Public-private sector pay gaps in Finland: A quantile regression analysis[*]

Abstract

This paper examines public-private sector wage differentials in Finland using a quantile regression method. We control for the endogeneity of the working sector and allow the returns of individual skills to vary between industries. The results suggest that men earn a premium of 3 percent in the public sector at the lower-end jobs. At the median and the upper end of the distribution, men's pay gap is negative, varying between 5 and 10 percent. Women, in turn, always earn more in the public sector (4-10 percent), and the premium is highest at the upper end of the earnings distribution.

[*] We thank an anonymous referee and the editor for helpful comments on a previous draft of this paper. The financial support of the Yrjö Jahnsson Foundation is gratefully acknowledged (grant number 5764).

1 Introduction

Wage differentials between workers in the public and private sectors have attracted a considerable amount of attention in empirical studies over the last three decades (see Disney 2007 for a survey). The focus of the research has been the treatment of the endogenous working sector, using the methods of Heckman (1979) and the decomposition of the observed pay gap into parts that are due to workers' characteristics and the rewards attributable to these characteristics (Oaxaca 1973, Neuman and Oaxaca 2004). Typically, employment in the public and private sectors has been found to be endogenously determined (e.g., Kanellopoulos 1997, Lassibille 1998, Adamchik and Bedi 2000, Christofides and Pashardes 2002, Tansel 2005). Although the results of decomposition analyses vary from one country to another, empirical findings have two common features. First, the unexplained pay gap is positive; public sector workers earn more than their private sector counterparts. Second, the pay gaps are higher for women than for men (e.g., Shapiro and Stelcner 1989, Kanellopoulos 1997, Prescott and Wandschneider 1999, Bender and Elliot 2002, Christofides *et al*. 2002, Heitmueller 2006).

Recent empirical research on pay gaps has utilised quantile regression (QR) methods that allow the pay gap to vary along the wage distribution (Mueller 1998, Blackaby, Murphy and O'Leary 1999, Melly 2005, Papapetrou 2006, Lucifora and Meurs 2006). The results indicate that the public sector pay premium is typically highest at the lower end of the wage distribution. These QR method studies, however, do not control for endogenous selection into the public and private sectors, which can potentially bias the estimated pay gaps.

Our wage gap analysis is based on longitudinal register-based data, consisting of 396,100 wage observations, from Statistics Finland for the period 1995-2004. This study aims to contribute to the existing literature in three ways. First, we analyse the pay gaps at different points of the wage distribution while simultaneously controlling for the endogenous selection of the working sector. We adopt the selection method proposed by Buchinsky (1998b, 2001), in which a polynomial of the basic selection term is employed in a regression of the QR model. The QR method itself is less restrictive than the mean regression, allowing for more a complex examination of returns to exogenous variables. Second, we allow for the possibility that returns to certain individual attributes (human capital) may vary across employees' work industries. In this respect, our study acknowledges empirical literature that reports considerable inter-industry wage differentials (e.g., Krueger and Summers 1988, Lucifora 1993, Gannon and Nolan 2004). This generalisation can be easily justified by assuming that certain skills that are essential, for example, in the production process, may not be valued in industries that manufacture different product or service lines. Third, we utilise a dummy coding to decompose the observed pay gaps into explained and unexplained parts (see Jann 2008 for a survey). This approach is a clear improvement over the basic decomposition because the coding method produces explicit estimates on the detailed contributions of the variables affecting the pay gap.

The rest of this paper is organised as follows. Section 2 describes the data and reports the average pay gaps across different earnings percentiles and between industries. Section 3 presents the econometric analysis of our study. Finally, Section 4 concludes the paper.

2 Data description and average pay gaps

The data used in the analysis are based on the various registers of Statistics Finland, including variables from the Longitudinal Census File and Longitudinal Employment Statistics from the 1970-2004 period. Data from various sources, including data on spouses and parents, have been merged using personal identifiers. The data represent a seven (7) percent random sample of the Finnish population in 2001 and comprise a comprehensive set of information on individual characteristics and the regions where the work places are located.

