Economics Bulletin

Volume 39, Issue 4

The Public-Private Sector Wage Differential Across Gender in Italy: a New Quantile-Based Decomposition Approach.

Carolina Castagnetti
University of Pavia

Luisa Rosti
University of pavia

Marina Töpfer
University of Erlangen-Nürnberg

Abstract

We use a novel approach that combines quantile wage gap decomposition to standard technique for panel regression. We apply this methodology to investigate the public sector wage gap throughout the wage distribution in Italy by gender. Controlling for unobserved (time-invariant) individual heterogeneity suggests no positive selection effect in the public sector. The analysis reveals a consistent public sector premium throughout the wage distribution and across gender independent from individual endowments (both observed and unobserved).

Citation: Carolina Castagnetti and Luisa Rosti and Marina Töpfer, (2019) "The Public-Private Sector Wage Differential Across Gender in Italy: a New Quantile-Based Decomposition Approach.", *Economics Bulletin*, Volume 39, Issue 4, pages 2533-2539
Contact: Carolina Castagnetti - carolina.castagnetti@unipv.it, Luisa Rosti - luisa.rosti@unipv.it, Marina Töpfer - marina.toepfer@fau.de.
Submitted: May 27, 2019. **Published:** November 12, 2019.

1 Introduction

Related literature on the public-private pay differential has shown that, in general, the wage gap tends to be higher for women and typically narrows as one moves up the earning distributions (e.g. Lucifora and Moeurs, 2006). Giordano et al. (2011) find different public-sector wage premia along the wage distribution and across gender. The premium is typically higher for female workers compared to their male counterparts. However, in countries with more pronounced wage compression like Italy, the premium across quantiles for women is flatter. There are several theoretical interpretations of differences in pay among sectors. Gregory and Borland (1999), among others, argue that these differences in wage structure are not surprising given that wage setting in the public sector occurs in a political environment whereas private-sector decision making occurs in a market environment.

We use a two-step procedure for computing a Oaxaca-Blinder-type decomposition (Oaxaca, 1973; Blinder, 1973) using panel data for Italy (see Castagnetti and Giorgetti, 2019) in order to assess how the public-private sector wage gap (PPWG) varies along the wage distribution. First, we estimate the PPWG on the pooled sample neglecting unobserved individual heterogeneity. Second, we estimate the PPWG by means of the approach of Canay (2011) and run Machado and Mata (2005) decompositions for quantile regressions. This allows us to account for individual (time-invariant) heterogeneity by fixed effects estimation along the wage distribution. As men and women may face different selection into public jobs, we conduct the analysis separately by gender.

In line with previous findings, we find different, though positive, PPWGs and thus sector-wage premia along the wage distribution (e.g. Lucifora and Moeurs, 2006). In particular, we show the presence of higher premia for female workers compared to their male counterparts. By decomposing the PPWG, we show that it is almost completely due to differences in returns at the lower part of the earnings distribution, while this component vanishes at the upper part of the distribution. The results change dramatically, when we take generally unobserved individual-specific heterogeneity into account. Differences in returns between the public and private sector become the most important drivers of the gap for both men and women at all parts of the wage distribution.

2 Econometric Methodology

We estimate the following wage equations using linear quantile regression for individual i, with $i = 1, \ldots, N$:

$$y_i = q_\theta(y_i|x_i) + u_{\theta i} = x_i'\beta_\theta + u_{\theta i} \tag{1}$$

where $q_\theta(y_i|x_i)$ is the *conditional quantile* of the dependent variable y (log wages), given the covariates x (individual characteristics). The distribution of the error term $u_{\theta i}$ is left unspecified and it is assumed that $q_\theta(u_{\theta i}|x_i) = 0$. We estimate the wage equation separately for the public and the private sector as well as for men and women separately at different quantiles, with $\theta = \{0.10, 0.25, 0.50, 0.75, 0.90\}$. The decomposition of the PPWG in an explained and an unexplained components is carried out using the simulation technique of Machado and Mata (2005) that generalizes the Oaxaca-Blinder (OB) decomposition to a quantile regression framework.[1] The advantage of the quantile decomposition is that we can

