ISSN: 2281-1346

Department of Economics & Management

UNIVERSITÀ DI PAVIA

DEM Working Paper Series

Assessing Changes in Wage Gaps: A New Approach

Carolina Castagnetti
(University of Pavia)

Luisa Rosti
(University of Pavia)

Marina Töpfer
(University of Erlangen-Nuremberg)

 # 170 (03-19)

Via San Felice, 5
I-27100 Pavia

economiaweb.unipv.it

Assessing Changes in Wage Gaps: A New Approach*

Carolina Castagnetti[1], Luisa Rosti[1], and Marina Töpfer[§2]

[1] *Department of Economics and Management, University of Pavia, Italy*

[2] *School of Business and Economics, University of Erlangen-Nuremberg, Germany*

February 26, 2019

Abstract

Little work has been done on directly estimating differences in pay gaps. Studies that estimate pay differentials generally compare them across different sub-samples or rely on decompositions that are based on the assumption of independent errors across samples. Both methods contain serious drawbacks that we overcome by proposing an extension of the Oaxaca-Blinder decomposition. Our proposed method overcomes both the index number and the indeterminacy problem of standard Oaxaca-Blinder decompositions. In addition, like the standard decomposition, our proposed approach can be extended beyond the mean by using linear unconditional quantile regressions and can be decomposed in detail. We present two empirical applications to illustrate the methodology.

Keywords: Pay Differential; Decomposition; Index Number Problem.

JEL - Classification: J7, J13, J310

*We thank Nabanita Datta Gupta for providing us with the Appendix of Gupta *et al.* 2006, where the standard errors of the Juhn *et al.* (1991) decomposition are derived.

§Correspondence to: University of Erlangen-Nuremberg, School of Business and Economics, Chair of Labor and Regional Economics, Lange Gasse 20, 90403 Nuremberg. E-mail: `marina.toepfer@fau.de`

1 Introduction

Gender gaps in the labour market have obtained attention from both policymakers and researchers, leading to the implementation of equal pay legislation and the promotion of equal opportunities. Diverse policies (anti-discrimination laws, board quotas for women, and family-friendly policies) have been implemented to counter gender-based pay differences, but these differences persist (see for example Godin, 2014; Blau and Kahn, 2017).

Gender-related wage gaps are found across sectors occupations, and countries, as well as over time. According to Eurostat, the average Gender Pay Gap (GPG) in the European Union lay at 16 percent in 2016, but this percentage differs widely across countries. Italy and Portugal have average GPGs well below 10 percent, while the gap in Germany, Great Britain, Estonia and Slovakia is above 20 percent (Eurostat, 2015).

In this paper, we propose an extension of the Oaxaca-Blinder (OB) decomposition (Blinder, 1973; Oaxaca, 1973) for estimating wage gaps. Our approach draws inferences directly from the changes in wage gaps between groups across sub-samples and compares the estimated components between groups and across sub-samples. Therefore, it can test whether a significant change has occurred in the explained or unexplained parts of the decomposition of interest, such as whether the components of the GPG have changed statistically significantly over time. Even though most applications of the OB decomposition are found in the labour market and discrimination literature (see Stanley and Jarrell (1998) and Weichselbaumer and Winter-Ebmer (2005) for meta studies), our method, like the standard OB decomposition, can be employed to study (the evolution of) group differences in any (continuous and unbounded) outcome variable. Also like the standard decomposition, our proposed approach can be extended beyond the mean by using linear unconditional quantile regressions and can be decomposed in detail.

We illustrate our method by presenting two empirical applications, one regarding the evolution of the GPG in Italy from 2005 to 2016 and one involving the private-public sector wage gap (PPWG) between women and men in 2016 in Italy. For each application, we compare the standard OB decomposition to our proposed approach. For the first case, our findings reveal differences in results when our proposed estimation methodology is applied compared to when the standard approach is applied.[1] Our proposed methodology reveals that human capital and individual characteristics are the only statistically significant force driving the convergence of the GPG at the bottom and middle of the wage distribution in the last decade in Italy, while at the top is a statistically significant reduction in the unexplained part. In contrast, comparing the components of the GPGs using the standard OB decomposition reveals differences in observable wage characteristics and that the unexplained part of the gap played a role in closing the gap over the last decade in Italy at all points of the wage distribution. For the second case, we can explain the difference in the PPWG between men and women only using our detailed decomposition. The results obtained again differ from the conclusions drawn using the standard approach.

In general, studies that examine changes in the GPG over time or between groups and sectors use either the Juhn *et al.* (1991) method (Blau and Kahn, 1997, 2006) or the double OB decomposition (Smith and Welch, 1989). Both methods rely on estimations obtained on different sub-samples, but their conclusions about what drives changes in wage gaps between groups may differ when these drivers are estimated directly or when conclusions are based on comparing results obtained on different sub-samples. Therefore, in the second case, when estimations are obtained on different sub-samples, it is not possible to draw inferences about the difference in the components of interest. By decomposing the wage gap of interest in an explained and an unexplained part for different sub-samples, the literature has identified various causes of the GPG (see Blau and Kahn, 2017, for an overview). GPGs also differ across time; in particular, declining gaps are observed in recent decades (see Blau and Kahn, 2006; England, 2006) because women are catching up to men in terms of education and experience in the labour market (Goldin, 2006) and because of technical development (Black and Spitz-Oener, 2010), changes in attitudes toward women in the labour market, less occupational segregation (Cotter, 2004; England, 2006; Mandel and Semyonov, 2014), and increasing numbers of anti-discrimination laws (Fortin, 2015). Most studies have found decreasing differences in the explained part(Godin, 2014) of the GPG, while the results in the literature have been ambiguous concerning the unexplained part. For example, (Mandel and Semyonov, 2014) found a reduction in the GPG over time, while (Blau

[1] That is, the OB decomposition and ex-post comparison of the decomposition results.

and Kahn, 2017), for example, found that the unexplained component remained comparatively stable. Differences in pay have also been revealed across sectors, particularly between the public and private sectors, as the PPWG differs for men and women (Arulampalam *et al.*, 2007; Lucifora and Meurs, 2006; Melly, 2005). Women generally prefer the public sector because of its fairer recruitment, selection criteria, and remuneration schemes and better implementation of anti-discrimination laws (Gornick and Jacobs, 1998; Grimshaw, 2000; Castagnetti and Giorgetti, 2018). As a result, the difference in pay by gender is generally smaller in the public sector compared to the private sector. However, regardless of sector, wages across the sector differ based on gender (Lucifora and Meurs, 2006).

Despite its popularity, the OB decomposition has several drawbacks,[2] the frequently most cited of which is the so-called index number problem. Solutions in the literature consist of estimating a pooled wage structure (Neumark, 1988; Oaxaca and Ransom, 1994) and assigning different weights to the two groups (Reimers, 1983; Cotton, 1988). The intercept-shift approach (see Fortin, 2008; Elder *et al.*, 2010; Magnani and Zhu, 2012, for examples) generalizes the approach of Neumark (1988) and Oaxaca and Ransom (1994), which allow for different intercepts in the pooled sample. Fortin (2008) re-wrote the decomposition of the GPG in terms of advantages for men and disadvantages for women by including the group indicator and parameter restrictions, so the decomposition no longer depends on the choice of the non-discriminatory wage structure.

However, as Lee (2015) stressed, Fortin (2008)'s intercept-shift approach set the reference parameter for the OB decomposition (i.e. the parameter that would prevail in a world with no discrimination) based on the variance in the difference among categories and not on the level of difference. The decomposition should rely on the level of difference (see Lee, 2015, for further details), as level and variance differences may not go in the same direction. For example, women may have a larger variance in average labour market experience than men do, while men may have larger levels of labour market experience than women do. Moreover, the reference intercept in Fortin (2008) is arbitrary, as the same OB decomposition holds with different reference intercepts.

[2]Blinder (1973) and Oaxaca (1973) were together cited more than 15,000 times as of 7 February, 2019, according to the Google Scholar citation statistics.

A second problem of the OB decomposition is known as the omitted group problem. In the case of categorical variables, the decomposition depends on the choice of the omitted group (see e.g. Jones, 1983; Oaxaca and Ransom, 1994), but this problem can be solved easily by using deviation contrast transforms of the categorical variables (Gardeazabal and Ugidos, 2004; Yun, 2005).

We show that our decomposition overcomes both the index number problem and the omitted group problem and does not suffer from arbitrary reference intercepts but relies on the level of difference. Like the OB approach, our method can provide a detailed decomposition. Further, using linear unconditional quantile regressions (Firpo *et al.*, 2009), our decomposition can be extended beyond the mean, so changes in inequality, measured by inter-quantile GPGs, can be derived easily.

The remainder of the paper is organized as follows. The next section provides an overview of the literature concerning the methods used to estimate changes in gaps over time. Section 3 outlines our proposed decomposition and Section 5 presents the extension proposed for solving the index number problem. Section 6 describes the data set, along with some descriptive statistics. Section 7 presents the empirical results of both the standard OB and our proposed decomposition. Finally, Section 8 concludes.

2 Changes in the Gender Pay Gap: Methods in Use in the Literature

The OB decomposition is the workhorse in empirical labour economics when it comes to decomposition methods (Fortin *et al.*, 2011). The approach allows researchers to study labour market outcomes between groups (e.g. gender, race) by decomposing differences in log wages using linear regression models.

Ordinary least squares (OLS) is used in the case of the mean , while linear probability models are applied (RIF-OLS) to quantiles. The idea is to construct counterfactuals so as to attribute one part of the gap to an explained part of the gap and one part to an unexplained part. The OB decomposition estimates Mincer-type wage equations separately for the two groups and then decomposes the wage differential into two components: endowments (explained) and coefficients (unexplained). Using men as the reference category, the decomposition takes the following form:

$$\begin{aligned}\bar{y}_M - \bar{y}_F &= \bar{x}_M^{'}\hat{\beta}_M - \bar{x}_F^{'}\hat{\beta}_F \\ &= (\bar{x}_M^{'} - \bar{x}_F^{'})\hat{\beta}_M + \bar{x}_F^{'}(\hat{\beta}_M - \hat{\beta}_F)\end{aligned} \tag{1}$$

where $\bar{y}_G$ is the dependent variable (e.g. the log of hourly wages) of group $G = (M, F)$ evaluated at the mean and $\bar{x}_G$ and $\hat{\beta}_G$ are $K \times 1$ vectors of average characteristics and estimated coefficients for group G, respectively.[3] The first term is the effect that is due to differences in observable characteristics, such as education and work experience. As different observed characteristics are expected to have different effects on earnings, the difference in observed characteristics is also referred to as the explained component, endowments, or quantity effect. The second term is the effect that is due to differences in returns on observable wage characteristics. This component is generally referred to as the unexplained part, coefficients, or price effect of the GPG. The unexplained part is often used as a measure of discrimination, but it incorporates the effect of group differences in unobserved predictors.[4]

The presence and degree of discrimination has been a controversial issue in the literature in large part because the wage equation cannot include all relevant variables because skills and individual productivity cannot be observed. Therefore, observationally equivalent people based on the characteristics in the wage equation may not be truly equivalent (omitted variable problem). As for omitted controls, the OB decomposition over-estimates the degree of discrimination, as the price effect is now the sum of discrimination and differences in unobserved characteristics.

To examine changes in wage gaps over time (or between groups/sectors), one may use a double OB decomposition, as proposed in Smith and Welch (1989):

[3] Alternative specifications consider different reference groups (e.g. female or pooled coefficient estimates; Oaxaca, 1973; Oaxaca and Ransom, 1994).

[4] The unexplained portion of the GPG may include effects of unobserved characteristics like individual productivity, motivation, and educational quality (Blau and Kahn, 2006).

$$
\begin{aligned}
\Delta(\bar{y}_{MT} - \bar{y}_{FT}) &= [(\bar{x}'_{FT} - \bar{x}'_{MT}) - (\bar{x}'_{Ft} - \bar{x}'_{Mt})]\hat{\beta}_{tM} \\
&+ (\bar{x}'_{FT} - \bar{x}'_{Ft})(\hat{\beta}_{tF} - \hat{\beta}_{tM}) \\
&+ (\bar{x}'_{FT} - \bar{x}'_{MT})'(\hat{\beta}_{TM} - \hat{\beta}_{tM}) \\
&+ \bar{x}'_{FT}[(\hat{\beta}_{TF} - \hat{\beta}_{TM}) - (\hat{\beta}_{tF} - \hat{\beta}_{tM})] \qquad (2)
\end{aligned}
$$

where the subscripts T and t refer to the ending period (current year) and the starting period (base year), respectively, and $\Delta(\bar{y}_{MT} - \bar{y}_{FT}) = (\bar{y}_{MT} - \bar{y}_{FT}) - (\bar{y}_{Mt} - \bar{y}_{Ft})$.

The first term in (2) measures the predicted change in group $(M - F)$ wages that occurs because of differences in observed characteristics over time $(T - t)$ that are valued at base-year group M parameter values. The second term measures group interactions. If individuals in group F are paid less than those in group M for a given characteristic, we have $(\hat{\beta}_{tF} - \hat{\beta}_{tM}) < 0$. Individuals in group F will lose relative to group M when average sets of endowments increase over time and based on gender. The third term measures year interactions, and the fourth term measures group-year interactions. The decomposition in (2) is conducted by comparing parameter estimates on different samples and periods, as inferences cannot be drawn on the single components of the decomposition.

Juhn *et al.* (1991) (JMP) proposed a decomposition equation for changes in wage differentials that considers the effect of unobserved skills on the GPG. Evaluated at the sample mean, the wage equation for group $G = (M, F)$ may be written as:

$$\bar{y}_G = \bar{x}'_G\hat{\beta} + \hat{\sigma}\hat{\theta}_G \qquad (3)$$

where $\hat{\beta}$ is the estimated parameter vector from a pooled wage regression,[5] $\hat{\theta}_G$ is the mean

[5]The JMP decomposition considers the estimation of only the non-discriminated group, assuming that the discriminated group is affected by the same economic forces that influence the non-discriminated group' wage distribution. Thus, the estimated prices of measured characteristics are assumed to affect both groups in the same way, and the residuals are decomposed into a portion that reflects the prices of unmeasured skills and a portion that reflects the quantities of unmeasured skills, with the former affecting both groups similarly (Yun, 2009). Therefore, the JMP decomposition relies on two strong assumptions: that one group's OLS estimations are unbiased and the other group's are biased, and that the level of discrimination is constant over time. Because of these issues and to address the index number problem, we present the JMP decomposition for the pooled wage regression.

standardized residual in group G, and $\hat{\sigma}$ is the mean standard error estimate (i.e. the average percentile rank). $\hat{\beta}$ represents an estimate of the vector of observed prices, and $\hat{\sigma}$ is an estimate of wage dispersion (which is often interpreted as an estimate of unobserved prices (see Blau and Kahn, 1997 and Gupta *et al.*, 2006)), and $\hat{\theta}_G$ represents some measure of generally (unobserved) labour market ability. Given (3), the wage differential between group M and F is:

$$\bar{y}_M - \bar{y}_F = (\bar{x}^{'}_M - \bar{x}^{'}_F)\hat{\beta} + (\hat{\theta}_M - \hat{\theta}_F)\hat{\sigma} \tag{4}$$

where the first component represents the explained part of the wage gap, the predicted gap, and the second component represents the residual or unexplained part of the wage gap, the residual gap. From (4), the change in the wage gap between years T and t is:

$$\begin{aligned} \Delta(\bar{y}_{MT} - \bar{y}_{FT}) &= \hat{\beta}_T \Delta(\bar{x}^{'}_{MT} - \bar{x}^{'}_{FT}) + (\bar{x}^{'}_{MT} - \bar{x}^{'}_{FT})\Delta\hat{\beta}_T + \\ &\quad \hat{\sigma}_T \Delta(\hat{\theta}_{FT} - \hat{\theta_{MT}}) + (\hat{\theta}_{Ft} - \hat{\theta}_{Mt})\Delta\hat{\sigma}_T \end{aligned} \tag{5}$$

where the first term represents the difference in mean endowments, the second term represents differences in returns to endowments, and the last two terms correspond to the change in the residual wage gap. In particular, the first term of the residual wage gap, termed the ranking effect, can be split between the effect of group F's movements in the wage distribution after adjusting for changes in human capital characteristics, $\hat{\sigma}_T \Delta(\hat{\theta}_{FT})$, and the effect of movements of group M in the wage distribution at time T after controlling for changes in the characteristics of human capital, the term $-\hat{\sigma}_T \Delta(\hat{\theta}_{MT})$. The last term is interpreted as the dispersion or unobserved prices effect (see Gupta *et al.* 2006). However, unlike OB-type decompositions of wage differentials, the JMP method provides coefficients and characteristics effects only at an aggregate level. Because of this shortcoming, the JMP method cannot be used for a detailed decomposition of the variation in the GPG over time. More important, as Suen (1997) stressed, the JMP decomposition of wage residuals into standard deviations (the price of unobserved skills) and percentile ranks (the level of unobserved skills) is unbiased only when the two measures are independent. Moreover, Juhn *et al.* (1991) and Juhn *et al.* (1993) did not derive the statistical distribution of the decomposition components. Inferences about the components can be made using the approach in Gupta *et al.* (2006), where the standard errors are derived. However, the standard errors are derived under two strong assumptions: that the standard deviation and percentile ranks are independent, and that the covariance between estimators in different time periods is approximately zero.