This study focuses on the post-recession period of 1995-2004. The dependent variable is the individual's annual wage and salary earnings (the logarithm), deflated in 2004 euros using the consumer price index. The data are truncated at the upper end of the earnings distribution because the reported annual earnings are capped at 72,000 (years 1995-2000) and 96,000 euros (years 2001-2004). The analysis is confined to those individuals who were full-year wage earners between 18 and 64 years of age with positive earnings. Self-employed individuals and individuals living in Åland were excluded from the analysis. The data are unbalanced panel data including 396,100 wage observations from 209,398 males and 186,702 females. Approximately 15 percent of males and 45 percent of females are public sector workers.

Table 1 reports the means of certain individual characteristics by working sector and sex. Similar to the existing literature, public sector workers are older and, in general, better educated than private sector workers (e.g., Christofides *et al*. 2002 and Tansel 2005). There is no clear discrepancy in average work experience, although public sector workers report a considerably

shorter mean duration in tenure, which reflects the use of temporary contracts among the public sector. Finally, a higher share of public sector workers are married or cohabiting, have children and are in upper-level occupations compared with private sector workers.

As Table 1 shows, the average annual earnings are slightly higher for employees in the private sector. This result applies for both sexes. However, the wage differential between the working sectors is not constant and varies at different wage percentiles as well by industry. These results are depicted in Table 2, which shows that workers are typically better off working in the public sector at the lower end of the distribution,[2] whereas the reverse is true at the upper end. For females, for example, the total average pay gap is positive at the 20th percentile (6%) and negative at the 80th percentile (-7%). This pattern is particularly clear in manufacturing, trade and finance & real estate for both sexes. The pay gap also varies by industry. The gaps are large and negative in agriculture, finance & real estate and construction. Transportation represents an opposite case, in which the gap is large and positive for both sexes. Although these findings are based on average wages, one obvious message emerges from the data: the public-private sector wage gap should be analysed by controlling for industry-specific differences in returns, and it should account for differences along the wage distribution.

[Tables 1-2 around here]

[2] We report the wages of the 20th and 80th percentiles instead of the 10th and 90th percentiles because the upper-end earnings are truncated by Statistics Finland, as noted above, and the lower-end earnings may be downward biased due to incomplete information on part-time employment.

3 Empirical analysis of pay gaps

3.1 Wage equations and decompositions

The quantile regression (QR) model introduced by Koenker and Bassett (1978) is more flexible than mean regression because it allows for the study of the effects of the covariates on the entire conditional earnings distribution. Decomposition calculated from the mean regression may show that the average public sector worker is paid economic rents when, in fact, the pay gap may be larger at the bottom of the wage distribution compared to the top of the same distribution (e.g., Mueller 1998). Therefore, we build our empirical analysis on the following earnings equation that is estimated by the QR method:[3]

(1) $\ln w_{it} = \beta_q Z_{it} + \varepsilon_{qit}$, $Quant_q(\ln w_{it} \mid Z_{it}) = \beta_q Z_{it}$, where

$$\beta_q Z_{it} = \sum_{j=1}^{J} \delta_{q,j} X_{j,it} + \sum_{k=1}^{K} \alpha_{q,k} Ind_{k,it} + \sum_{k=1}^{K} \sum_{l=1}^{L} \gamma_{q,kl} HC_{l,it} Ind_{k,it} + \theta_q \lambda_{it}$$

where w_{it} is the annual wage obtained by an individual *i* in year *t*. *Ind* is a categorical variable reflecting an individual's industry (agriculture, manufacturing, construction, trade, transportation, finance & real estate, education and health & social work). *X* stands for a vector of other observable characteristics. This includes work experience and its squared term, tenure,

[3] The most important properties of QR is summarised in Koenker (2001). Asplund (2010), who explores the sources of increased wage differentials in the Finnish private sector, also provides a helpful description of the QR method.