[1] The OB decomposition for the PPWG is given by

$$\bar{y}_{public} - \bar{y}_{private} = \bar{X}_{public}^{'}\hat{\beta}_{public} - \bar{X}_{private}^{'}\hat{\beta}_{private} = (\bar{X}_{public}^{'} - \bar{X}_{private}^{'})\hat{\beta}_{public} + \bar{X}_{private}^{'}(\hat{\beta}_{public} - \hat{\beta}_{private})$$

estimate the unexplained component of the wage gap across the distribution of income, that is, at any quantile of the wage distribution. While in the OB setting, the wage gap is divided by means of a counterfactual wage structure, the Machado and Mata (2005) decomposition is based on the construction of a counterfactual distribution of $y^{private}$, i.e. a distribution of what would be for an individual employed in the private sector if she or he had been rewarded according to the public sector pay schedule.

Let $S^G \in \{public, private\}$ represent public and private observations for $G = (male, female)$, so that we have samples $\left\{(y^G_{i,S}, x^G_{i,S}) : i = 1, \ldots, n^G_S\right\}$ for all populations S and gender G, and we can estimate $q_\theta(y^G_S)$ separately for each group.

Formally, the Machado-Mata approach to estimate the counterfactual distribution of $y^G_{private}$ can be summarized as follows:[2]

1. Draw a random sample θ^*_i, $i = 1, 2, \ldots, 5000$ from a uniform distribution $U[0, 1]$.

2. For each θ_i, estimate $\beta^G_{private}(\theta)$ and $\beta^G_{public}(\theta)$ as

$$\hat{\beta}^G_S(\theta^*_i) = \arg\min_{\beta\in\mathbb{R}^p} \sum_{j=1}^{n^G_S} \rho_{\theta^*_i}(y^G_{j,S} - x^{G\prime}_{j,S}\beta) \quad S = public, private$$

   using the public and private dataset, respectively.

3. Randomly draw $5,000$ private sector employees with replacement and use their characteristics ($x^G_{*private}$) to predict the wages using the estimated coefficients $\beta^G_{public}(\theta)$ generating a set of predicted wages, $\tilde{y}^G_{private}(\theta) = x^{\prime G}_{private}\hat{\beta}^G_{public}(\theta)$. The empirical c.d.f. of these values is the estimated counterfactual distribution, namely what an individual employed in the private sector would have earned if she or he were paid according to the public sector pay schedule. The counterfactual distribution is computed for males and females, respectively.

4. Then compare the counterfactual distribution with the empirical public and private distributions whose θquantiles are defined by $\hat{y}^G_{public}(\theta) = x^{G'}_{public}\hat{\beta}^G_{public}(\theta)$ and $\hat{y}^G_{private}(\theta) = x^{G'}_{private}\hat{\beta}^G_{private}(\theta)$, respectively.

For each gender, $G = (female, male)$, we decompose the PPWG as:

$$y^G_{public}(\theta) - y^G_{private}(\theta) = [\hat{y}^G_{public}(\theta) - \tilde{y}^G_{private}(\theta)] + [\tilde{y}^G_{private}(\theta) - \hat{y}^G_{private}(\theta)] + residual \qquad (2)$$

where $y^G_S(\theta)$ is the observed log wage for gender G in sector $S = (public, private)$ at θ, $\hat{y}^G_S(\theta)$ is the estimated log wage for gender G in sector S at θ, and $\tilde{y}^G_{private}(\theta)$ is the estimated

---

where $\bar{y}_{public}$ and $\bar{y}_{private}$ are the log hourly wages of individuals employed in the public and private sector evaluated at the mean, respectively, with $\bar{X}_S$ and $\hat{\beta}_S$ being vectors of average characteristics and estimated coefficients for $S = (public, private)$. The first term is known as the *characteristics effect* that evaluates the PPWG in terms of endowments at the rate of return of the public sector. As different endowments should have different effects on earnings, the difference in endowments represents the *explained component* of the OB decomposition. The second term is the *coefficients effect* evaluating the PPWG in terms of different returns for the private sector employees's characteristics. This part is called the *unexplained component* because of the traditional application of OB to gender wage gap decomposition.