3 Proposed Decomposition

The method we propose starts from the OB decomposition Gelbach (2016) proposed, which divides cross-specification differences in OLS estimates of the female coefficient in a path-independent way. Following Gelbach (2016), we rely on the omitted variable bias (OVB) formula to estimate the decomposition consistently, conditional on all covariates. As in the standard OB framework, sequencing problems do not occur when the OVB formula is used for decompositions[6]

The Role of the Intercepts

In the decomposition proposed, the intercept terms, $(\hat{\alpha}_M - \hat{\alpha}_F)$, play an important role. The group difference in the intercepts is generally attributed to the second term in (1) and is referred to as the group difference in starting points. Blinder (1973) called this part the unexplained part of discrimination, as interpretation of the difference in the intercepts may not be straightforward. The intercept coefficients are influenced by the reference group(s) used for the indicator variable(s), and the intercept is influenced by the choice of scale for continuous variables in the model. According to Jones (1983), the problem is critical, and interpretation of the intercept is arbitrary, so the intercept term is uninterpretable.

Relying on the OVB formula's sequential decomposition of the wage gap, we propose a different interpretation of the difference in the intercepts. Consider the following linear model for the wage regression on the sample composed of both groups of interest, males and females:

$$y = \iota\alpha + F\alpha_1 + X\beta_1 + XF\beta_2 + \epsilon_1 \tag{6}$$

[6] When one starts from a base specification and adds regressors sequentially, the order of addition influences the coefficient estimates.

where ι is a vector of ones, F is a vector of a dummy variable that equals 1 if the individual is female and zero otherwise, X is the matrix of regressors, XF is the interaction effect, and ϵ_1 is the vector of error terms. Using Gelbach (2016)'s terminology, the specification in (6) represents the full model. The OLS estimate of $\hat{\alpha}_1$ is equal to: $\hat{\beta}_F - \hat{\beta}_M$ where

$$y = \iota\beta_F + X\beta_{1F} + \epsilon_F \text{ for females}$$

and

$$y = \iota\beta_M + X\beta_{1M} + \epsilon_M \text{ for males}$$

are the two wage regressions for the female group and the male group, respectively. The difference in the average observed y between the two groups (i.e. $\bar{y}_F - \bar{y}_M$) is given by $\hat{\gamma}_F$ in the following regression:

$$y = \iota\gamma + F\gamma_F + \epsilon_2 \text{ for the whole sample} \tag{7}$$

where F is a vector of a dummy variable that equals 1 for female and zero otherwise. The model in (7) represents the base model. The difference between the base and full model reads as:

$$\hat{\gamma}_F - \hat{\alpha}_1 = \alpha_1^{base} - \alpha_1^{full}$$

which represents the part of the gap that is explained by the regressors (X, XF) and which can be decomposed as the sum of two components by means of the OVB formula, as shown in the next section. If α_1^{full} were equal to zero, the model would explain all of the observed GPG. α_1^{full} is the part of the GPG that cannot be explained by the quantity and the price effect, so instead of attributing the difference in the intercepts to the price component without a clear interpretation of its source, we focus the analysis on the components that can be attributed to either part of the decomposition-that is, to differences in endowments (the explained component) or to part of the differences in remuneration (the unexplained component).

Changes in Wages over Time

We present the derivation of the method in the case of the GPG over time. Using our proposed approach, we can (directly) draw inferences about the differences in the GPG across years and investigate the main contributors to the change in the GPG over time. Using the aggregate decomposition, we can distinguish between the explained and unexplained component and then use the detailed approach to attribute the change to gender differences in educational attainment, labour market presence, and occupational or sectoral sorting. The method proposed can be applied to various cases of group differences in outcome variables over time, sectors, countries, and so on. To estimate the wage equation separately by G (gender) and J (year), we use:

$$y_{GJ} = \iota\alpha_{GJ} + X_{GJ}\beta_{GJ} + \epsilon_{GJ} \tag{8}$$

with $G = F, M$ (for $F =$ female and $M =$ male), $J = t, T$ (for $t =$ starting period and $T =$ ending period), and where y_{GJ} is the $N \times 1$ vector of logarithmic wages of G in year J, α_{GJ} is the intercept, ι is the $N \times 1$ vector of constants, and X is a $N \times K$ matrix of exogenous regressors. β_{FJ} is the corresponding $K \times 1$ vector of coefficients, and ϵ_{FJ} is a $N \times 1$ vector containing the error terms. The estimation provides four sets of parameter estimates of the same dimension, given the assumption that the set of regressors is the same for the four combinations considered. Evaluating the estimation at the mean, given the property that OLS estimates must go through the mean of the data, equation (8) becomes:

$$\widehat{\bar{y}}_{GJ} = \bar{y}_{GJ} = \hat{\alpha}_{GJ} + \bar{x}'_{GJ}\hat{\beta}_{GJ} \tag{9}$$

where $\hat{\alpha}_{GJ}$ is the intercept estimate, $\bar{x}_{GJ}$ is the $K \times 1$ column vector of sample means of observable characteristics in X:

$$\bar{x}'_{GJ} = \left[\bar{x}_{1,GJ}, \bar{x}_{2,GJ}, \ldots, \bar{x}_{K,GJ}\right]$$

and $\hat{\beta}_{GJ}$ is the corresponding $K \times 1$ vector of parameter estimates.

To estimate the joint model as in Gelbach (2016), we distinguish between two sets of regressors, X_1 and X_2, where X_1, represents the regressors of the base specification, which contains only a constant, an interaction term between the gender and year dummies, and the group and time dummies themselves:

$$X_1 = \Big[1, FJ, F, J\Big]$$

where

$$F = \begin{cases} 1 & \text{if } female \\ 0 & \text{if } male \end{cases} \qquad J = \begin{cases} 1 & \text{if } year = t \\ 0 & \text{if } year = T \end{cases}$$

The base model is defined as follows:

$$y = \alpha_0^{base} + FJ\alpha_1^{base} + F\alpha_2^{base} + J\alpha_3^{base} + \epsilon^{base} \tag{10}$$

The second set of regressors, X_2, of dimension $(N \times 4K)$, contains the matrix of characteristics X and the interactions of the gender and year dummies with X:

$$X_2 = [X, FX, JX, FJX] \tag{11}$$

where FX and JX are the interaction variables between the regressors X, that is, the female dummy F and the starting period dummy J, respectively. Thus, FJX represents the interaction variable among regressors X. The full model is then defined as follows:

$$y = \alpha_0^{full} + FJ\alpha_1^{full} + F\alpha_2^{full} + J\alpha_3^{full} + X\beta_1 + FX\beta_2 + JX\beta_3 + FJX\beta_4 + \epsilon^{full} \tag{12}$$

The link between the parameters of the full model and the four equations represented in (8) follows straightforwardly.[7]

Next, we consider the set of regressors X_2 as omitted variables. Using the OVB formula, we

[7] Appendix A reports the relationship between the two sets of estimation results.

have:

$$\hat{\alpha}^{base} = \hat{\alpha}^{full} + (X_1'X_1)^{-1}X_1'X_2\hat{\beta}^{full} \tag{13}$$

where the vector of parameter estimates from the base model (10) is:

$$\hat{\alpha}^{base'} = (\hat{\alpha}_0^{base}\ \hat{\alpha}_1^{base}\ \hat{\alpha}_2^{base}\ \hat{\alpha}_3^{base}) \tag{14}$$

and $\hat{\alpha}^{full}$ is the 4×1 vector that contains the coefficient estimates of X_1 from the full model (12): $(X_1'X_1)^{-1}X_1'X_2$ is the linear projection of X_2 on X_1 and

$$\hat{\beta}^{full'} = (\hat{\beta}_1\ \hat{\beta}_2\ \hat{\beta}_3\ \hat{\beta}_4) \tag{15}$$

is the $(1 \times 4K)$ vector of coefficients from the full model (12). Model (13) can be decomposed as follows:

$$\hat{\alpha}^{base} = \hat{\alpha}^{full} + \hat{\delta} \tag{16}$$

where $\hat{\delta} \equiv \hat{\alpha}^{base} - \hat{\alpha}^{full} = (X_1'X_1)^{-1}\, X_1'X_2\hat{\beta}^{full}$ and

$$\hat{\delta} = \hat{\delta}_X + \hat{\delta}_{FX} + \hat{\delta}_{JX} + \hat{\delta}_{FJX} \tag{17}$$

where $\hat{\delta}_S = \hat{\Gamma}^S\hat{\beta}_S^{full}$, with $\hat{\Gamma}^S = (X_1'X_1)^{-1}\, X_1'S$ of dimension $(4 \times K)$ and S is the portion of the matrix (11) that corresponds to the set of regressors S, for $S = X, \ldots, FJX$ in (11).[8]

The Decomposition

Our interest relies in the estimation and decomposition of the GPG across two periods, t and T:

$$\Delta_T - \Delta_t = \left(\bar{y}_{MT} - \bar{y}_{FT}\right) - \left(\bar{y}_{MT} - \bar{y}_{Ft}\right)$$

[8] Accordingly, $\hat{\delta}_X = \hat{\Gamma}^X\hat{\beta}_X^{full}$, with $\hat{\Gamma}^X = (X_1'X_1)^{-1}\, X_1'X$ of dimension $(4 \times K)$ and $\hat{\beta}_X^{full}$ is the $(K \times 1)$ vector $\hat{\beta}_1$ in (15).

with Δ_T being the GPG in T. Thus, it is easy to show that:

$$\Delta_T = \left(\bar{y}_{MT} - \bar{y}_{FT}\right)$$
$$= -\hat{\alpha}_2^{base}$$

and

$$\Delta_t = \left(\bar{y}_{Mt} - \bar{y}_{Ft}\right)$$
$$= -\hat{\alpha}_1^{base} - \hat{\alpha}_2^{base}$$

Therefore:

$$\Delta_T - \Delta_t = \hat{\alpha}_1^{base}$$

Therefore, given (14), we are interested in the second row of $\hat{\alpha}^{base}$, i.e. $\hat{\alpha}_1^{base}$, in order to obtain the differences of the respective wage gaps; $\Delta_T - \Delta_t$. Following (13) and (16)-(17), we decompose $\hat{\alpha}_1^{base}$ accordingly. In particular, in decomposing $\hat{\alpha}_1^{base}$ we refer to the second row of $(X_1'X_1)^{-1}X_1'X_2$:

$$(X_1'X_1)^{-1}X_1'X = \begin{bmatrix} / \\ (\bar{x}'_{MT} - \bar{x}'_{FT}) - (\bar{x}'_{Mt} - \bar{x}'_{Ft}) \\ / \\ / \end{bmatrix}$$

$$(X_1'X_1)^{-1}X_1'FX = \begin{bmatrix} / \\ (\bar{x}'_{Ft} - \bar{x}'_{FT}) \\ / \\ / \end{bmatrix}$$

$$(X_1'X_1)^{-1}X_1'JX = \begin{bmatrix} / \\ (\bar{x}'_{Ft} - \bar{x}'_{Mt}) \\ / \\ / \end{bmatrix}$$

and

$$(X_1'X_1)^{-1}X_1'FJX = \begin{bmatrix} / \\ \bar{x}'_{Ft} \\ / \\ / \end{bmatrix}$$

Thus, the second row of equation (13) (i.e. the change in the wage gap evaluated at the mean) is:

$$\begin{aligned} \hat{\alpha_1}^{base} &= \underbrace{(\hat{\alpha}_{MT} - \hat{\alpha}_{FT}) - (\hat{\alpha}_{Mt} - \hat{\alpha}_{Ft})}_{\hat{\alpha}_1^{full}} + [(\bar{x}'_{MT} - \bar{x}'_{FT}) - (\bar{x}'_{Mt} - \bar{x}'_{Ft})]\underbrace{\hat{\beta}_{MT}}_{\hat{\beta}_1} \\ &+ (\bar{x}'_{Ft} - \bar{x}'_{FT})\underbrace{(\hat{\beta}_{FT} - \hat{\beta}_{MT})}_{\hat{\beta}_2} \\ &+ (\bar{x}'_{Ft} - \bar{x}'_{Mt})\underbrace{(\hat{\beta}_{Mt} - \hat{\beta}_{MT})}_{\hat{\beta}_3} \\ &+ \bar{x}'_{Ft}\underbrace{[(\hat{\beta}_{MT} - \hat{\beta}_{FT}) - (\hat{\beta}_{Mt} - \hat{\beta}_{Ft})]}_{\hat{\beta}_4} \\ &= \Delta_T - \Delta_t \end{aligned}$$

or, equivalently:

$$\hat{\alpha_1}^{base} = \hat{\alpha_1}^{full} + \hat{\delta}_1 + \hat{\delta}_2 + \hat{\delta}_3 + \hat{\delta}_4$$

where:

$$\begin{aligned}
\hat{\delta}_1 &= (\bar{x}_{Mt} - \bar{x}_{Ft})\hat{\beta}_{Mt} - (\bar{x}_{MT} - \bar{x}_{FT})\hat{\beta}_{Mt} \\
\hat{\delta}_2 &= \bar{x}_{Ft}(\hat{\beta}_{Mt} - \hat{\beta}_{Ft}) + \bar{x}_{FT}(\hat{\beta}_{Ft} - \hat{\beta}_{Mt}) \\
\hat{\delta}_3 &= (\bar{x}_{FT} - \bar{x}_{MT})\hat{\beta}_{MT} - (\bar{x}_{FT} - \bar{x}_{MT})\hat{\beta}_{Mt} \\
\hat{\delta}_4 &= \bar{x}_{FT}(\hat{\beta}_{Mt} - \hat{\beta}_{Ft}) - \bar{x}_{FT}(\hat{\beta}_{MT} - \hat{\beta}_{FT})
\end{aligned}$$

which can be re-written as a double OB decomposition:

$$\hat{\alpha_1}^{base} - \hat{\alpha_1}^{full} = \underbrace{(\hat{Q}_t + K)}_{\hat{\delta}_1} + \underbrace{(\hat{P}_t + W)}_{\hat{\delta}_2} + \underbrace{(-\hat{Q}_T - K)}_{\hat{\delta}_3} + \underbrace{(-\hat{P}_T - W)}_{\hat{\delta}_4}$$

where $\hat{Q}_t = (\bar{x}_{Mt} - \bar{x}_{Ft})\hat{\beta}_{Mt}$, is the estimated quantity effect and $\hat{P}_t = \bar{x}_{Ft}(\hat{\beta}_{Mt} - \hat{\beta}_{Ft})$, the estimated price effect in period t, and $\hat{Q}_T = (\bar{x}_{MT} - \bar{x}_{FT})\hat{\beta}_{MT}$, and $\hat{P}_T = \bar{x}_{FT}(\hat{\beta}_{MT} - \hat{\beta}_{FT})$, the estimated quantity and price effect in period T, respectively.

