gender wage gap* | | |
| | | *(in percentage points)* | | |
| 5 | **-6.91** | -7.25 | -13.39 | -1.40 |
| 10 | **-2.06** | -2.88 | -6.88 | 0.95 |
| 25 | **-1.16** | -0.90 | -3.43 | 1.65 |
| 50 | **-1.16** | -0.38 | -2.99 | 2.33 |
| 75 | **-0.51** | -1.50 | -5.00 | 2.03 |
| 90 | **0.80** | -2.76 | -7.96 | 2.43 |
| 95 | 0.87 | -6.59 | -12.26 | -0.60 |

[†] Actual gender wage gaps lying in the 95% confidence interval of the simulated ones are in bold.

[§] Monte Carlo confidence intervals computed by 999 replications.

Figure 4: Goodness-of-Fit: The Gender Wage Gap and Its Variation over Time
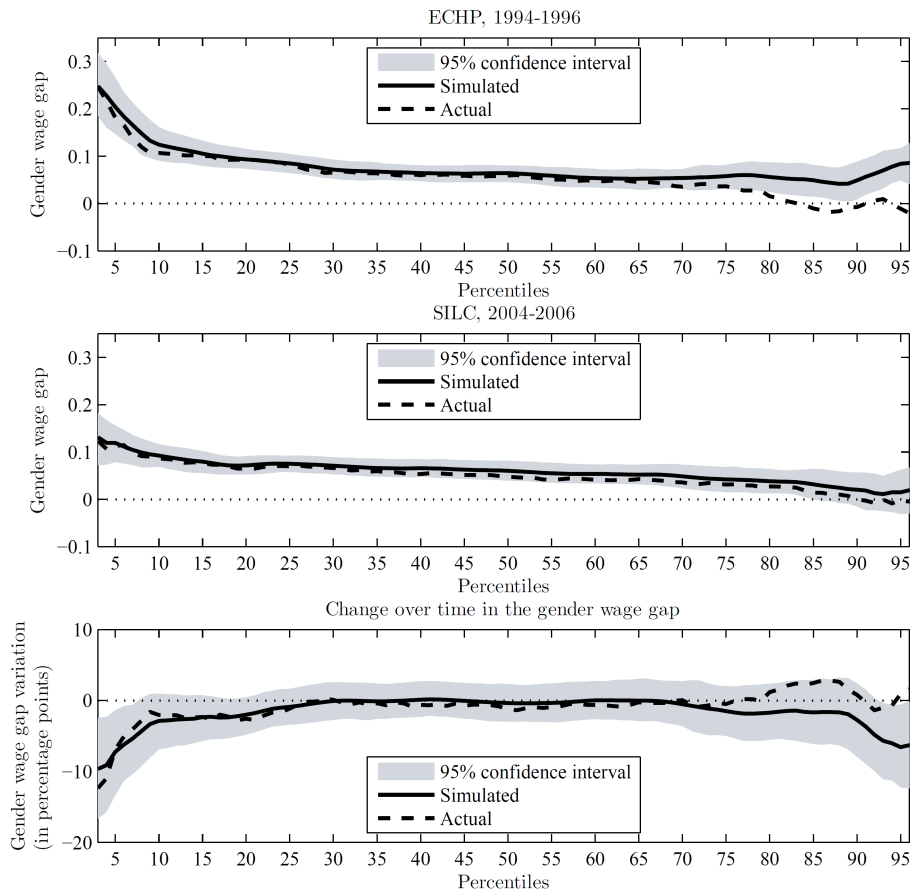


*Notes*: The grey areas are Monte Carlo 95% confidence intervals, computed by 999 replications.

# References

**Addabbo, T. and D. Favaro**, "Education and Wage Differentials by Gender in Italy," 2007. CHILD Working Paper No. 4, Centre for Household, Income, Labour and Demographic economics, Italy.

— , — , **and S. Magrini**, "The Distribution of the Gender Wage Gap in Italy: Does Education Matter?," 2007. RECent Working Paper No. 4, Center for Economic Research, University of Modena and Reggio Emilia.

**Albrecht, J., A. van Vuuren, and S. Vroman**, "Counterfactual Distributions with Sample Selection Adjustments: Econometric Theory and an Application to the Netherlands," *Labour Economics*, 2009, *16* (4), 383–396.

**Barbieri, P. and G. Cutuli**, "Equal Job, Unequal Pay – Fixed Term Contracts and Wage Differentials in the Italian Labour Market," 2009. Working Paper No. 45, Department of Sociology and Social Reseach, University of Trento, Trento.

— **and P. Sestito**, "Temporary Workers in Italy: Who Are They and Where They End Up," *Labour*, 2008, *22* (1), 127–166.

— **and S. Scherer**, "Increasing Labour Market Instability? The Case of Italy," in H. Blossfeld, S. Buchholz, E. Bukodi, and K. Kurz, eds., *Young Workers, Globalization and the Labor Market: Comparing Early Working Life in Eleven Countries*, Cheltenham: Edwar Elgar, 2008, chapter 7.

**Barling, J. and D.G. Gallagher**, "Part-Time Employment," in C.L. Cooper and I.T. Robertson, eds., *International Review of Industrial and Organizational Psychology, Vol. 11*, New York: John Wiley, 1996, pp. 243–277.

**Blau, F.D. and L.M. Khan**, "Swimming Upstream: Trends in the Gender Wage Differential in the 1980s," *Journal of Labor Economics*, 1997, *15* (1), 1–42.