education years, field of education, marital status, presence of children, native language, socio-economic group, major province and year dummies. We consider the possibility that returns to certain individual attributes may vary between industries, and we augment the model with interactions between categorical variable *Ind* and a vector *HC*. The latter includes continuous variables of the Mincerian Human Capital theory (Mincer 1974), including work experience, tenure and education years. λ_{it} refers to the selectivity term that is of an unknown functional form for a specific quantile, which can be corrected in a semi-parametric fashion (see Buchinsky 1998b, 2001). Here, we adopt a simplified version of the approach, in which the standard Heckman selection term is estimated in the first stage by a probit model for public-private sector employment choice (e.g., Kanellopoulos 1997, Dustmann and van Soest 1998, Lassibille 1998, Adamchik *et al*. 2000, Christofides *et al*. 2002 and Tansel 2005). In the second stage, a polynomial of the estimated selection term is used as a regression in the QR model. $Quant_q(lnw_{it}|Z_{it})$ denotes the q*th* conditional quantile of a wage given variable vector *Z*. As noted earlier, we will focus on the 0.20th, 0.50th and 0.80th quantiles in the regression analyses. Finally, ε_{qit} is a random error term, and δ, α, γ and θ are the parameters to be estimated.

The wage equations are estimated separately for the public and private sectors and by sex. Following Neuman *et al*. (2004), the conditional wage gaps are calculated by decomposing the difference in observed mean log wages between the public sector (*pu*) and the private sector (*pr*) as follows:

$$(2)\quad \ln \overline{W}_{pu}^{q} - \ln \overline{W}_{pr}^{q} = (\overline{Z}_{pu} - \overline{Z}_{pr})^{q} \hat{\beta}_{pr}^{q} + \overline{Z}_{pu}^{q} (\hat{\beta}_{pu} - \hat{\beta}_{pr})^{q}$$

where the first term on the right-hand side of the equation captures the total differences in the individual's characteristics (explained part) weighted by the parameters from the model for the private sector (*pr).* The last term measures the gap that is due to differences in the parameters (unexplained part or estimated pay gap) weighted by the means of the public sector workers (*pu).*

The detailed contribution of a single variable or a set of variables is typically of special interest. However, the contributions to the unexplained part depend on arbitrary scaling, i.e., the choice of the omitted base category (see Jann 2008 for a survey). One solution, adopted here, is to estimate group models using the standard dummy coding and then to transform the coefficient vectors so that deviations from the grand mean are expressed and the coefficient for the base category is added.[4] In short, we explicitly estimate the extent to which the pay gap is due to the specific individual characteristics (parameter δ), selectivity (parameter θ), the basic return of working in a specific industry (parameter α), and how workers are rewarded for skills in these industries (parameter γ).

3.2 Wage determination by quantiles

Tables 3-4 report the results from our QR model. To save space, we do not report the parameter estimates of the categorical variables but simply show the importance of these controls by F-

[4] This decomposition method enables interactions between categorical and continuous variables. Therefore, we excluded all of the categorical and dummy variables from vector *HC*. The exclusion of these controls should not be a problem because the role of basic human capital variables (experience and education) accounts for a large share of an individual's wage determination process.

statistics in the lower parts of the tables. The selection term (Lambda) is statistically significant and negative for the public sector wage equation in each quantile and sex, the point estimate is approximately -0.30. This result indicates that the average pay is lower for those public sector workers who choose to work in the public sector. The selection term is positive for the private sector wage equation across different quantiles and sex. These results clearly suggest that public/private sector employment is not exogenously determined.[5]

The individual parameter estimates for wage equations are well defined and have the expected signs. For males, the return to an additional year of education is higher in the public sector than in the private sector (approximately 7% versus 5%, respectively). This result is qualitatively similar to that of Shapiro *et al*. (1989). We also find evidence, in line with Budria (2006), that a return to schooling for males is higher at the upper end of the distribution in both working sectors. For females, the numbers are opposite: the returns are, on average, lower in the public sector than in the private sector (approximately 5% versus 6%, respectively), which is in accordance with Shapiro *et al*. (1989) and Lassibille (1998).

[5] The results from the first-stage probit model are shown in Table A1 in the Appendix. The exclusion restriction variables are age and parent's socio-economic status. The estimates are in line with international findings. For example, age is positively correlated with public sector employment (Lassibille 1998, Adamchik *et al*. 2000 and Christofides *et al*. 2002). Occupation, industry and field of education have similarly expected signs, capturing existing employment differences between the public and private sectors (Kanellopoulos, 1997). In line with Lassibille (1998), public sector workers are more likely to be employed in lower productivity regions. The probability of being a public sector worker is higher among men whose parents have worked in general government, which is in agreement with Dustman *et al*. (1998) and Christofides *et al*. (2002).