[2] The decomposition proposed by Machado and Mata (2005) grounds on the probability integral transformation theorem from elementary statistics: if U is uniformly distributed on $[0, 1]$, then $F^{-1}(U)$ has distribution F. Thus, for a given x_i and a random $\theta \sim U[0, 1]$, $x'_i\beta(\theta)$ has the same distribution as $y_i|x_i$. If, instead of keeping x_i fixed, we draw a random x from the population, $x'\beta(\theta)$ as the same distribution of y.

counterfactual log wage at θ. The counterfactual represents the wage that an individual employed in the private sector would get, if she or he had been rewarded according to the public sector pay schedule. The residual term captures the changes unaccounted for by the estimation method. In particular the residual term comprises the simulation errors and the sampling errors. Both errors asymptotically vanish (Melly, 2005).

The first part of the wage differential is the so-called *characteristics* effect, as it is due to the different distribution of covariates in the two sectors. The second addend in (2) represents the so-called *coefficients* effect. It is obtained by evaluating private-sector characteristics using two different conditional distributions.[3]

2.1 Quantile Regression for Panel

To take into account the unobserved individual heterogeneity in explaining the wage gap across the distribution, we extend our empirical analysis by exploiting the longitudinal structure of the data. Therefore, we consider the following quantile regression fixed effect model (hereafter FE-QR):

$$Q_\theta(y_{it}|x_{it}) = \alpha_i + x'_{it}\beta_\theta \tag{3}$$

$$y_{it} = x'_{it}\beta_\theta + u_{\theta it} \tag{4}$$

While estimation methods for cross-sectional conditional quantile regression models are well developed, corresponding methods for panel data (especially FE models) have received attention only recently. One problem associated with FE-QR is that, as it is the case with non-linear panel data models, the method of differencing out the fixed effects used for the conditional linear mean model does not carry over to the conditional quantiles. Koenker (2004) proposes to treat each individual effect as a parameter to estimate[4] by means of a penalized estimation method. However, controlling fixed effects by directly estimating them is not without difficulty - known as incidental parameter problem (Neyman and Scott, 1948), which manifest itself in inconsistency of the common parameters when the number of individuals goes to infinity while the number of time period is fixed.[5]

A second problem arises because the objective function is not differentiable. The implication is that standard asymptotic analysis of panel data model is not directly applicable to QR. Kato and Galvao (2016) propose the smoothing of the objective function and study the properties of the estimator. They show that the estimator is asymptotically normally distributed and propose a bias correction for the estimator's mean. Flores et al. (2014) estimate a two-way fixed effects model where both effects vary over quantiles. Flores et al. (2014) account for the problem of *quantile crossing* adopting the method proposed by Chernozhukov et al. (2010) to transform the original estimated quantiles into monotonic ones. However, the objective function they consider is not smooth and they rely on a Monte Carlo experiment to show the small bias in their estimates. Alternative approaches that do not consider the case of unobserved heterogeneity represented by the classical individual effects α_i, are introduced by Harding and Lamarche (2014), who propose a quantile regression estimator for a model with a multifactor error structure and interactive effects potentially correlated with covariates.

[3] The standard errors are computed following Chernozhukov et al. (2013).

[4] The individual parameters are assumed to have a pure location shift effect on the conditional distribution of the dependent variable.

[5] The analysis of an incidental parameter problem in FE-QR is described in Graham et al. (2009) and Kato and Galvao (2016).