— **and** — , "Analyzing the Gender Pay Gap," *Quarterly Review of Economics and Finance*, 1999, *39* (5), 625–646.

**Blundell, R., A. Gosling, H. Ichimura, and C. Meghir**, "Changes in the Distribution of Male and Female Wages Accounting for Employment Composition using Bounds," *Econometrica*, 2007, *75* (2), 323–363.

**Boeri, T. and G. Garibaldi**, "Two Tier Reforms of Employment Protection: A Honeymoon Effect?," *Economic Journal*, 2007, *117*, 357–385.

**Buchinsky, M.**, "Changes in the U.S. Wage Structure 1963–1987: Application of Quantile Regression," *Econometrica*, 1994, *62* (2), 405–458.

**Chevalier, A.**, "Education, Occupation and Career Expectations: Determinants of the Gender Pay Gap for UK Graduates," *Oxford Bulletin of Economics and Statistics*, 2007, *69* (6), 819–842.

**Cockx, B. and M. Picchio**, "Are Short-lived Jobs Stepping Stones to Long-lasting Jobs?," *Oxford Bulletin of Economics and Statistics*, 2011, *forthcoming.*

**Delsen, L.**, *Atypical Employment: An International Perspective. Causes, Consequences and Policy*, Groningen: Wolters-Noordhof, 1995.

**DiNardo, J., N.M. Fortin, and T. Lemieux**, "Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach," *Econometrica*, 1996, *64* (5), 1001–1044.

**Donald, S.G., D.A. Green, and H.J. Paarsch**, "Differences in Wage Distributions between Canada and the United States: An Application of a Flexible Estimator of Distribution Functions in the Presence of Covariates," *Review of Economic Studies*, 2000, *67* (4), 609–633.

**Eurostat**, *Description of Target Variables: Cross-sectional and Longitudinal, Doc. EU-SILC 065/2004* 2004.

**Fortin, N.M. and T. Lemieux**, "Rank Regressions, Wage Distributions, and the Gender Gap," *Journal of Human Resources*, 1998, *33* (3), 610–643.

**Gallagher, D.G. and M. Sverke**, "Contingent Employment Contracts: Are Existing Theories Still Relevant?," *Economic and Industrial Democracy*, 2005, *26* (2), 181–203.

**Goldin, C. and C. Rouse**, "Orchestrating Impartiality: The impact of "Blind" Auditions on Female Musicians," *American Economic Review*, 2000, *90* (4), 715–741.

**Goslinga, S. and M. Sverke**, "Atypical Work and Trade Union Membership: Union Attitudes and Union Turnover among Traditional vs Atypically Employed Union Members," *Economic and Industrial Democracy*, 2003, *24* (2), 290–312.

**Heckman, J.J. and B. Singer**, "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 1984, *52* (2), 271–320.

**Kugler, A. and G. Pica**, "Effects of Employment Protection on Worker and Job Flows: Evidence from the 1990 Italian Reform," *Labour Economics*, 2008, *15* (1), 78–95.

**Lazear, E.P.**, "Job Security Provisions and Employment," *Quarterly Journal of Economics*, 1990, *105* (3), 699–726.

**Machado, J.A. and J. Mata**, "Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression," *Journal of Applied Econometrics*, 2005, *20* (4), 445–465.

**Mroz, T.A.**, "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions," *Econometrica*, 1987, *55* (4), 765–799.

**Mulligan, C.B. and Y. Rubinstein**, "Selection, Investment, and Women's Relative Wages Over Time," *Quarterly Journal of Economics*, 2008, *123* (3), 1061–1110.

**Niederle, M. and L. Vesterlund**, "Do Women Shy Away From Competition? Do Men Compete Too Much?," *Quarterly Journal of Economics*, 3 2007, *122*, 1067–1101.

**OECD**, *Employment Outlook*, Paris: OECD, 2004.

**Olivetti, C. and B. Petrongolo**, "Unequal Pay or Unequal Employment? A Cross-Country Analysis of Gender Gaps," *Journal of Labor Economics*, 2008, *26* (4), 621–654.

**Picchio, M. and C. Mussida**, "Gender Wage Gap: A Semi-Parametric Approach With Sample Selection Correction," *Labour Economics*, 2011, *forthcoming.*

**Rustichelli, E.**, "I differenziali retributivi di genere," 2005. Quaderni Spinn, Ministero del Lavoro, 79–106.

**Schindler, M.**, "The Italian Labor Market: Recent Trends, Institutions and Reform Options," 2009. IMF Working Paper No. 47, International Monetary Fund.

**Sciulli, D.**, "Labour Market Transitions in Italy: The Effects of the Treu Law on Flows from an Atypical Contract," 2006. CEIS Working Paper No. 232, Tor Vergata University, Rome.

— , "Making the Italian Labour Market More Flexible: An Evaluation of the Treu Reform," 2006. Economics Working Paper No. 34, Department of Economics, University Carlos III, Madrid.

**Van den Berg, G.J. and M. Lindeboom**, "Attrition in Panel Survey Data and the Estimation of Multi-State Labor Market Models," *Journal of Human Resources*, 1998, *33* (2), 458–478.

— **, — , and G. Ridder**, "Attrition in Longitudinal Panel Data and the Empirical Analysis of Dynamic Labour Market Behaviour," *Journal of Applied Econometrics*, 1994, *9* (4), 421–435.