In line with García-Pérez and Jimeno (2005), we find that for males, the return to tenure is higher in the public sector. The estimate is approximately 3-4 percent in the public sector and 1 percent in the private sector across different quantiles. For females, the return to tenure is approximately 1 percent in both sectors. Similar to the earlier literature, married or cohabiting men earn more than single men, and this wage effect is higher for public sector employees, as in Dustmann *et al*. (1998) and Adamchik *et al*. (2000). The presence of small children increases men's wages significantly for both sectors, with a higher effect for private sector employees. This result is in line with Kanellopoulos (1997). The reverse is true for females: being married or cohabiting and having children decreases women's wages significantly. This general finding is in line with other Finnish evidence (e.g., Napari 2008). The negative wage effect (around 2 %) is stronger for employees in the private sector.

Swedish-speaking men experience small earnings cuts (approximately 1%) in the private sector across different quantiles. In the public sector, they earn a small premium at the median. Swedish-speaking women employees, in turn, earn a negative pay premium in both sectors (approximately 1- 3%). Finally, the wage equations show that occupation, field of education, industry and regional attributes are important for pay determination in both sectors and sex; see the F-test results in the lower parts of the tables. The F-tests from the industry-level interaction terms support our hypothesis that returns to skills vary between industries. This effect is more significant in the public sector wage equations and among men.

[Tables 3-4 around here]

3.3 Decomposing the pay gaps

Figures 1-2 depict the decomposition results for men and women. The main finding is that the public sector employees are, on average, better off at the lower parts of the earnings distribution and worse off at upper end of the wage distribution; see the first bars in the figures. The average pay gap is positive at the lowest quantile (2 % for men and 5 % for women) and negative at the median (approximately -1 %) and at the highest quantile (-6 %). The pay gaps stemming from the differences in characteristics (explained part) are positive for men and negative for women at the median and at the highest quantile. At the lowest quantile, the differences are statistically zero; see the middle bars in the figures.

Our primary interest is in the unexplained parts of the total pay gaps; see the last bars in the figures. These results suggest that men in the public sector earn a positive pay premium at the lowest quantile (3 %), whereas the pay gap becomes negative at the median (-5 %) and at the upper part of the earnings distribution (-10 %). In sum, the public sector pay advantage is centred at those jobs where the pay is the lowest. These results are in line with those of Mueller (1998), Blackaby *et al*. (1999), Lucifora *et al.* (2006) and Papapetrou (2006). The comparable results for women show that public sector workers earn a positive pay premium in each quantile. The premium increases along the earnings distribution, with a 4 percent pay gap at the lowest quantile, 7 percent at the median and 10 percent at the highest quantile. These results, demonstrating that women are always overpaid in the public sector, are in accordance with those of Mueller (1998) and Lucifora *et al.* (2006), especially at those jobs where the pay is the highest, which is not expected by the other international findings.

[Figures 1-2 around here]

Tables 5-6 provide detailed information on the role of observable characteristics and their related returns in the wage gaps. The main finding is that employees in the public sector are better educated, are in upper-level occupations and have attained education in fields that pay more in all quantiles. However, they have less work experience and fewer tenure years, and they are employed in industries that generally pay less. The effect of industry is highly negative at the highest quantile for both sexes (approximately -11 %). The results confirm earlier findings that stress the role of individual attributes in explaining public-private pay differentials. The results also show that skills differ between industries. In particular, public sector workers are more skilled (in terms of work experience, tenure and education) in manufacturing and education, whereas private sector workers are more skilled in health & social work and trade. Private sector women are also more skilled in finance & real estate, whereas the reverse is true for men.