We follow the approach proposed by Canay (2011) under the assumption that the individual parameters (fixed effects) have a pure location shift effect. Canay (2011) proposes an easy-to-use two-step estimator. In the first step, the individual effects α_i are estimated by traditional mean estimations (for instance estimation in first differences or by means of the within estimator). Next, the corrected wages $\hat{y}_{it} = y_{it} - \hat{\alpha}_i$ are estimated on the other covariates using traditional quantile regression. Given $\hat{y}_{it}$, we estimate the wages by quantile regression and we apply the Machado and Mata (2005) method to decompose the wage gap in its components.

3 Data

The analysis is based on the panel data set drawn from the ISFOL-PLUS survey (2005-2010).[6] The dependent variable is the log-hourly net income (adjusted to the 2010 level). We select a group of about thirty independent variables, which include years of schooling, labor market experience and tenure, type of contract (part-time, short-term), family characteristics (civil status, presence of pre-schooling age children), occupation and industry dummies. Further, we add geographic variables (dummies for individuals living in Northern regions and dummies for individuals living in urban areas), as well as personal skills that may reveal individual ability such as knowledge of English and knowledge about how to use a computer for particular basic tasks. A first look[7] on the main variables considered shows that public-sector employees are, on average, better educated, possess better indicators of personal skills and have more years of experience and tenure than private-sector employees.

4 Pooled Analysis

The results for the PPWG are presented in Figure 1 for females and males, respectively. For both samples, we observe that there is a positive wage premium for those working in the public sector, which shrinks as we move towards higher quantiles of the income distribution. These results confirm the findings of Disney and Gosling (1998), Lucifora and Moeurs (2006) and Campos and Centeno (2012), among others. Disney (2007) considers several explanation for the observed differences in pay between the public and the private sectors. One justification often found in the literature for the PPWG is that governments are less competitively-driven than the private sector and more inclined to equity and fairness in wage settlements, which translates into higher earnings than market levels at the bottom and moderate remuneration to top level managers. Another line of argument concerns the potential existence of compensating differentials, i.e. in-kind advantages and fringe benefits that would offset pay differences. However, this type of gain is to be found mostly in the public sector, where employees benefit from job protection and more advantageous pension plans. Other structural differences, due to for instance fundamentally different occupational compositions between sectors, is controlled for by *Occupational dummies* included in the estimation. Bargain et al. (2018) explain the disparity in pay between the two sectors by means of the differences in workers' unobserved characteristics; they show that when unobserved heterogeneity is controlled for by standard panel techniques, the public-sector wage premium vanishes.

[6]Isfol: Istituto per lo Sviluppo della Formazione Professionale dei Lavoratori.

[7]Descriptive statistics on the variables used are available upon request.