The proposed decomposition approach can be estimated beyond the mean by using the linear unconditional quantile regression (RIF-OLS) introduced by Firpo *et al.* (2009). Instead of using y as a dependent variable, a nonlinear transformation of y, we use the recentered influence function (RIF) of y at the unconditional quantile Q_τ; $RIF(y; Q_\tau)$.[9]

4 Inference

The asymptotic distribution of $\sqrt{N}(\hat{\delta} - \delta)$, with $\hat{\delta} = (\hat{\delta}_1, \ldots, \hat{\delta}_4)$ derived in Gelbach (2016), is summarized in Appendix B. Given the distribution of the parameters $\hat{\delta}$, the proposed decomposition allows inferences about the dynamic of the single components of the GPG to be carried out.

If the interest relies on, for instance, investigating the convergence of the GPG, whether the convergence can be explained by the convergence of the level of endowments (explained or quantity components) or by changes in prices (unexplained or price components) can be determined. The hypothesis that the convergence is driven by changes in observed characteristics

[9]For additional details on the RIF-OLS approach, see Appendix C.

can be tested by:

$$H_0 : \delta_1 + \delta_3 = 0$$

, which is equivalent to testing for the H_0 that there was no change in endowments between M and F over time:

$$H_0 : Q_t = Q_T$$

Changes in the remuneration scheme between men and women can be controlled for by testing whether the components of the price effects have been stable over time:

$$H_0 : \delta_2 + \delta_4 = 0$$

which is equivalent to testing:

$$H_0 : P_t = P_T$$

where H_0 indicates no change in prices between M and F over time. Each $\hat{\delta}$ can be decomposed into its single components; for instance, the contribution to the GPG of labour market experience to the quantity component in period t can be extracted from $\hat{\delta}_1$. Each $\hat{\delta}$ is given by

$$\hat{\delta}_i = \sum_{k=1}^{K} \hat{\delta}_{ik} = \sum_{k=1}^{K} \hat{\Gamma}_{ik} \hat{\beta}_k \text{ for i=1, ..., 4} \tag{18}$$

where K are the regressors considered in the analysis. (18) shows how the detailed decomposition can be managed.

5 The Index Number Problem

As is well known in the literature, the OB decomposition is not unique, so the choosing a non-discriminatory wage structure leads to different results (Oaxaca and Ransom, 1994; Cotton, 1988; Fortin *et al.*, 2011). Several approaches have been proposed to circumvent this problem (Reimers, 1983; Cotton, 1988; Neumark, 1988; Oaxaca and Ransom, 1994; Fortin, 2008).

We propose an extension of our method that provides a wage decomposition that is invariant to the reference category adopted. Considering the standard case of the GPG, Fortin (2008) included gender intercept shifts along with an identification restriction in the regression of females and males pooled together:

$$y_i = \gamma_0 + \gamma_{0F} F_i + \gamma_{0M} M_i + X_i \gamma + \epsilon_i$$

subject to:

$$\gamma_{0F} + \gamma_{0M} = 0$$

where F_i (M_i) is equal to one if the individual is female (male), and zero otherwise. The identification restriction, $\gamma_{0F} + \gamma_{0M} = 0$, requires that the pooled wage equation represents a non-discriminatory wage structure, that is, a wage structure in which the advantage of men is equal to the disadvantage of women. The first component on the RHS, $(\bar{X}_M - \bar{X}_F)\hat{\gamma}$, is the explained part, while $\hat{\gamma}_{0M}$ and $\hat{\gamma}_{0F}$ are the advantage of men and the disadvantage of women, respectively. In particular:

$$\bar{y}_M - \bar{y}_F = (\bar{X}_M - \bar{X}_F)\hat{\gamma} + (\hat{\gamma}_{0M} - \hat{\gamma}_{0F}) \tag{19}$$

where $\hat{\alpha}_M, \hat{\alpha}_F, \hat{\beta}_M$, and $\hat{\beta}_F$ are the estimated coefficients of the wage equations for men and women, respectively :

$$\begin{aligned} y_{iM} &= \alpha_M + X_M \beta_M + \epsilon_{iM} && (20) \\ y_{iF} &= \alpha_F + X_F \beta_F + \epsilon_{iF} && (21) \end{aligned}$$

Solution for Differences in Wages over Time

The extension of the decomposition described above to changes in wage gaps and over time, sector, country, and so on follows straightforwardly. The set of regressors considered in Section 3 becomes:

$$
\begin{aligned}
X_1 &= [1, (F-M)J, (F-M), J] \\
X_2 &= [X, (F-M)X, JX, (F-M)JX]
\end{aligned}
$$

Then base model reads as:

$$
y_i = \gamma_0^{base} + (F_i - M_i)J_i\gamma_{FJ}^{base} + (F_i - M_i)\gamma_F^{base} + J_i\gamma_J^{base} + \epsilon_i^{base} \tag{22}
$$

while the full model is defined as:

$$
\begin{aligned}
y_i &= \gamma_0^{full} + (F_i - M_i)J_i\gamma_{FJ}^{full} + (F_i - M_i)\gamma_F^{full} + J_i\gamma_J^{full} + \\
&+ X_i\gamma + (F_i - M_i)X_i\gamma_{XF} + J_iX_i\gamma_{XJ} + (F_i - M_i)J_iX_i\gamma_{XJF} + \epsilon_i^{full}
\end{aligned} \tag{23}
$$

$(\gamma_0^{base}\ \gamma_{FJ}^{base}\ \gamma_F^{base}\ \gamma_J^{base})$ is the vector of coefficients estimates of X_1 from the base model (22), and $(\gamma_0^{full}\ \gamma_{FJ}^{full}\ \gamma_F^{full}\ \gamma_J^{full})$ is the vector containing the coefficient estimates of X_1 from the full model (23), while $(\gamma\ \gamma_{XF}\ \gamma_{XJ}\ \gamma_{XJF})$ is the vector of coefficients estimates of X_2 from the full model (23). The linear projection of X with respect to X_1 is equal to:

$$
(X_1'X_1)^{-1}X_1'X = \begin{bmatrix} / \\ -[(\bar{x}_{Mt} - \bar{x}_{Ft}) - (\bar{x}_{MT} - \bar{x}_{FT})]/2 \\ / \\ / \end{bmatrix}
$$

The linear projection of $(F-M)X$ with respect to X_1 is equal to:

$$
(X_1'X_1)^{-1}X_1'(F-M)X = \begin{bmatrix} / \\ [(\bar{x}_{Mt} + \bar{x}_{Ft}) - (\bar{x}_{MT} + \bar{x}_{FT})]/2 \\ / \\ / \end{bmatrix}
$$

The linear projection of JX with respect to X_1 is equal to:

$$(X_1'X_1)^{-1}X_1'JX = \begin{bmatrix} / \\ (\bar{x}_{Ft} - \bar{x}_{Mt})/2 \\ / \\ / \end{bmatrix}$$

and the linear projection of $(F - M)JX$ with respect to X_1 is equal to:

$$(X_1'X_1)^{-1}X_1'(F - M)JX = \begin{bmatrix} / \\ (\bar{x}_{Ft} + \bar{x}_{Mt})/2 \\ / \\ / \end{bmatrix}$$

It can be easily shown that:

$$\begin{aligned} \hat{\gamma}_{FJ}^{base} &= \frac{\Big(\bar{y}_{MT} - \bar{y}_{FT}\Big) - \Big(\bar{y}_{Mt} - \bar{y}_{Ft}\Big)}{2} \\ &= \frac{\Delta GPG}{2} \end{aligned}$$

and

$$\hat{\gamma}_{FJ}^{full} = \frac{(\hat{\alpha}_{MT} - \hat{\alpha}_{FT}) - (\hat{\alpha}_{Mt} - \hat{\alpha}_{Ft})}{2}$$

where, analogously to Section 3, $\hat{\alpha}_{Mt}, \hat{\alpha}_{MT}, \hat{\alpha}_{Ft}, \hat{\alpha}_{FT}$ are the estimated coefficients of the wage equations for men and women in period t and T, respectively:

$$y_{iMt} = \alpha_{Mt} + X_{Mt}\beta_{Mt} + \epsilon_{iMt} \tag{24}$$

$$y_{iFt} = \alpha_{Ft} + X_{Ft}\beta_{Ft} + \epsilon_{iFt} \tag{25}$$

$$y_{iMT} = \alpha_{MT} + X_{MT}\beta_{MT} + \epsilon_{iMT} \tag{26}$$

$$y_{iFT} = \alpha_{FT} + X_{FT}\beta_{FT} + \epsilon_{iFT} \tag{27}$$

Hence, the relationship:

$$
\begin{aligned}
\hat{\gamma}_{FJ}^{base} &= \hat{\gamma}_{FJ}^{full} + (X_1^{'}X_1)^{-1}X_1^{'}X\hat{\gamma} + (X_1^{'}X_1)^{-1}X_1^{'}X(F-M)\hat{\gamma}_{XF} + \\
&+ (X_1^{'}X_1)^{-1}X_1^{'}XJ\hat{\gamma}_{XY} + (X_1^{'}X_1)^{-1}X_1^{'}XJ(F-M)\hat{\gamma}_{XJF}
\end{aligned}
$$

can be re-written in terms of the change in the pay gap between M and F over time, ΔGPG (for the case of the GPG over time), as:

$$
\begin{aligned}
2\hat{\gamma}_{FJ}^{base} &= \Delta GPG = \\
&= [(\hat{\alpha}_{MT} - \hat{\alpha}_{FT}) - (\hat{\alpha}_{Mt} - \hat{\alpha}_{Ft})] + (\Delta\bar{x}_T - \Delta\bar{x}_t)\hat{\gamma} + \\
&\quad +(\sum \bar{x}_t - \sum \bar{x}_T)\hat{\gamma}_{XF} - \Delta\bar{x}_t\hat{\gamma}_{XY} + \sum \bar{x}_t\hat{\gamma}_{XYF}
\end{aligned}
$$

where $\Delta\bar{x}_{Year}$ is the difference between the average level of observed characteristics of men and women in a certain year, with $Year = t, T$ and $\sum \bar{x}_{Year}$ representing the sum of observable labour market characteristics present for men and women in year$=$$Year$. The model can be re-written in terms of the OVB formula as follows:

$$
\begin{aligned}
&2\hat{\gamma}_{FJ}^{base} = \hat{\gamma}_{FJ}^{full} + \hat{\delta}_A + \hat{\delta}_B + \hat{\delta}_C + \hat{\delta}_D \\
&\hat{P} + \hat{Q} = \hat{\delta}_A + \hat{\delta}_B + \hat{\delta}_C + \hat{\delta}_D
\end{aligned}
$$

with P accounting for the price effect and Q for the quantity effect. In particular,

$$
\begin{aligned}
\hat{\delta}_A &= \underbrace{(\bar{x}_{MT} - \bar{x}_{FT})\hat{\gamma}}_{Q_T} - \underbrace{(\bar{x}_{Mt} - \bar{x}_{Ft})\hat{\gamma}}_{W} \\
\hat{\delta}_B &= \underbrace{\bar{x}_{MT}(\hat{\beta}_{MT} - \hat{\gamma}) - \bar{x}_{FT}(\hat{\beta}_{FT} - \hat{\gamma})}_{P_T} + \underbrace{(\bar{x}_{Mt} + \bar{x}_{Ft})\hat{\gamma}_{XF}}_{K} \\
\hat{\delta}_C &= \underbrace{(\bar{x}_{Mt} - \bar{x}_{Ft})(\hat{\gamma} + \hat{\gamma}_{XJF})}_{Q_t} - \underbrace{(\bar{x}_{Mt} - \bar{x}_{Ft})\hat{\gamma}}_{W} \\
\hat{\delta}_D &= -\underbrace{[\bar{x}_{Mt}(\hat{\beta}_{Mt} - \hat{\gamma} - \hat{\gamma}_{XJ}) - \bar{x}_{Ft}(\hat{\beta}_{Ft} - \hat{\gamma} - \hat{\gamma}_{XY})]}_{P_t} - \underbrace{(\bar{x}_{Mt} + \bar{x}_{Ft})\hat{\gamma}_{XF}}_{K}
\end{aligned}
$$

The invariant decomposition for the standard case of a wage differential like the GPG is shown in Appendix D, while a discussion of the intercept-shift and pooled sample approach can be found in Appendix E. In particular, Appendix E shows that our invariant decomposition for the variance over time does not suffer from arbitrary intercepts.

6 Data and Descriptive Statistics

We use the 2016 and 2005 results of the survey ISFOL Plus from the Italian Institute for the Development of Vocational Training for Workers (ISFOL) to analyse the GPG's evolution over time (pooling the two cross-sections of 2005 and 2016) and the latest release, the cross-section of 2016, to analyse the PPWG between men and women. Our analysis focuses on full-time employees between eighteen and sixty-four years of age who work more than thirty-five hours per week. The analysis is constrained to earnings from the employee's job that yields the highest income. These selection criteria led to a sample size of 9,185 in 2005 and 10,148 in 2016. The 2005 sample contains 3,983 women (43%) and 5,202 men (57%), while the 2016 sample contains 4,359 (43%) women and 5,789 (57%) men. In 2016, 2,084 women (48% of women in the sample) and 2,033 men (35% of the men in the sample) were employed in the public sector, suggesting that women favour the public sector, which is in line with results in the literature (see Section 1) because of its more egalitarian pay schemes.

Table 1 and 2 report the means and standard deviations for the human capital variables included in the analysis for the two cases under consideration, respectively . Table 1 shows that men's average earnings were higher than women's in both 2005 and 2016, even though women's average educational attainment was higher than that of men, a gap that increased from 2005 to 2016. Men still outperformed women in terms of labour market characteristics like labour market experience and job market tenure . The difference in the average labour market characteristics between men and women did not change substantially from 2005 to 2016. The proportion of married individuals decreased over the decade for both men and women. The geographic controls are largely stable across both gender and time.

Table 2 shows that average wages are highest for public-sector workers and that the GPG in the public sector is smaller than it is in the private sector. Average educational attainment is

Table 1: Descriptive Statistics Case 1: Women & Men in 2005 & 2016

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Women 2005		Women 2016		Men 2005		Men 2016	
Variable	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
Net Log Hourly Wages	1.851	0.368	2.072	0.439	2.003	0.403	2.159	0.459
Schooling (in Years)	12.94	2.693	13.86	2.408	12.22	2.902	13.06	2.642
Experience (in Years)	17.16	11.77	20.89	12.64	21.12	12.86	24.72	13.13
Tenure (in Years)	11.82	10.54	15.42	12.04	14.65	11.76	18.11	12.91
Married (Dummy)	0.563	0.496	0.344	0.475	0.592	0.491	0.358	0.479
North (Dummy)	0.523	0.500	0.530	0.499	0.457	0.498	0.470	0.499
Centre (Dummy)	0.208	0.406	0.215	0.411	0.188	0.390	0.203	0.403
Observations	3,983		4,359		5,202		5,789	

higher for women than for men and even higher in the public sector. Experience and job market tenure is higher for men in both sectors, but the gender difference in these characteristics is less pronounced in the public sector. Fewer public servants are married than are private-sector employees, but there are no substantial sector or gender differences in the geographic position of the employees in the sample.

Table 2: Descriptive Statistics Case 2: Public & Private Sector by Gender

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Private Sector Men		Private Sector Women		Public Sector Men		Public Sector Women	
Variable	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
Net Log Hourly Wages	2.086	0.471	1.952	0.461	2.295	0.402	2.203	0.372
Schooling (in Years)	12.70	2.686	13.51	2.553	13.73	2.419	14.25	2.176
Experience (in Years)	22.42	13.63	16.95	12.13	28.97	10.95	25.19	11.75
Tenure (in Years)	15.12	12.68	11.22	10.47	23.64	11.43	20.01	11.98
Married (Dummy)	0.422	0.494	0.454	0.498	0.239	0.426	0.223	0.416
North (Dummy)	0.547	0.498	0.613	0.487	0.327	0.469	0.440	0.497
Centre (Dummy)	0.199	0.400	0.202	0.402	0.211	0.408	0.229	0.421
Observations	3,756		2,275		2,033		2,084	

7 Empirical Results

This section presents two empirical applications of our proposed decomposition that are robust to variances with respect to categorical variables and the choice of the reference category. The first empirical application shows the convergence of the GPG during the last decade (2005-2016) in Italy, while the second estimates the difference in the PPWG between men and women directly. In both cases, we compare the results and implications from the standard approach with those of our decomposition approach.

In addition to looking at the average change in the pay gaps over time and by sector, we extend our proposed decomposition throughout the wage distribution using unconditional quantile regression (RIF-OLS). Apart from the aggregate decomposition, we also show the detailed decompositions for both cases using our proposed approach.

Empirical Results Case 1: The Gender Pay Gap between 2016 and 2005

Figure 1 shows the estimates of the GPG in 2005 and 2016, as well as our decomposition approach along the wage distribution. The convergence of the GPG was particularly pronounced at the top and bottom (except the first to fifth percentile) of the distribution. Apart from the very bottom of the distribution, the GPG was always lower in 2005 than it was in 2016, and the reduction in the GPG was always statistically significant.

Figure 2 shows that the most pronounced changes were in the explained (Q_t and Q_T) components, while the unexplained components (P_t and P_T) changed relatively more . Moreover, the unexplained component in the ending period (P_T), 2016, was much more volatile across the wage distribution than in 2005. In 2005, the unexplained part, P_t, was positively increasing toward the top, but in 2016 (P_T), it turned negative at the upper part of the distribution.

Table 3 shows the standard OB decomposition in the male-reference category (Panel A) and Fortin's regression-compatible decomposition (Panel B). In line with the literature that considers GPGs along the wage distribution, we find substantial differences in the gap at various points of the distribution. The 2016 GPG is lower than the 2005 GPG, yet the results suggest that the reduction was not evenly distributed across the distribution. The reduction was especially pronounced at the bottom of the distribution (from 11.8% to 2.9%). The explained component

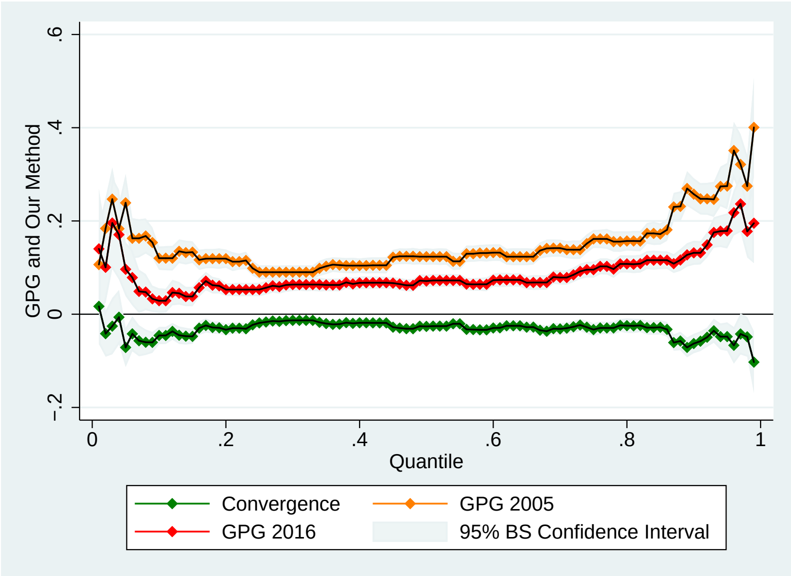

Figure 1: Gender Pay Gap in 2005 & 2016 and Convergence over Time along the Wage Distribution

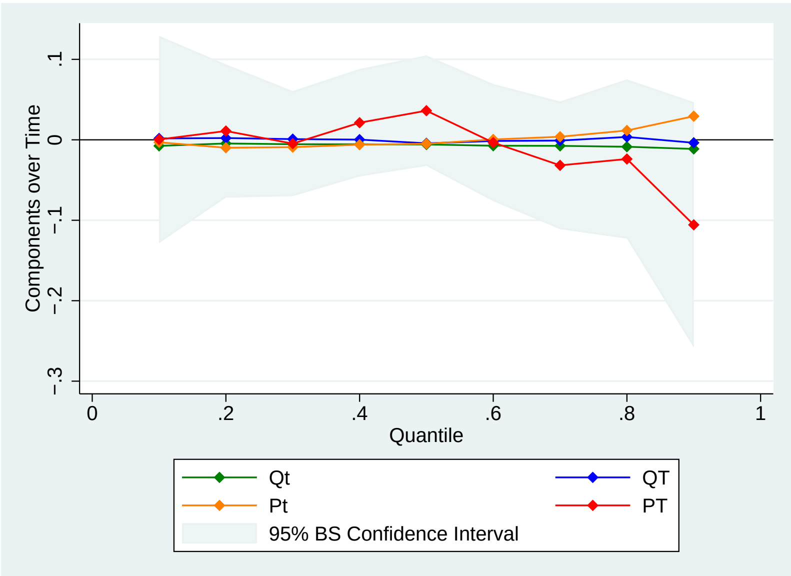

Figure 2: Our Decomposition Approach, Components of the Gender Pay Gap in 2005 & 2016 along the Wage Distribution

in 2005 was never statistically significant, while it became statistically significant at all points (except the 10th percentile) in 2016, when the explained component became not only significant but even negative. This result suggests that women outperform men in general human capital and labour market characteristics, although both the reduction and the unexplained component itself are relatively small. Therefore, one may conclude that there was no significant change in the explained component over the decade in Italy.

On the other hand, the unexplained component, although it was the main driver of the gap in both 2005 and 2016, was substantially smaller in 2016 than it was in 2006, suggesting a more pronounced decrease in the unexplained component over time. The results of the detailed decompositions that use the standard OB and regression-compatible approach are presented in Table E.1 in Appendix F.

Table 3: Decomposition of the Gender Pay Gap between 2016-2005 (Case 1) at the Mean & Selected Percentiles, Standard Approach

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2005	2016	2005	2016	2005	2016	2005	2016
Percentile	Mean	Mean	10.	10.	50.	50.	90.	90.
Panel A: Decomposition with Male-Reference Category								
GPG	0.152***	0.087***	0.118***	0.029**	0.121***	0.072***	0.252***	0.133***
	(0.008)	(0.009)	(0.013)	(0.013)	(0.009)	(0.007)	(0.017)	(0.013)
Explained	0.006	-0.012*	0.012	-0.009	0.009	-0.015***	-0.008	-0.038***
	(0.006)	(0.007)	(0.007)	(0.009)	(0.006)	(0.005)	(0.015)	(0.010)
Unexplained	0.146***	0.100***	0.106***	0.038***	0.112***	0.086***	0.261***	0.171***
	(0.008)	(0.009)	(0.014)	(0.015)	(0.009)	(0.007)	(0.020)	(0.015)
Panel B: Regression-Compatible Decomposition								
GPG	0.152***	0.087***	0.118***	0.029**	0.121***	0.072***	0.252***	0.133***
	(0.008)	(0.009)	(0.013)	(0.013)	(0.009)	(0.007)	(0.017)	(0.013)
Explained	-0.001	-0.014**	-0.002	-0.015**	0.004	-0.017***	0.006	-0.021***
	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)	(0.004)	(0.011)	(0.008)
Unexplained	0.152***	0.101***	0.120***	0.044***	0.117***	0.089***	0.246***	0.154***
	(0.007)	(0.009)	(0.013)	(0.014)	(0.008)	(0.006)	(0.018)	(0.014)

9,185 observations in 2005 & 10,148 observations in 2016. Robust standard errors in parentheses. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table 4 shows the results of our proposed decomposition. The coefficient estimate of α_1^{base} is negative and statistically significant, suggesting that the GPG significantly decreased between 2016-2005 in Italy. The convergence amounts to —6.8— log points at the mean and is relatively more pronounced at the bottom and top of the wage distribution. This result is in line with what we expected from the standard estimation technique (Table 3), but now we can also conclude that the reduction in the GPG was statistically significant along the wage distribution.

We find no statistically significant coefficient estimate $\hat{\alpha}_1^{full}$ throughout the wage distribution, so our (full) model can completely explain the convergence. The decomposition shows that the only significant component was gender differences in human capital characteristics in 2016, which is again in line with the results from the standard decomposition. However, the t-test on the coefficient estimates of the explained components in both periods suggests a statistically significant difference at the 5% significance level. The unexplained component is statistically significant only in 2016 at the top of the wage distribution. In addition, we do not reject the H_0 that the price component was stable over the decade in Italy, so the observed convergence of 6.8 log points at the mean may have been due only to changes in the explained component, while the unexplained part did not change significantly.

However, we find substantial differences in the component that drove the convergence across the wage distribution. Only at the bottom and the mean of the distribution was the convergence due to the explained component, while at the top it was due to a statistically significant change in the price component. To explain what drove the change in the explained and unexplained part, we decompose the change in the GPG over time in some detail. Table 5 shows that the convergence of the GPG is due only to women's catching up in education and labour market characteristics and narrowing the gap in demographic characteristics and occupational/sectoral sorting at the median of the wage distribution.[10]

At the bottom and top of the distribution only changes in observable labour market characteristics contributed statistically significantly to reducing the GPG from 2005 to 2016 in Italy. The mean of the distribution suggests that changes in the occupational/sectoral sorting of women and men over time significantly affected the convergence of the gap. Remuneration between men and women narrowed significantly based on educational attainment only at the median of the the distribution. Returns t o demographic characteristics narrowed in the middle and top of the wage distribution, and gender differences in returns to labour market characteristics like experience and job tenure were equalized only at the top.

[10]The detailed decomposition of the standard approach is shown in Table E.1.

Table 4: Proposed Regression-Compatible Decomposition of the Gender Pay Gap between 2016-2005 (Case 1) at the Mean & Selected Percentiles

Percentile	(1) Mean	(2) 10.	(3) 50.	(4) 90.
$\Delta GPG = 2 * \gamma_{FJ}^{\text{base}}$	-0.068***	-0.091***	-0.052***	-0.126***
	(0.006)	(0.009)	(0.005)	(0.011)
$\Delta GPG = 2 * \gamma_{FJ}^{\text{full}}$	0.061	-0.075	-0.095	0.057
	(0.044)	(0.083)	(0.037)	(0.062)
Explained: Q_T (X)	-0.009***	-0.007**	-0.006**	-0.011***
	(0.003)	(0.004)	(0.002)	(0.004)
Explained Q_t (XY)	0.002	0.002	-0.004	-0.004
	(0.003)	(0.005)	(0.003)	(0.006)
$H_0 : Q_T = Q_t$				
(χ^2-Statistic, P-value)	(4.98,0.026)	(2.84,0.09)	(1.07,0.30)	(1.58,0.21)
Unexplained: P_T (fmX)	0.003	-0.003	-0.005	0.029***
	(0.003)	(0.006)	(0.003)	(0.007)
Unexplained: P_t (fmXY)	-0.060	0.00	0.036	-0.106
	(0.040)	(0.066)	(0.035)	(0.078)
$H_0 : P_T = P_t$	(2.64,0.104)	(1.37,0.24)	(2.39,0.12)	(3.99,0.05)

Observations: 19,333. Bootstrapped standard errors in parentheses (500 replications). X represents the set of observable characteristics included in the regression, XY is the interaction of the set of covariates used with the time indicator. fmX and $fmXY$ are the corresponding interactions with the $(F - M)$ gender dummy. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$

Table 5: Proposed Regression-Compatible Detailed Decomposition of the Gender Pay Gap between 2016-2005 (Case 1) at the Mean & Selected Percentiles

Percentile	(1) Mean	(2) 10.	(3) 50.	(4) 90.
$\Delta GPG = 2 * \gamma_{FJ}^{\text{base}}$	-0.068***	-0.091***	-0.052***	-0.126***
	(0.006)	(0.009)	(0.005)	(0.011)
$\Delta GPG = 2 * \gamma_{FJ}^{\text{full}}$	0.061	-0.075	-0.095	0.057
	(0.044)	(0.083)	(0.037)	(0.062)
Explained: Q_T (X)				
X_HC	-0.001	-0.001	-0.001	-0.002
	(0.001)	(0.001)	(0.001)	(0.002)
X_LM	-0.003	-0.005**	-0.002	-0.001
	(0.002)	(0.003)	(0.001)	(0.002)
X_Demo	0.001	0.000	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
X_OccInd	-0.005***	-0.002	-0.004***	-0.009***
	(0.001)	(0.002)	(0.001)	(0.003)
Explained: Q_t (XY)				
XY_HC	0.000	-0.001	0.002***	0.003
	(0.001)	(0.002)	(0.001)	(0.002)
XY_LM	-0.001	0.004	-0.006***	-0.006**
	(0.001)	(0.002)	(0.001)	(0.003)
XY_Demo	-0.000	-0.001	-0.003***	0.002
	(0.001)	(0.001)	(0.001)	(0.001)
XY_OccInd	0.003	-0.000	0.002	-0.002
	(0.003)	(0.005)	(0.003)	(0.006)
$H_0: Q_T\_HC = Q_t\_HC$				
(χ^2-Statistic, P-value)	(1.03,0.31)	(1.28,0.26)	(9.92,0.002)	(4.98,0.023)
$H_0: Q_T\_LM = Q_t\_LM$	(0.75,0.39)	(6.41,0.01)	(8.02,0.005)	(3.38,0.07)
$H_0: Q_T\_Demo = Q_t\_Demo$	(0.77,0.38)	(1.82,0.18)	(8.05,0.005)	(1.47,0.23)
$H_0: Q_T\_OccInd = Q_t\_OccInd$	(4.42,0.04)	(1.43,0.23)	(4.67,0.03)	(1.3,0.26)
Unexplained: P_T (fmX)				
fmX_HC	-0.004**	-0.008***	-0.004***	0.007**
	(0.002)	(0.003)	(0.001)	(0.003)
fmX_LM	0.002	-0.003	0.000	0.004
	(0.001)	(0.002)	(0.001)	(0.003)
fmX_Demo	0.004	0.004	0.004**	0.008*
	(0.002)	(0.004)	(0.002)	(0.004)
fmX_OccInd	0.002	0.004	-0.006***	0.011**
	(0.002)	(0.004)	(0.002)	(0.005)
Unexplained: P_t (fmXY)				
fmXY_HC	-0.032	0.034	0.050*	-0.058
	(0.033)	(0.054)	(0.029)	(0.064)
fmXY_LM	-0.019	-0.021	0.001	-0.036
	(0.016)	(0.027)	(0.014)	(0.032)
fmXY_Demo	-0.003*	-0.002	-0.006***	-0.003
	(0.002)	(0.003)	(0.001)	(0.003)
fmXY_OccInd	-0.006	-0.010	-0.009	-0.008
	(0.012)	(0.020)	(0.011)	(0.024)
$H_0: P_T\_HC = P_t\_HC$				
(χ^2-Statistic, P-value)	(0.81,0.37)	(1.9,0.17)	(4.81,0.03)	(1.88,0.17)
$H_0: P_T\_LM = P_t\_LM$	(1.72,0.19)	(1.89,0.17)	(1.11,0.29)	(2.75,0.098)
$H_0: P_T\_Demo = P_t\_Demo$	(3.46,0.06)	(2.04,0.15)	(10.28,0.001)	(3.24,0.072)
$H_0: P_T\_OccInd = P_t\_OccInd$	(0.37,0.54)	(2.12,0.15)	(1.0,0.32)	(0.84,0.36)

Observations: 19,333. Bootstrapped standard errors in parentheses (500 replications). X represents the set of observable characteristics included in the regression, XY is the interaction of the set of covariates used with the time indicator. fmX and $fmXY$ are the corresponding interactions with the $(F-M)$ gender dummy. The set of covariates X is split in the following way: HC contains years of schooling, LM includes labor market experience, its square, job tenure as well as firm-size dummies, $Demo$ includes a dummy for being married and place of residence (North & Centre) and $OccInd$ occupational and sectoral dummies. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Empirical Results Case 2: Private-Public Sector Wage Gap between Women and Men

Here we discuss the results of the second empirical application of -that is, the change in the PPWG between men and women from 2005 to 2016 in Italy. Figure 3 shows the estimated (raw) PPWG for men and women in 2016, as well as our decomposition approach along the wage distribution. The PPWGs are especially pronounced at the bottom of the wage distribution. The difference in the PPWG between men and women is especially pronounced at the very bottom and top of the distribution, while it is relatively small elsewhere in the distribution.