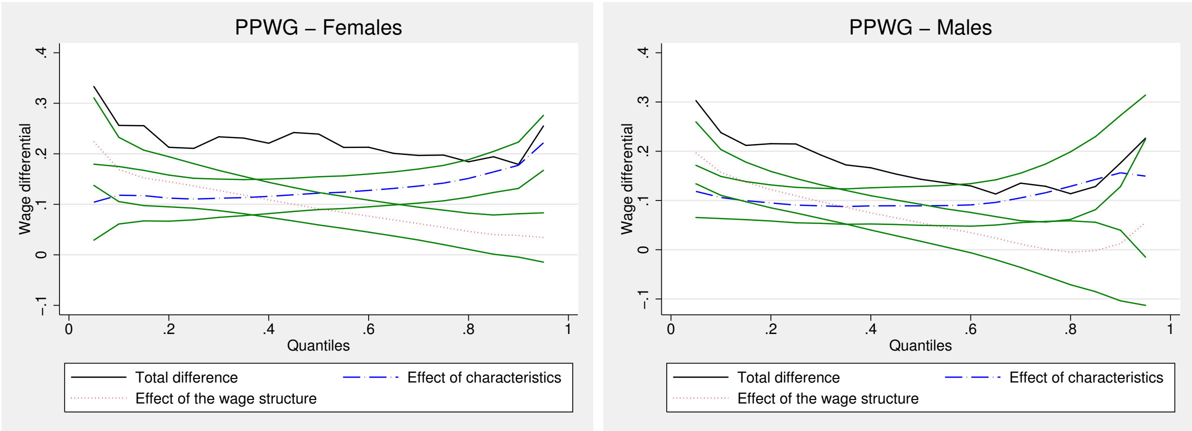

Figure 1: Public-Private sector wage gap decomposition by gender – Pooled sample

In terms of the decomposition of measurable attributes (*Effect of characteristics*) and differences in the return to the same attributes (*Effect of the wage structure or coefficients effect*)[8], we observe that the estimated unexplained component of the PPWG varies strongly with θ. In particular, in both samples the portion of the PPWG accounted for by differences in the wage structure declines monotonically, to almost zero, from lower to upper quantiles. That is, a significant portion of the differential, in the lower part of the wage distribution, can be accounted for by differences in returns, while in the upper part there are almost no differences. For males, the sharp decline is more evident, and the estimated public sector wage premium due to differences in returns becomes negative at top deciles, implying that there are significant differences in individual (observed) characteristics and occupations across sectors and quantiles.

5 Longitudinal Analysis

The analysis in Figure 1 shows that using QR and Machado-Mata decomposition both males and females receive a (slightly decreasing) positive premium in the public sector along the wage distribution. This result is in line with the results of Lucifora and Moeurs (2006) and Campos and Centeno (2012), among others. In order to better understand, whether these findings depend on a positive selection in the public sector, we control for endogenous selection using fixed effect estimation as proposed by Canay (2011) and then decompose the pay gap by the Machado and Mata (2005) method. Indeed, adding fixed effects to the wage equations allows to capture (time-invariant) individual heterogeneity that remains generally unobserved. Differences in the composition of the PPWG by gender may therefore reveal different sorting behavior in the public or private sector (Bargain et al., 2018).

In order to prevent any distortion due to switching among sectors, we compute the decomposition by restricting the sample to no-sector switchers, present in at least three waves of the data. In this way, we attenuate also the bias that could arise due to measurement errors due to misreported or miscoded values for the variable referring to working in the public sector.[9]

[8]This component is generally interpreted as the wage premium.

[9]We compute the decomposition also by restricting the sample to sector- switchers in only one direction. The results are similar to the results presented here.

Differently from what is shown by Bargain et al. (2018) and Campos and Centeno (2012) that provide evidence of a positive selection effect into public-sector jobs, we provide evidence of a public-sector premium independent from the individual abilities or endowments. The public premium does not arise because better-endowed individuals choose to work in the public rather than in the private sector. Instead, it is the difference in the wage structure that causes the premium. This effect is particularly pronounced for women and for both male and female low-income earners.

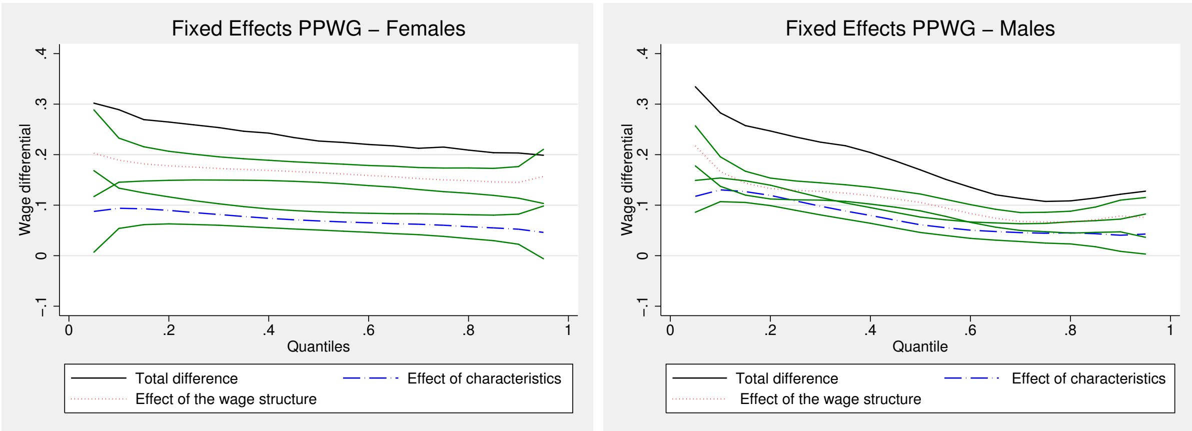

Figure 2: Public-Private sector wage gap decomposition by gender – FE-quantile estimation