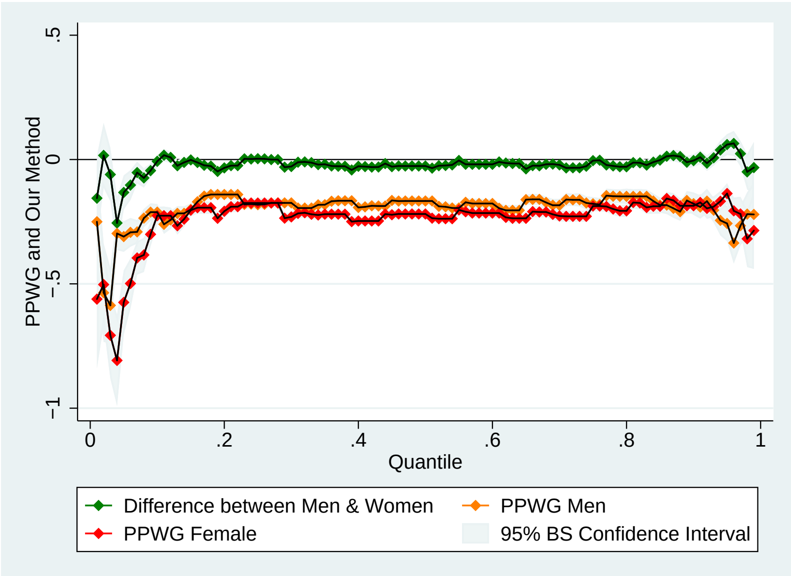

Figure 3: PPWG for Women & Men as well as Difference by Gender

Figure 4 shows that the explained components diverged the most at the bottom and top of the distribution (Q_t and Q_T, where t = men and T = women), and the corresponding unexplained components (P_t and P_T) are also more pronounced at these points of the distribution. In the middle of the distribution, the components tend to be constant.

Table 6 shows the decomposition results of the standard OB method using male reference parameters and the regression-compatible decomposition of Fortin (2008). We find negative PPWGs at all points of the distribution that are due to higher pay levels in the public sector. For women, the unexplained component is more pronounced than it is for men, suggesting differences in the unexplained component between women and men across sectors .

Table 7 shows a positive and statistically significant difference in the PPWG between women and men of $|-0.04|$ log points. By construction, this result is in line with the standard approach's estimation results (Table 6). We find a statistically significant difference between the men's and women's PPWG at the mean and median of the distribution. The negative sign indicates that the sectoral earnings gap is most pronounced for women, so women earn substantially less in the

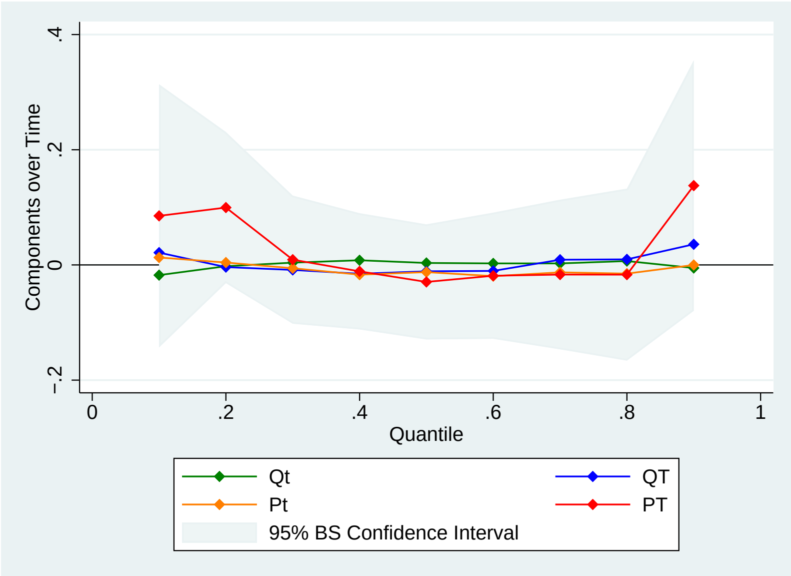

Figure 4: Our Decomposition Approach, Components of the PPWG

private sector than they do in the public sector and compared to their male colleagues in the private sector. The unexplained component is never statistically significantly different, while the explained component contributes substantially to the pay divergence at the bottom and top at the 10% significance level. The significant difference between men's and women's PPWG in the middle of the wage distribution cannot be explained by the aggregate decomposition.

The detailed decomposition of the PPWG between women in men, shown in the Table 8, allows us to attribute the change to gender differences in education, experience in the labour market, demographic characteristics, and occupational sorting. However, these characteristics' contributions vary across the wage distribution. Gender differences in remuneration that are due to occupational (and sectoral) sorting contribute statistically significantly to the PPWG between men and women only at the 10th percentile. Thus, sorting of higher-earning individuals, particularly higher-earning women in public-sector employment, explains the difference in the PPWG between women and men.[11]

[11] The detailed decomposition results from using the standard OB and regression-compatible decomposition are shown in Table E.2 in Appendix F.

Table 6: Decomposition of the Private-Public Sector Wage Gap by Gender (Case 2) at the Mean & Selected Percentiles, Standard Approach

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Men	Women	Men	Women	Men	Women	Men	Women
Percentile	Mean		10.		50.		90.	
Panel A: Decomposition with Male-Reference Category								
PPWG	-0.210***	-0.252***	-0.214***	-0.228***	-0.168***	-0.221***	-0.179***	-0.188***
	(0.012)	(0.013)	(0.018)	(0.018)	(0.009)	(0.009)	(0.023)	(0.016)
Explained	-0.134***	-0.131***	-0.113**	-0.147***	-0.101***	-0.125***	-0.153***	-0.114***
	(0.024)	(0.022)	(0.046)	(0.035)	(0.017)	(0.015)	(0.036)	(0.020)
Unexplained	-0.076***	-0.121***	-0.101**	-0.081**	-0.068***	-0.097***	-0.026	-0.074***
	(0.026)	(0.024)	(0.048)	(0.038)	(0.018)	(0.017)	(0.041)	(0.025)
Panel B: Regression-Compatible Decomposition								
PPWG	-0.210***	-0.252***	-0.214***	-0.228***	-0.168***	-0.221***	-0.179***	-0.188***
	(0.012)	(0.013)	(0.017)	(0.018)	(0.009)	(0.009)	(0.023)	(0.016)
Explained	-0.157***	-0.153***	-0.117***	-0.147***	-0.123***	-0.136***	-0.224***	-0.169***
	(0.012)	(0.014)	(0.019)	(0.018)	(0.009)	(0.009)	(0.021)	(0.016)
Unexplained	-0.052***	-0.099***	-0.096***	-0.081***	-0.045***	-0.085***	0.045*	-0.018
	(0.015)	(0.018)	(0.026)	(0.024)	(0.011)	(0.012)	(0.027)	(0.020)

Observations: 5,789 men and 4,359 women. Robust standard errors in parentheses. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$

8 Conclusion

This paper proposes a new method that allows changes in pay gaps between groups and across samples to be decomposed directly. For illustration, we show two empirical applications of the method using data for Italy, but the methodology can be used to study (the evolution of) group differences in any (continuous and unbounded) outcome variable. Our alternative decomposition method overcomes essential flaws in the standard OB method (the index number and omitted group problems) and allows inferences to be drawn on the difference between two wage gaps. The proposed method can easily be extended beyond the mean by using linear unconditional quantile regressions and allows both aggregate and detailed decomposition. We propose to decompose the GPG following the intercept-shift approach proposed by Fortin (2008) and applying the OVB formula (as proposed by Gelbach, 2016). Thus, when we conduct a detailed decomposition, we can associate the single parts of the explained component with specific covariates, and the invariance problem with respect to categorical variables can be solved (Gardeazabal and Ugidos, 2004; Fortin, 2008).

Table 7: Proposed Regression-Compatible Decomposition of the Private-Public Wage Gap between Women and Men (Case 2) at the Mean & Selected Percentiles

	(1)	(2)	(3)	(4)
Percentile	Mean	10.	50.	90.
$\Delta GPG = 2 * \gamma_{FJ}^{\text{base}}$	-0.042**	-0.014	-0.052***	-0.008
	(0.008)	(0.012)	(0.006)	(0.014)
$\Delta GPG = 2 * \gamma_{FJ}^{\text{full}}$	-0.169	-0.276	0.050	-0.354**
	(0.078)	(0.113)	(0.058)	(0.088)
Explained: Q_T (X)	-0.004	-0.018*	0.003	-0.005
	(0.007)	(0.011)	(0.005)	(0.010)
Explained Q_t (XY)	0.010	0.021	-0.011	0.036*
	(0.013)	(0.022)	(0.010)	(0.021)
$H_0 : Q_T = Q_t$				
χ^2-Statistic, P-value	0.65, 0.42	2.98, 0.085	2.1, 0.15	2.92, 0.087
Unexplained: P_T (fmX)	-0.007	0.013	-0.012*	-0.000
	(0.010)	(0.016)	(0.007)	(0.015)
Unexplained: P_t (fmXY)	0.044	0.085	-0.029	0.138
	(0.071)	(0.117)	(0.051)	(0.111)
$H_0 : P_T = P_t$				
χ^2-Statistic, P-value	0.49, 0.48	1.64, 0.20	1.18, 0.28	2.09, 0.15

Observations: 10,148. Bootstrapped standard errors in parentheses (500 replications). X represents the set of observable characteristics included in the regression, XY is the interaction of the set of covariates used with the time indicator. fmX and $fmXY$ are the corresponding interactions with the $(F - M)$ gender dummy. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$

Table 8: Proposed Regression-Compatible Detailed Decomposition of the Private-Public Wage Gap between Women and Men (Case 2) at the Mean & Selected Percentiles

Percentile	(1) Mean	(2) 10.	(3) 50.	(4) 90.
$\Delta GPG = 2 * \gamma_{FJ}^{\text{base}}$	-0.042**	-0.014	-0.052***	-0.008
	(0.008)	(0.012)	(0.006)	(0.014)
$\Delta GPG = 2 * \gamma_{FJ}^{\text{full}}$	-0.169	-0.276	0.050	-0.354**
	(0.078)	(0.113)	(0.058)	(0.088)
Explained: Q_T (X)				
X_HC	0.005***	0.004**	0.004***	0.005***
	(0.002)	(0.002)	(0.002)	(0.002)
X_LM	-0.006**	-0.006*	-0.005***	-0.006*
	(0.003)	(0.004)	(0.002)	(0.003)
X_Demo	-0.003***	-0.004***	-0.002***	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
X_OccInd	-0.000	-0.012	0.006	-0.004
	(0.006)	(0.010)	(0.005)	(0.010)
Explained: Q_t (XY)				
XY_HC	-0.004*	-0.007*	-0.006***	0.008**
	(0.002)	(0.004)	(0.002)	(0.004)
XY_LM	0.007	0.003	0.008	0.016
	(0.008)	(0.012)	(0.005)	(0.012)
XY_Demo	0.005*	0.009*	0.001	0.007
	(0.003)	(0.005)	(0.002)	(0.004)
XY_OccInd	0.002	0.017	-0.014	0.005
	(0.012)	(0.019)	(0.008)	(0.018)
$H_0: Q_T\_HC = Q_t\_HC$				
(χ^2-Statistic, P-value)	(7.57, 0.006)	(7.1, 0.008)	(13.55, 0.0002)	(1.29, 0.26)
$H_0: Q_T\_LM = Q_t\_LM$	(3.05, 0.08)	(1.67, 0.20)	(6.69, 0.0097)	(4.49, 0.03)
$H_0: Q_T\_Demo = Q_t\_Demo$	(6.13, 0.01)	(8.02, 0.005)	(1.53, 0.22)	(3.16, 0.076)
$H_0: Q_T\_OccInd = Q_t\_OccInd$	(0.02, 0.88)	(2.28, 0.13)	(3.14, 0.08)	(0.73, 0.39)
Unexplained: P_T (fmX)				
fmX_HC	-0.000	0.010**	-0.001	-0.006
	(0.002)	(0.004)	(0.002)	(0.004)
fmX_LM	-0.002	-0.015***	-0.002	0.003
	(0.003)	(0.005)	(0.002)	(0.005)
fmX_Demo	0.002	0.010***	0.000	0.001
	(0.001)	(0.003)	(0.001)	(0.002)
fmX_OccInd	-0.007	0.008	-0.010	0.001
	(0.009)	(0.015)	(0.007)	(0.015)
Unexplained: P_t (fmXY)				
fmXY_HC	0.028	0.118	-0.034	0.092
	(0.058)	(0.095)	(0.041)	(0.090)
fmXY_LM	0.038	-0.006	0.013	0.036
	(0.032)	(0.053)	(0.023)	(0.050)
fmXY_Demo	0.001	0.002	-0.000	-0.001
	(0.001)	(0.002)	(0.001)	(0.002)
fmXY_OccInd	-0.023	-0.028	-0.008	0.011
	(0.020)	(0.033)	(0.014)	(0.031)
$H_0: P_T\_HC = P_t\_HC$				
(χ^2-Statistic, P-value)	(0.25, 0.62)	(2.42, 0.12)	(1.81, 0.18)	(1.65, 0.20)
$H_0: P_T\_LM = P_t\_LM$	(1.40, 0.24)	(1.14, 0.29)	(1.41, 0.24)	(1.41, 0.24)
$H_0: P_T\_Demo = P_t\_Demo$	(0.34, 0.56)	(5.45, 0.02)	(1.05, 0.3)	(1.03, 0.31)
$H_0: P_T\_OccInd = P_t\_OccInd$	(0.63, 0.43)	(1.58, 0.21)	(1.0, 0.32)	(0.73, 0.39)

Observations: 10,148. Bootstrapped standard errors in parentheses (500 replications). X represents the set of observable characteristics included in the regression, XY is the interaction of the set of covariates used with the time indicator. fmX and $fmXY$ are the corresponding interactions with the $(F-M)$ gender dummy. The set of covariates X is split in the following way: HC contains years of schooling, LM includes labor market experience, its square, job tenure as well as firm-size dummies, $Demo$ includes a dummy for being married and place of residence (North & Centre) and $OccInd$ occupational and sectoral dummies. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

The first application adds to the discussion of the convergence of the GPG over time. Changes in the GPG over the last several decades have been widely discussed in the literature, and determining the reasons for the narrowing is of significant interest, especially with regard to policy implications (Blau and Kahn, 2006,0; Godin, 2014). We find a statistically significant convergence of the GPG over period from 2005 to 2016 in Italy that can be explained only by a reduction in the differences in observable labour market characteristics at the mean and the bottom of the wage distribution. At the top of the distribution, changes in the unexplained component have led to a significant reduction in the GPG. The change in the price component that results when using the standard OB decomposition and estimating the GPG separately for 2005 and 2016-that is, following the double OB decomposition-might have suggested that the implementation of anti-discrimination laws and changing attitudes toward women in the labour market have influenced the narrowing of the pay gap over time as well, but these policies and changes in social norms appear to have been less effective than expected (except at the top of the distribution). In fact, even if the unexplained part is found to be the main contributor to the GPG in a particular year, it becomes insignificant when changes in the GPG are estimated directly over time.

The results for the second empirical application, the PPWG between women and men, point to differences in employee-sorting in the public and private sectors, which are found to be a main driver of the differential. The decomposition we carried out allows us to determine whether both parts of the price effect-the difference in the intercepts and the difference in remuneration-drive the change in the pay gap in a statistically significant way. We conclude that both parts contribute to the difference in the PPWG between women and men, although they do so differently along the wage distribution.

All in all, our decomposition method helps to clarify what led to the narrowing of the GPG from 2005 to 2016 in Italy and what drives the difference in the PPWG between women and men. Policymakers would benefit from taking our inferences about what drives the difference in the respective pay gaps into consideration. Our proposed decomposition approach leads to additional insights on the composition of differences in gaps. The method can be extended to the choice of the reference category (Reimers, 1983; Cotton, 1988; Neumark, 1988; Oaxaca and

Ransom, 1994; Fortin, 2008) as well as to the indeterminacy problem (Lee, 2015).

References

Arulampalam, W., Booth, A. L. and Bryan, M. L. (2007). Is there a Glass Ceiling over Europe? Exploring the Gender Pay Gap across the Wage Distribution. *Industrial and Labor Relations Review*, **60** (2), 163–186.

Black, S. E. and Spitz-Oener, A. (2010). Explaining Women's Success: Technological Change and the Skill Content of Women's Work. *The Review of Economics and Statistics*, **92** (1), 187–194.

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

— and — (2006). The US Gender Pay Gap in the 1990s: Slowing Convergence. *Industrial Labor Relations Review*, (60), 45–66.

— and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, **55** (3), 789–865.

Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *Journal of Human Resources*, **8** (4), 436–455.

Castagnetti, C. and Giorgetti, M. (2018). Understanding the gender wage-gap differential between the public and private sectors in italy: A quantile approach. *Economic Modelling*, pp. 245–265.

Cotter, J. V. R., D.A.; Hermsen (2004). *Gender Inequality at Work*. Russell Sage Foundation.

Cotton, J. (1988). On the Decomposition of Wage Differentials. *The Review of Economics and Statistics*, **70** (2), 236–43.

Elder, T. E., Goddeeris, J. H. and Haider, S. J. (2010). Unexplained gaps and Oaxaca–Blinder decompositions. *Labour Economics*, **17** (1), 284–290.

England, P. (2006). Towards Gender Equality: Progress and Bottlenecks. In F.D. Blau, M.C. Brinton & D.B. Grusky (Eds.). *The Declining Significance of Gender?*, (63), 245–265.

Eurostat (2015). Gender Pay Gap in Unadjusted Form. http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=earn_gr_gpgr2&lang=en, accessed: 2019-02-07.

Firpo, S., Fortin, N. M. and Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica*, **77** (3), 953–973.

Fortin, N. (2008). The Gender Wage Gap among Young Adults in the United States: The Importance of Money versus People. *Journal of Human Resources*, **43** (4).

— (2015). Gender Role Attitudes and Women's Labor Market Participation: Opting-out AIDS and the Persistent Appeal of Housewifery. *Annales of Economics and Statistics*, (117-118), 379–401.

—, Lemieux, T. and Firpo, S. (2011). *Decomposition Methods in Economics*, Elsevier, *Handbook of Labor Economics*, vol. 4, chap. 1, pp. 1–102.

Gardeazabal, J. and Ugidos, A. (2004). More on Identification in Detailed Wage Decompositions. *Review of Economics and Statistics*, **86** (4), 1034–1036.

Gelbach, J. B. (2016). When Do Covariates Matter? And Which Ones, and How Much? *Journal of Labor Economics*, **34** (2), 509–543.

Godin, C. (2014). A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, **104** (4), 1091–1119.

Goldin, C. (2006). The Quiet Revolution that Transformed Women's Employment, Education, and Family. *American Economic Review*, **96** (2), 1–21.

Gornick, J. and Jacobs, J. (1998). Gender, the Welfare State, and Public Employment: A Comparative Study of Seven Industrialized Countries. *American Sociological Review*, (63), 688–710.

Grimshaw, D. (2000). Public Sector Employment, Wage Inequality and the Gender Pay Gap in the UK. *International Review of Applied Economics*, (14), 427–448.

Gupta, A., Oaxaca, R. and Smith, N. (2006). Swimming upstream, floating downstream: Comparing women's relative wage progress in the united states and denmark. *Industrial and Labor Relations Review*, **59** (2), 243–266.

Jones, F. (1983). On decomposing the wage gap: A critical comment on blinder's method. *The Journal of Human Resources*, **18**, 126–130.

Juhn, C., Murphy, K. M. and Pierce, B. (1991). *Workers and their wages: Changing patterns in the United States*, Washington DC: AEI Press, chap. Accounting for the Slowdown in Black- White Wage Convergence.

—, — and — (1993). Wage Inequality and the Rise in Returns to Skill. *Journal of Political Economy*, **101** (3), 410–442.

Lee, M.-J. (2015). Reference Parameters in Blinder-Oaxaca Decomposition: Pooled-Sample versus Intercept-Shift Approaches. *The Journal of Economic Inequality*, **13** (1), 69–82.

Lucifora, C. and Meurs, D. (2006). The Public Sector Pay Gap In France, Great Britain and Italy. *Review of Income and Wealth*, **52** (1), 43–59.

Magnani, E. and Zhu, R. (2012). Gender wage differentials among rural-urban migrants in china. *Regional Science and Urban Economics*, **42**, 779–793.

Mandel, H. and Semyonov, M. (2014). Gender Pay Gap and Employment Sector: Sources of Earnings Disparities in the United States, 19702010. *Demography*, **51** (5), 1597–1618.

Melly, B. (2005). Public-Private Sector Wage Differentials in Germany: Evidence from Quantile Regression. *Empirical Economics*, **30** (2), 505–520.

Neumark, D. (1988). Employers' Discriminatory Behavior and the Estimation of Wage Discrimination. *Journal of Human Resources*, **23** (3), 279–295.

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, **14** (3), 693–709.

— and Ransom, M. R. (1994). On Discrimination and the Decomposition of Wage Differentials. *Journal of Econometrics*, **61** (1), 5–21.

— and — (1999). Identification in Detailed Wage Decompositions. *Review of Economics and Statistics*, **81** (1), 154–157.

Reimers, C. W. (1983). Labor Market Discrimination Against Hispanic and Black Men. *The Review of Economics and Statistics*, **65**.

Smith, J. and Welch, F. (1989). Black Economic Progress after Myrdal. *Journal of Economic Literature*, **27** (2), 519–564.

Stanley, T. D. and Jarrell, S. B. (1998). Gender Wage Discrimination Bias? A Meta-Regression Analysis. *Journal of Human Resources*, **33** (4), 947–973.

Suen, W. (1997). Decomposing wage residuals: Unmeasured skill or statistical artifact? *Journal of Labor Economics*, **15**, 555–66.

Weichselbaumer, D. and Winter-Ebmer, R. (2005). A Meta-Analysis of the International Gender Wage Gap. *Journal of Economic Surveys*, **19** (3), 479–511.

Yun, M.-S. (2005). A simple solution to the identification problem in detailed wage decompositions. *Economic inquiry*, **43** (4), 766–772.

— (2009). Wage Differentials, Discrimination and Inequality: A Cautionary Note on the Juhn, Murphy and Pierce Decomposition Method. *Scottish Journal of Political Economy*, **56** (1), 114–122.

Appendices

A Parameter Link

The links of the parameters presented in Section 3 are presented here in detail.

1. When F=1, i.e. Female, J=1, i.e. year t:
   - $\hat{\alpha}_{Ft} = \hat{\alpha}_0^{full} + \hat{\alpha}_1^{full} + \hat{\alpha}_2^{full} + \hat{\alpha}_3^{full}$
   - $\hat{\beta}_{Ft} = \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 + \hat{\beta}_4$
2. When F=0, i.e. Male, J=1, i.e. year t, we get:
   - $\hat{\alpha}_{Mt} = \hat{\alpha}_0^{full} + \hat{\alpha}_3^{full}$
   - $\hat{\beta}_{Mt} = \hat{\beta}_1 + \hat{\beta}_3$
3. When F=1, i.e. Female, J=0, i.e. year T, we get:
   - $\hat{\alpha}_{FT} = \hat{\alpha}_0^{full} + \hat{\alpha}_2^{full}$
   - $\hat{\beta}_{FT} = \hat{\beta}_1 + \hat{\beta}_2$
4. When F=0, i.e. Male, J=0, i.e. year T, we get:
   - $\hat{\alpha}_{MT} = \hat{\alpha}_0^{full}$
   - $\hat{\beta}_{MT} = \hat{\beta}_1$

Re-arranging terms slightly, gives us:

$$
\begin{aligned}
\hat{\alpha}_0^{full} &= \hat{\alpha}_{MT} \\
\hat{\alpha}_2^{full} &= \hat{\alpha}_{FT} - \hat{\alpha}_{MT} \\
\hat{\alpha}_3^{full} &= \hat{\alpha}_{MT} - \hat{\alpha}_{Mt} \\
\hat{\alpha}_1^{full} &= \hat{\alpha}_{Ft} - \hat{\alpha}_{MT} - \hat{\alpha}_{FT} + \hat{\alpha}_{MT} - \hat{\alpha}_{Mt} + \hat{\alpha}_{MT} \\
&= (\hat{\alpha}_{MT} - \hat{\alpha}_{FT}) - (\hat{\alpha}_{Mt} - \hat{\alpha}_{Ft})
\end{aligned}
$$

$$
\begin{aligned}
\hat{\beta}_1 &= \hat{\beta}_{MT} \\
\hat{\beta}_2 &= \hat{\beta}_{FT} - \hat{\beta}_{MT} \\
\hat{\beta}_3 &= \hat{\beta}_{Mt} - \hat{\beta}_{MT} \\
\hat{\beta}_4 &= \hat{\beta}_{MT} - \hat{\beta}_{FT} - \hat{\beta}_{Mt} + \hat{\beta}_{Ft} \\
&= (\hat{\beta}_{MT} - \hat{\beta}_{FT}) - (\hat{\beta}_{Mt} - \hat{\beta}_{Ft})
\end{aligned}
$$

B Inference

The derivation of the asymptotic distribution of $\sqrt{N}\hat{\delta} = \sqrt{N}(\hat{\delta}_X \hat{\delta}_{FX} \hat{\delta}_{JX} \hat{\delta}_{FJX})$ follows the same line of argument in Gelbach (2016). In particular, given that all estimators involved in the decomposition are asymptotically normal and given that the decomposition involves continuously differentiable functions of these estimators, joint asymptotic normality of the decomposition components follows from the delta method.

The elements of the decomposition, $\hat{\delta}$ in (17) can be written as:

$$\sqrt{N}(\hat{\delta} - \delta) = \sqrt{N}(\hat{\Gamma}\hat{\beta^{full}} - \Gamma\beta^{full}) \tag{B.1}$$

where

$$X_2 = X_1\Gamma + W \tag{B.2}$$

with W, matrix $(N \times 4K)$ of the error terms. (B.1) can be written as:

$$\sqrt{N}(\hat{\delta} - \delta) = \hat{\Gamma}\sqrt{N}(\hat{\beta^{full}} - \beta^{full}) + \sqrt{N}(\hat{\Gamma} - \Gamma)\beta^{full} \tag{B.3}$$

where $\hat{\Gamma} = (X_1^{'}X_1)^{-1}X_1^{'}X_2$ and given that:

$$\hat{\Gamma} - \Gamma = (X_1^{'}X_1)^{-1}X_1^{'}W$$

(B.3) can be expressed as:

$$\sqrt{N}(\hat{\delta} - \delta) = \hat{\Gamma}\sqrt{N}(\hat{\beta^{full}} - \beta^{full}) + (\frac{X_1^{'}X_1}{N})^{-1}\frac{X_1^{'}W}{\sqrt{N}}\beta^{full} \tag{B.4}$$

The asymptotic variance of the vector $\hat{\delta}$ is given by:

$$\begin{aligned} AsyCov(\hat{\delta}) = & \underbrace{\hat{\Gamma}AsyVar(\hat{\beta}^{full})\hat{\Gamma}'}_{I} \\ & + \underbrace{(\frac{X_1^{'}X_1}{N})^{-1}\text{plim}(\frac{X_1^{'}W\beta^{full}\beta^{full'}W^{'}X_1}{N})(\frac{X_1^{'}X_1}{N})^{-1}}_{II} \\ & + \underbrace{\hat{\Gamma}AsyCov(\sqrt{N}(\hat{\beta^{full}} - \beta^{full}), \frac{X_1^{'}W\beta^{full}}{\sqrt{N}})(\frac{X_1^{'}X_1}{N})^{-1}}_{III} \\ & + \underbrace{(\frac{X_1^{'}X_1}{N})^{-1}AsyCov(\frac{X_1^{'}W\beta^{full}}{\sqrt{N}}, \sqrt{N}(\hat{\beta^{full}} - \beta^{full}))\hat{\Gamma}'}_{IV} \end{aligned} \tag{B.5}$$

where the consistent estimators for the matrices $Q = E[x_{1,i}x_{1,i}^{'}]$ and Γ have been already substituted in (B.4) by their consistent estimators: $\frac{X_1^{'}X_1}{N}$ and $\hat{\Gamma}$, respectively . Term I in (B.4) entails the asymptotic variance of $\hat{\beta^{full}}$ that can be consistently estimated under standard assumptions. In particular, consider the vector of all the parameters estimated from the full model $\hat{\beta} = (\hat{\alpha}^{full'}\hat{\beta}^{full'})$:

$$var(\hat{\beta}) = (X^{'}X)^{-1}(X^{'}\Sigma X)(X^{'}X)^{-1} \tag{B.6}$$

where Σ is the variance covariance matrix of the error terms ϵ^{full} in the full specification (12) and $X = [X_1 X_2]$. The asymptotic variance of $\hat{\beta}^{full}$ is the sub-block of $var(\hat{\beta})$ corresponding to the variables in X_2.