6 Conclusion

In this paper, we analyze the wage gap between the public and private sector for Italian women and men. Using quantile regression methods, we perform the analysis for both cross-section and panel data. The cross-sectional (or pooled) analysis shows that there is a positive public-sector wage premium for men and women along the wage distribution. Both females and males are better paid in public-sector jobs. This effect is particularly pronounced at the lowest quantiles and higher for women than for men at all points of the distribution. Controlling additionally for unobserved heterogeneity reduces the public premium only slightly. We find no evidence of strong positive selection into the public sector. The results suggest that the positive public-private sector wage is mainly due to the different remuneration schemes across the two sectors.

References

Bargain, O., Etienne, A., and Melly, B. (2018). Public sector pay gaps over the long-run: Evidence from panel administrative data. Technical report, IZA DP No. 11924.

Blinder, A. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 2(8):436–455.

Campos, M. M. and Centeno, M. (2012). Public-private wage gaps in the period prior to the adoption of the euro: an application based on longitudinal data. Technical report, Banco de Portugal. Working Papers.

Canay, I. (2011). A simple approach to quantile regression for panel data. *The Econometrics Journal*, 14:368–386.

Castagnetti, C. and Giorgetti, M. (2019). Understanding the gender wage gap differential between public and private sector in italy: aquantile approach for panel data. *Economic Modelling*, 78:240–261.

Chernozhukov, V., Fernaandez-Val, I., and Galichon, A. (2010). Quantile and probability curves without crossing. *Econometrica*, 78:1093–1125.

Chernozhukov, V., Fernandez-Val, I., and Melly, B. (2013). Inference on counterfactual distributions. *Econometrica*, 81:2205–2268.

Disney, R. (2007). Public-private sector wage differentials around the world: methods and evidence. University of Nottingham.

Disney, R. and Gosling, A. (1998). Does it pay to work in the public sector? *Fiscal Studies*, 19:347–374.

Flores, C., Flores-Lagunes, A., and Kapetanakis, D. (2014). Lessons from quantile panel estimation of the environmental kuznets curve. *Econometric Reviews*, 8:815–853.

Giordano, R., Depalo, D., Pereira, C., Eugene, B., Papapetrou, E., Perez, J., Reiss, L., and Roter, M. (2011). The public sector pay gap in a selection of euro area countries. ECB Working Paper Series, No. 1406.

Graham, B., Hahn, J., and Powell, J. (2009). The incidental parameter problem in a non-differentiable panel data model. *Economics Letters*, 105:181–182.

Gregory, R. and Borland, J. (1999). *Handbook of Labor Economics*, chapter Recent developments in public sector labor markets, pages 3573–3630. O. Ashenfelter and D. Card.

Harding, M. and Lamarche, C. (2014). Estimating and testing a quantile regression model with interactive effects. *Journal of Econometrics*, 178:101–113.

Kato, K. and Galvao, A. (2016). Smoothed quantile regression for panel data. *Journal of Econometrics*, 193:92–112.

Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91:74–89.

Lucifora, C. and Moeurs, D. (2006). The public sector pay gap in france, great britain and italy. *Review of Income and Wealth*, 52:43–59.

Machado, J. and Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, 20:445–465.

Melly, B. (2005). Public-private sector wage differentials in germany: evidence from quantile regression. *Empirical Economics*, 30:505–520.

Neyman, J. and Scott, E. (1948). Consistent estimates based on partially consistent observations. *Econometrica*, 16:1–32.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 2(14):693–709.