By organizing the observations for group and sector (for instance gender and period) ϵ^{full} can be thought as $\epsilon^{full'} = (\epsilon_{Ft}^{full'} \epsilon_{Mt}^{full'} \epsilon_{FT}^{full'} \epsilon_{MT}^{full'})$ where F = female and M = male and t = *starting period* and T = *ending period*. It follows that

$$\Sigma = \begin{bmatrix} \sigma_{Ft}^2 \mathbf{1}_{N_{Ft},N_{Ft}} & \sigma_{Ft,Mt}\mathbf{1}_{N_{Ft},N_{Mt}} & \sigma_{Ft,FT}\mathbf{1}_{N_{Ft},N_{FT}} & \sigma_{Ft,MT}\mathbf{1}_{N_{Ft},N_{MT}} \\ \sigma_{Ft,Mt}\mathbf{1}_{N_{Mt},N_{Ft}} & \sigma_{Mt}^2\mathbf{1}_{N_{Mt},N_{Mt}} & \cdots & \cdots \\ \sigma_{Ft,FT}\mathbf{1}_{N_{FT},N_{Ft}} & \cdots & \cdots & \cdots \\ \sigma_{MT,Ft}\mathbf{1}_{N_{MT},N_{Ft}} & \cdots & \cdots & \cdots \end{bmatrix}$$

where $\mathbf{1}_{K,L}$ is a $(K \times L)$ matrix with unit and N_{FJ} is the number of observations for category F at time J. The $AsyVar(\hat{\beta}^{full})$ can be estimated consistently by taking the appropriate sub-block of a consistent estimate of (B.6) where the single components in Σ are obtained from the consistent estimates of the OLS residual from the full model: $\hat{\epsilon}^{full} = Y - X\hat{\beta}^{full}$, i.e. $\hat{\sigma}_{Ft}^2 = \frac{\hat{\epsilon}_{Ft}'\hat{\epsilon}_{Ft}}{N_{Ft}-K}$.

The estimation of the middle part of Term II in (B.4) can be obtained by using the consistent estimates of β^{full} and W, i.e. $\hat{\beta}^{full}$ and $\hat{W} = X_2 - X_1\hat{\Gamma}$. The estimation of terms III and IV requires the estimation of the covariance between $\sqrt{N}(\hat{\beta}^{full} - \beta^{full})$ and $\frac{X_1 W \beta^{full}}{\sqrt{N}}$. Given standard assumption on the error terms, the consistent estimation of the covariance is given by the columns corresponding to the variables X_2 the matrix below:

$$plim(\frac{X'X}{N})^{-1} plim(\frac{X^1 \epsilon^{\text{full}} \beta^{\text{full}'} W' X_1'}{N})$$

where β^{full}, ϵ^{full} and W are substituted by their corresponding consistent estimators.

C Unconditional Quantile Regression

To be precise, the RIF-OLS regression model allows us to estimate the effect of explanatory variables X on the unconditional quantile Q_τ of an outcome variable Y. The RIF is estimated in quantile regressions by first calculating the sample quantile $\hat{Q}_\tau$ and computing the density at $\hat{Q}_\tau$-that is, $f(\hat{Q}_\tau)$-using kernel methods Firpo *et al.* (2009).

This approach relies on the indicator function $\mathbb{1}\{Y_t \leq Q_\tau\}$, which takes the value of one if the condition in $\{\cdot\}$ is true, and zero otherwise. Estimates for each observation i of the RIF $\widehat{RIF}(Y_{it}; Q_\tau)$ are then obtained by inserting $\hat{Q}_\tau$ and $f(\hat{Q}_\tau)$ in the aggregate RIF function, defined as:

$$\begin{aligned} RIF(Y_t; Q_\tau) &= Q_\tau + IF(Y_t; Q_\tau) \\ &= Q_\tau + \frac{\tau - \mathbb{1}\{Y_t \leq Q_\tau\}}{f_Y(Q_\tau)} \\ &= \frac{1}{f_{Y_t}(Q_\tau)}\mathbb{1}\{Y_t > Q_\tau\} + Q_\tau - \frac{1}{f_{Y_t}(Q_\tau)}(1-\tau) \end{aligned} \tag{E.1}$$

where the RIF is the first-order approximation of the quantile Q_τ, and $IF(Y_t; Q_\tau)$ represents the influence function for the τth quantile. It measures the (marginal) influence of an observation at Y on the sample quantile. Adding the quantile Q_τ to the influence function yields the RIF. The probability density of Y at time t is evaluated at Q_τ is $f_{Y_T}(Q_\tau)$. The model can then be estimated by OLS using the RIFs as dependent variables. 132009Firpo *et al.*Firpo, Fortin and Lemieux () modelled the conditional expectation of the RIF-regression function $E[RIF(Y_t; Q_\tau)|X]$ as a function of explanatory variables X in the UQR:

$$E[RIF(Y_t; Q_\tau)|X] = g_{Q_\tau}(X) \tag{E.2}$$

where a linear function $X\beta_\tau$ is specified for $g_{Q_T}(X)$. The explanatory variables X contain time-varying controls like labour market experience and job tenure as well as time-constant controls like education. The average derivative of the unconditional quantile regression $E_X\left[\frac{dg_{Q_T}(X)}{dX}\right]$ captures the marginal effect of a small location shift in the distribution of covariates on the τth UQ of Y_t, keeping everything else constant. Therefore, the coefficients β_τ can be unconditionally interpreted as $E[RIF(Y_t; Q_\tau)] = E_X\big[E\big(RIF(Y_t; Q_\tau)|X\big)\big] = E(X)\beta_\tau$. That is, the unconditional expectations $E[RIF(Y_t; Q_\tau)]$ using the LIE allow for the interpretation of the unconditional mean: On the other hand, the interpretation of the conditional mean is valid only in the context of CQRs: $Q_\tau(Y_t|X) = X\beta_\tau^{CQR}$, where β_τ^{CQR} can be interpreted as the effect of X on the τth CQ of Y given X. The LIE does not apply here; $Q_\tau \neq E_X[Q_\tau(Y_t|X)] = E(X)\beta_\tau^{CQR}$, where Q_τ is the UQ. Hence, β_τ^{CQR} cannot be interpreted as the effect of increasing the mean value of X in the UQ Q_τ. In UQR, the coefficients β_τ can be estimated by OLS in the following way:

$$Q_\tau = E[RIF(Y_t; Q_\tau)] = E_X[RIF(Y_t; Q_\tau)|X] = E(X)\beta_\tau \tag{E.3}$$

D Solving the Index Number Problem for the Level of the Gender Pay Gap

Section 5 presents the solution to the indeterminacy problem for the variation over time of the GPG. This Appendix shows how to solve the indeterminacy problem for the level of the GPG within the OVB decomposition. Our aim is to have a wage decomposition invariant to the reference category adopted. Following Fortin (2008), we include gender intercept shifts along with an identification restriction, in the regression of females and males pooled together, when considering the standard case of the GPG:

$$y_i = \gamma_0 + \gamma_{0F}F_i + \gamma_{0M}M_i + X_i\gamma + \epsilon_i$$

subject to:

$$\gamma_{0F} + \gamma_{0M} = 0$$

where F_i (M_i) is equal to one if the individual is female (male) and zero otherwise. The identification restriction, $\gamma_{0F} + \gamma_{0M} = 0$, imposes that the pooled wage equation truly represents a non-discriminatory wage structure, i.e. a wage structure where the advantage of men is equal to the disadvantage of women:

$$\bar{y}_M - \bar{y}_F = (\bar{X}_M - \bar{X}_F)\hat{\gamma} + (\hat{\gamma}_{0M} - \hat{\gamma}_{0F}) \tag{D.1}$$

The first component on the RHS, $(\bar{X}_M - \bar{X}_F)\hat{\gamma}$, is the explained part, while $\hat{\gamma}_{0M}$ and $\hat{\gamma}_{0F}$ are the *advantage of men* and the disadvantage of women, respectively. In particular:

$$\begin{aligned} \hat{\gamma}_{0M} &= \bar{X}_M(\hat{\beta}_M - \hat{\gamma}) + (\hat{\alpha}_M - \hat{\gamma}_0) \qquad \textit{advantage of men} \\ \hat{\gamma}_{0F} &= \bar{X}_F(\hat{\beta}_F - \hat{\gamma}) + (\hat{\alpha}_F - \hat{\gamma}_0) \qquad \text{disadvantage of women.} \end{aligned}$$

where $\hat{\alpha}_M, \hat{\alpha}_F, \hat{\beta}_M, \hat{\beta}_F$ are the estimated coefficients of the wage equations for men and women, respectively:

$$y_{iM} = \alpha_M + X_M\beta_M + \epsilon_{iM} \tag{D.2}$$

$$y_{iF} = \alpha_F + X_F\beta_F + \epsilon_{iF} \tag{D.3}$$

In order to recast the wage decomposition of the *full model* with the conditional decomposition framework proposed in Section 3 we estimate the following wage equation:

$$y_i = \gamma_0 + \gamma_{0F}F_i + \gamma_{0M}M_i + X_i\gamma + X_iF_i\gamma_{XF} + X_iM_i\gamma_{XM} + \nu_i \tag{D.4}$$

subject to:

$$\begin{aligned} \gamma_{0F} + \gamma_{0M} &= 0 \\ \gamma_{X_kF} + \gamma_{X_kM} &= 0 \text{ for } k = 1 \dots K \end{aligned}$$

where γ_{X_kF} and γ_{X_kM} are the parameters of the interaction term between the *kth* regressor X_k and the dummy F and M, respectively. The error term is represented by v_i. Evaluating equation (D.4) at the mean yields:

$$\begin{aligned}\bar{y}_M &= \hat{\gamma}_0 + \hat{\gamma}_{0M} + \bar{X}_M\hat{\gamma} + \bar{X}_M\hat{\gamma}_{XM} \\ \bar{y}_F &= \hat{\gamma}_0 + \hat{\gamma}_{0F} + \bar{X}_F\hat{\gamma} + \bar{X}_F\hat{\gamma}_{XF}\end{aligned}$$

Hence, the GPG is given by:

$$\bar{y}_M - \bar{y}_F = (\hat{\gamma}_{0M} - \hat{\gamma}_{0F}) + (\bar{X}_M - \bar{X}_F)\hat{\gamma} + \bar{X}_M\hat{\gamma}_{XM} - \bar{X}_F\hat{\gamma}_{XF} \tag{D.5}$$

$$= 2\hat{\gamma}_{0M} + (\bar{X}_M - \bar{X}_F)\hat{\gamma} + (\bar{X}_M + \bar{X}_F)\hat{\gamma}_{XM} \tag{D.6}$$

First, we observe that there exists the following relationship between the parameter estimates of equations (D.2)-(D.3) and (D.4):

$$\begin{aligned}\hat{\gamma} - \hat{\gamma}_{XM} &= \hat{\beta}_F \\ \hat{\gamma}_0 - \hat{\gamma}_{0M} &= \hat{\alpha}_F \\ \hat{\gamma} + \hat{\gamma}_{XM} &= \hat{\beta}_M \\ \hat{\gamma}_0 + \hat{\gamma}_{0M} &= \hat{\alpha}_M\end{aligned}$$

Therefore, the GPG of (D.6) can be re-written in terms of the Fortin-decomposition as:

$$\bar{y}_M - \bar{y}_F = (\hat{\alpha}_M - \hat{\gamma}_0) - (\hat{\alpha}_F - \hat{\gamma}_0) + (\bar{X}_M - \bar{X}_F)\hat{\gamma} + \bar{X}_M(\hat{\beta}_M - \hat{\gamma}) - \bar{X}_F(\hat{\beta}_F - \hat{\gamma}) \tag{D.7}$$

$$= (\bar{X}_M - \bar{X}_F)\hat{\gamma} + \underbrace{[\bar{X}_M(\hat{\beta}_M - \hat{\gamma}) + (\hat{\alpha}_M - \hat{\gamma}_0)]}_{\text{advantage of men}} - \underbrace{[\bar{X}_F(\hat{\beta}_F - \hat{\gamma}) + (\hat{\alpha}_F - \hat{\gamma}_0)]}_{\text{disadvantage of women}} \tag{D.8}$$

Second, the estimation can be recast in terms of sequential decomposition by considering the following base model:

$$y_i = \gamma_0^{base} + (M_i - F_i)\gamma_{0M}^{base} + \epsilon_i^{base} \tag{D.9}$$

where the set of regressors of the base model is given by $X_1 = \Big[1, (M - F)\Big]$, the constant and the difference between the two dummy variables F and M. The full model is defined as follows:

$$y_i = \gamma_0^{full} + (M_i - F_i)\gamma_{0M}^{full} + X_i\gamma + X_i(M_i - F_i)\gamma_{XM} + \epsilon_i^{full} \tag{D.10}$$

where $X_2 = \Big[X, X(M-F)\Big]$. $X(M-F)$ is the interaction between the matrix of regressors X and the vector that contains the difference between the two dummy variables M and F. By the OVB formula the following relationship holds:

$$\begin{bmatrix} \hat{\gamma}_0^{base} \\ \hat{\gamma}_{0M}^{base} \end{bmatrix} = \begin{bmatrix} \hat{\gamma}_0^{full} \\ \hat{\gamma}_{0M}^{full} \end{bmatrix} + (X_1'X_1)^{-1}X_1'X_2 \begin{bmatrix} \hat{\gamma} \\ \hat{\gamma}_{XM} \end{bmatrix} \tag{D.11}$$

where $(\hat{\gamma}_0^{base}\ \hat{\gamma}_{0M}^{base})'$ is the vector of coefficient estimates of X_1 from the base model (D.9); $(\hat{\gamma}_0^{full}\ \hat{\gamma}_{0M}^{full})'$ is the vector containing the coefficient estimates of X_1 from the full model (D.10) and $(\hat{\gamma}\ \hat{\gamma}_{XM})'$ is the vector of coefficients estimates of X_2 from the full model (D.10). Observe that:

$$\begin{bmatrix} \hat{\gamma}_0^{base} \\ \hat{\gamma}_{0M}^{base} \end{bmatrix} = \begin{bmatrix} \frac{\bar{y}_M + \bar{y}_F}{2} \\ \frac{\bar{y}_M - \bar{y}_F}{2} \end{bmatrix} \tag{D.12}$$

and $\hat{\gamma}_{0M}^{full}$ is equal to $\frac{\hat{\alpha}_M - \hat{\alpha}_F}{2}$.

Given (D.12), our interest relies on the second row of equation (D.11), that represents the decomposition of the GPG. We observe that the linear projection of X with respect to X_1 is equal to:

$$(X_1'X_1)^{-1}X_1'X = \begin{bmatrix} / \\ (\bar{X}_M' - \bar{X}_F')/2 \end{bmatrix}$$

The linear projection of $X(M-F)$ with respect to X_1 is equal to:

$$(X_1'X_1)^{-1}X_1'X(F-M) = \begin{bmatrix} / \\ (\bar{X}_M' + \bar{X}_F')/2 \end{bmatrix}$$

Given (D.11), we observe that:

$$
\begin{aligned}
2\hat{\gamma}_{0M}^{base} &= 2(\bar{y}_M - \bar{y}_F) = \Delta(y) = 2\hat{\gamma}_{0M}^{full} + (\bar{X}_M - \bar{X}_F)\hat{\gamma} + (\bar{X}_M + \bar{X}_F)\hat{\gamma}_{XM} \\
&= 2\hat{\gamma}_{0M}^{full} + (\bar{X}_M - \bar{X}_F)\hat{\gamma} + \bar{X}_M(\hat{\beta}_M - \hat{\gamma}) + -\bar{X}_F(\hat{\beta}_F - \hat{\gamma}) \\
&= (\bar{X}_M - \bar{X}_F)\hat{\gamma} + \underbrace{[\bar{X}_M(\hat{\beta}_M - \hat{\gamma}) + (\hat{\alpha}_M - \hat{\gamma}_0)]}_{\text{advantage of men}} - \underbrace{[\bar{X}_F(\hat{\beta}_F - \hat{\gamma}) + (\hat{\alpha}_F - \hat{\gamma}_0)]}_{\text{disadvantage of women}}
\end{aligned}
$$

that completes the proof of the decomposition equivalence.

D.1 Invariance Decomposition with respect to Categorical Variables

A second type of identification issue arises when dummy variables are considered in the wage decomposition. Oaxaca and Ransom (1999) show that the assignment of the unexplained part of the GPG to specific variables is not invariant to the choice of reference groups. This problem can be easily solved by imposing the following parameters restrictions as proposed by Gardeazabal and Ugidos (2004) and Yun (2005):

$$\sum_{j=1}^{C_k} \gamma_{jk} = 0, \;\; k \in C \tag{D.13}$$

where C denotes the set of categorical variables, and C_k the number of categories for variable k. The *neutral*, i.e. non-sensitive to any left-out category, Oaxaca-Blinder decomposition follows. The zero-sum restriction on the coefficients for the single categories lead to express the effects as deviations from the grand mean. The zero-sum restriction (D.13) is applied to the wage equation, when female and male wages are estimated separately as well as to the pooled regression with gender dummies. The latter is additionally estimated with the identification restriction $\gamma_{0M} + \gamma_{0F} = 0$ on the gender parameters. Thereby, the intercepts, α_M, β_{0F} and γ_0, are no longer influenced by the choice of the reference category and the single parts of the endowments effect can be associated to specific covariates (Fortin, 2008). The restriction (D.13) can also be applied to the method proposed in Section 3 leading to indicator variables that, in case of categorical variables, are invariant to the choice of the left-out category.

E Intercept-shift Approach versus Pooled-sample Approach

The Intercept-shift Approach versus the Pooled-sample Approach Lee (2015) showed that the intercept-shift approach proposed by Fortin (2008) presents two drawbacks. First, the reference parameter for the OB decomposition, that is, the parameter that would prevail in a *fair* world with no discrimination, relies on the difference in the variance among categories. Second, the reference intercept is arbitrary: the same OB decomposition holds with vastly different reference intercepts. However, our proposed decomposition does not suffer from any of these issues, as it arises from a specification that allows different intercepts and slopes. In addition, the constraints imposed on the parameters that identify the counterfactual reference parameters are such that men's advantage is equal to women's disadvantage. In fact, in our model the slope that would prevail under no discrimination, γ, is the sample average of the group slopes: α_M and α_F:

$$\gamma = 0.5\alpha_M + 0.5\alpha_F$$

That is, it is equivalent to considering the weights proposed by Reimers (1983).[12] Moreover, the constraint:

$$\alpha_F - \gamma_{0F} = \alpha_M + \gamma_{0F}$$

prevents the indeterminacy problem shown by Lee (2015) in eq. (6) page 74 . It turns out that, in our model, the intercept indeterminacy problem Lee (2015) highlighted is ruled out by imposing the constraint that men's advantage should be equal women's disadvantage.

[12]See also Lee, 2015, p.72.

F Further Empirical Results

Table E.1 presents the detailed decomposition of the standard approaches: the OB (Panel A) and Fortin's regression-compatible decomposition (Panel B). The results show that gender differences in education are always higher for men than they are for women, while differences in labour market experience and job tenure are higher for men at all points of the wage distribution. Again, the standard approach reveals no substantial changes in the explained components from 2005 to 2016 in Italy. The unexplained components' point estimates suggest more differences over time; in particular, the remuneration scheme between men and women changed with respect to human capital and labour market characteristics.

Table E.1: Detailed Decomposition of the Gender Pay Gap in 2005 & 2016 (Case 1) at the Mean & Selected Percentiles, Standard Approach

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2005	2016	2005	2016	2005	2016	2005	2016
Percentile	Mean		10.		50.		90.	
GPG	0.155***	0.087***	0.087***	0.029**	0.123***	0.071***	0.258***	0.132***
	(0.008)	(0.009)	(0.009)	(0.013)	(0.009)	(0.007)	(0.018)	(0.013)
Panel A: Decomposition with Male-Reference Category								
Explained:								
X_HC	-0.023***	-0.023***	-0.007***	-0.011***	-0.016***	-0.017***	-0.045***	-0.041***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.005)	(0.004)
X_LM	0.040***	0.036***	0.030***	0.040***	0.038***	0.028***	0.058***	0.018***
	(0.004)	(0.004)	(0.004)	(0.006)	(0.003)	(0.003)	(0.008)	(0.006)
X_Demo	-0.003**	-0.004***	-0.003*	-0.008***	0.002	-0.004***	-0.010***	-0.002
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.003)	(0.002)
X_OccInd	-0.009*	-0.021***	-0.007	-0.022***	-0.010**	-0.018***	-0.016	-0.029***
	(0.005)	(0.005)	(0.006)	(0.007)	(0.005)	(0.003)	(0.013)	(0.008)
Total	0.008	-0.010	0.015**	-0.007	0.011*	-0.012***	-0.007	-0.036***
	(0.007)	(0.007)	(0.007)	(0.009)	(0.006)	(0.005)	(0.015)	(0.010)

Unexplained:								
X_HC	-0.056	-0.132**	-0.300***	-0.248***	-0.218***	-0.122***	0.310***	0.205**
	(0.039)	(0.058)	(0.075)	(0.092)	(0.044)	(0.042)	(0.095)	(0.087)
X_LM	0.032	0.037	0.011	-0.016	0.053**	0.025	0.053	0.009
	(0.021)	(0.024)	(0.030)	(0.041)	(0.021)	(0.018)	(0.059)	(0.050)
X_Demo	0.001	0.004	-0.006*	0.004	0.007***	0.005**	0.003	0.008*
	(0.002)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)	(0.005)	(0.004)
X_OccInd	-0.029*	-0.030	-0.055*	-0.058	0.007	-0.018	0.005	0.001
	(0.016)	(0.022)	(0.034)	(0.036)	(0.019)	(0.016)	(0.036)	(0.031)
Constant	0.172***	0.233***	0.484***	0.409***	0.289***	0.194***	-0.251**	-0.194*
	(0.047)	(0.067)	(0.091)	(0.107)	(0.053)	(0.049)	(0.113)	(0.101)
Total	0.147***	0.096***	0.105***	0.036**	0.112***	0.084***	0.265***	0.167***
	(0.008)	(0.009)	(0.014)	(0.015)	(0.009)	(0.007)	(0.020)	(0.015)
Panel B: Regression-Compatible Decomposition								
Explained:								
X_HC	-0.024***	-0.026***	-0.014***	-0.017***	-0.021***	-0.020***	-0.039***	-0.037***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)
X_LM	0.040***	0.030***	0.031***	0.041***	0.038***	0.022***	0.056***	0.021***
	(0.003)	(0.003)	(0.003)	(0.005)	(0.003)	(0.002)	(0.006)	(0.004)
X_Demo	-0.004***	-0.004***	-0.007***	-0.008***	0.001	-0.003***	-0.006***	-0.001
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
X_OccInd	-0.012***	-0.018***	-0.013**	-0.024***	-0.014***	-0.018***	0.001	-0.016***
	(0.004)	(0.004)	(0.006)	(0.006)	(0.004)	(0.003)	(0.009)	(0.006)
Total	0.002	-0.012**	0.002	-0.013*	0.007	-0.015***	0.008	-0.019**
	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)	(0.004)	(0.011)	(0.008)
Unexplained:								
X_HC	-0.055	-0.128**	-0.293***	-0.242**	-0.213***	-0.119***	0.304***	0.201***
	(0.039)	(0.053)	(0.078)	(0.099)	(0.044)	(0.041)	(0.087)	(0.072)
X_LM	0.031	0.043	0.010	-0.017	0.052**	0.031	0.055	0.006
	(0.024)	(0.027)	(0.035)	(0.049)	(0.022)	(0.019)	(0.061)	(0.048)
X_Demo	0.002	0.003	-0.002	0.004	0.007***	0.004**	-0.002	0.007**
	(0.002)	(0.002)	(0.003)	(0.004)	(0.002)	(0.002)	(0.005)	(0.003)

X_OccInd	-0.026**	-0.033	-0.050**	-0.056	0.010	-0.018	-0.012	-0.012
	(0.012)	(0.021)	(0.022)	(0.052)	(0.019)	(0.013)	(0.024)	(0.019)
Constant	0.172***	0.233***	0.484***	0.409***	0.289***	0.194***	-0.251**	-0.194**
	(0.049)	(0.067)	(0.097)	(0.130)	(0.054)	(0.047)	(0.104)	(0.083)
Total	0.153***	0.099***	0.118***	0.042***	0.117***	0.086***	0.250***	0.150***
	(0.007)	(0.009)	(0.014)	(0.014)	(0.008)	(0.006)	(0.018)	(0.014)

9,185 observations in 2005 & 10,148 observations in 2016. Robust standard errors in parentheses. X represents the set of observable characteristics included in the regression, XY is the interaction of the set of covariates used with the time indicator. fmX and $fmXY$ are the corresponding interactions with the $(F-M)$ gender dummy. The set of covariates X is split in the following way: HC contains years of schooling, LM includes labor market experience, its square, job tenure as well as firm-size dummies, $Demo$ includes a dummy for being married and place of residence (North & Centre) and $OccInd$ occupational and sectoral dummies. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table E.2: Detailed Decomposition of the Private-Public Sector Wage Gap by Gender (Case 2) at the Mean & Selected Quantiles, Standard Approach

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Men	Women	Men	Women	Men	Women	Men	Women
Percentile	Mean		10.		50.		90.	
PPWG	-0.210***	-0.252***	-0.210***	-0.252***	-0.210***	-0.252***	-0.210***	-0.252***
	(0.012)	(0.013)	(0.012)	(0.013)	(0.012)	(0.013)	(0.012)	(0.013)
Panel A: Decomposition with Male-Reference Category								
Explained:								
X_HC	-0.027***	-0.027***	-0.027***	-0.027***	-0.027***	-0.027***	-0.027***	-0.027***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
X_LM	-0.084***	-0.098***	-0.084***	-0.098***	-0.084***	-0.098***	-0.084***	-0.098***
	(0.007)	(0.010)	(0.007)	(0.010)	(0.007)	(0.010)	(0.007)	(0.010)
X_Demo	0.002	0.005	0.002	0.005	0.002	0.005	0.002	0.005
	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
X_OccInd	-0.026	-0.010	-0.026	-0.010	-0.026	-0.010	-0.026	-0.010

	(0.024)	(0.020)	(0.024)	(0.020)	(0.024)	(0.020)	(0.024)	(0.020)
Total	-0.134***	-0.131***	-0.134***	-0.131***	-0.134***	-0.131***	-0.134***	-0.131***
	(0.024)	(0.022)	(0.024)	(0.022)	(0.024)	(0.022)	(0.024)	(0.022)
Unexplained:								
X_HC	-0.052	0.014	-0.052	0.014	-0.052	0.014	-0.052	0.014
	(0.074)	(0.092)	(0.074)	(0.092)	(0.074)	(0.092)	(0.074)	(0.092)
X_LM	0.003	0.093**	0.003	0.093**	0.003	0.093**	0.003	0.093**
	(0.048)	(0.044)	(0.048)	(0.044)	(0.048)	(0.044)	(0.048)	(0.044)
X_Demo	0.029**	0.031	0.029**	0.031	0.029**	0.031	0.029**	0.031
	(0.014)	(0.020)	(0.014)	(0.020)	(0.014)	(0.020)	(0.014)	(0.020)
X_OccInd	0.055	0.207*	0.055	0.207*	0.055	0.207*	0.055	0.207*
	(0.090)	(0.121)	(0.090)	(0.121)	(0.090)	(0.121)	(0.090)	(0.121)
Constant	-0.109	-0.467***	-0.109	-0.467***	-0.109	-0.467***	-0.109	-0.467***
	(0.131)	(0.164)	(0.131)	(0.164)	(0.131)	(0.164)	(0.131)	(0.164)
Total	-0.076***	-0.121***	-0.076***	-0.121***	-0.076***	-0.121***	-0.076***	-0.121***
	(0.026)	(0.024)	(0.026)	(0.024)	(0.026)	(0.024)	(0.026)	(0.024)
Panel B: Regression-Compatible Decomposition								
Explained:								
X_HC	-0.029***	-0.027***	-0.029***	-0.027***	-0.029***	-0.027***	-0.029***	-0.027***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
X_LM	-0.089***	-0.095***	-0.089***	-0.095***	-0.089***	-0.095***	-0.089***	-0.095***
	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
X_Demo	-0.003	0.001	-0.003	0.001	-0.003	0.001	-0.003	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
X_OccInd	-0.037***	-0.031***	-0.037***	-0.031***	-0.037***	-0.031***	-0.037***	-0.031***
	(0.010)	(0.012)	(0.010)	(0.012)	(0.010)	(0.012)	(0.010)	(0.012)
Total	-0.157***	-0.153***	-0.157***	-0.153***	-0.157***	-0.153***	-0.157***	-0.153***
	(0.012)	(0.014)	(0.012)	(0.014)	(0.012)	(0.014)	(0.012)	(0.014)
Unexplained:								
X_HC	-0.050	0.015	-0.050	0.015	-0.050	0.015	-0.050	0.015
	(0.070)	(0.082)	(0.070)	(0.082)	(0.070)	(0.082)	(0.070)	(0.082)
X_LM	0.008	0.090**	0.008	0.090**	0.008	0.090**	0.008	0.090**

	(0.049)	(0.043)	(0.049)	(0.043)	(0.049)	(0.043)	(0.049)	(0.043)
X_Demo	0.034**	0.036	0.034**	0.036	0.034**	0.036	0.034**	0.036
	(0.015)	(0.024)	(0.015)	(0.024)	(0.015)	(0.024)	(0.015)	(0.024)
X_OccInd	0.065	0.229	0.065	0.229	0.065	0.229	0.065	0.229
	(0.070)	(0.154)	(0.070)	(0.154)	(0.070)	(0.154)	(0.070)	(0.154)
Constant	-0.109	-0.467**	-0.109	-0.467**	-0.109	-0.467**	-0.109	-0.467**
	(0.121)	(0.202)	(0.121)	(0.202)	(0.121)	(0.202)	(0.121)	(0.202)
Total	-0.052***	-0.099***	-0.052***	-0.099***	-0.052***	-0.099***	-0.052***	-0.099***
	(0.015)	(0.018)	(0.015)	(0.018)	(0.015)	(0.018)	(0.015)	(0.018)

9,185 observations in 2005 & 10,148 observations in 2016. Robust standard errors in parentheses. X represents the set of observable characteristics included in the regression, XY is the interaction of the set of covariates used with the time indicator. fmX and $fmXY$ are the corresponding interactions with the $(F-M)$ gender dummy. The set of covariates X is split in the following way: HC contains years of schooling, LM includes labor market experience, its square, job tenure as well as firm-size dummies, $Demo$ includes a dummy for being married and place of residence (North & Centre) and $OccInd$ occupational and sectoral dummies. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.