The estimated pay gaps for men are mainly due to four characteristics: education, occupation, field of education and industry. The public sector pays more for education, especially at those jobs where the pay level is the highest. However, private sector returns to occupation, field of education and industry exceed those in the public sector. The role of industry accounts for a considerable part of this finding (approximately -17%). The pay gaps for women are, in turn, mainly due to experience, education and industry. The public sector pays more for education at the highest quantile but pays less for experience in all quantiles, and this effect is higher at the lower end of the earnings distribution. The industry-affiliation follows the same pattern as for men. The within-industry results suggest that public sector employees are particularly worse off

in finance & real estate and better off in health & social work. The effects are stronger at the lowest quantile in both industries and higher for men.

[Tables 5-6 around here]

3.4 The role of selection bias and industry-specific prices

Finally, we examined the extent of possible bias arising in estimated pay gaps in a case where the interaction and selection terms are omitted from the wage equations. The results are depicted in Table 7. The first row reports the estimated pay gaps obtained from a basic QR model. In the second row, the model is augmented with a selection term, and in the third row, the model controls for the industry-specific interaction terms. The fourth row reports the result from a model that simultaneously controls both the endogenous selection and industry-specific interaction terms (i.e., the results from Tables 5 and 6). The general finding is that the omission of these two controls treats male and female workers differently. The overall results for males remain practically unchanged: the selection term slightly improves the relative position of public sector male workers, whereas industry-specific interaction terms decrease it. The results for females show more variability, although the role of industry-specific interaction terms follows the same pattern as for men. Our experiments indicate that exclusion of endogenous selection from the wage equation biases the pay gap estimates upward at the lower part of the wage distribution and downward at the upper part of the distribution. Because our basic model yields estimates that are generally in line with international evidence (i.e., the pay gaps are lower at higher parts of the earnings distribution), our tentative conclusion is that the estimated pay gaps

in previous international studies may be biased upward at the lowest quantile and biased downward at the highest quantile.

[Table 7 around here]

4 Conclusions

This study analyses public and private sector pay gaps using data for the period 1995-2004. We simultaneously control for three important factors in the analysis. First, we estimate earnings equations by a quantile regression method that allows the pay gap to vary along the earnings distribution. Second, we control for endogenous selection into the public and private sectors. Third, we consider the possibility that returns to skills differ between industries. By applying the new decomposition method by Jann (2008), we produce detailed estimates on the determinants of the pay gaps. The results imply that the omission of the endogenously determined working sector might bias numerical estimates. We find that the bias is particularly strong among female workers. The role of inter-industry variations in returns to skills is also important, and it decreases the estimated pay advantages for public sector employees.

When does it pay to work in the public sector? Our QR results suggest that public sector female employees earn a premium in each quantile, and the premium is largest at the highest wage levels. Our results for men suggest that employees in the public sector are better off at the lower parts of the earnings distribution and worse off at the higher pay levels. This finding partially explains why transitions from the public sector to the private sector increase at higher skill levels

(Borjas 2003) and why it is generally harder for the public sector to attract and retain highly skilled male workers by means of a wage policy (Lewis and Frank 2002). Our analysis suggests that the industry also matters. In particular, public sector workers are better rewarded in health & social work, and private sector workers are better rewarded in finance & real estate.

References

Asplund, R. (2010) "Sources of increased wage differentials in the Finnish private sector", *Finnish Economic Papers*, 23, 43-61.

Adamchik, V., and A. Bedi (2000). "Wage Differentials in the Public and the Private Sectors: Evidence from an Economy in Transition", *Labor Economics*, 7, 203-224.

Bender, K., and R. Elliot (2002). "The Role of Job Attributes in Understanding the Public-Private Wage Differential." *Industrial Relations*, 41, 407-421.

Blackaby, D., Murphy, P., and N. O'Leary (1999). "The Payment of Public Sector Workers in the UK; Reconciliation with North American Findings." *Economics Letters* 65, 239-243.

Borjas, G. (2003). "The Wage Structure and the Sorting of Workers into the Public Sector". In *For the People: Can We Fix Public Service?* 29-54. Eds. J. D. Donahue and J. S. Nye. Brookings Institution Press.

Buchinsky, M. (1998b). "The Dynamic of Changes in the Female Wage Distribution in the USA: A Quantile Regression Approach." *Journal of Applied Econometrics* 13, 1-30.

Buchinsky, M. (2001). "Quantile Regression with Sample Selection: Estimating Women's return to Education in the U.S." *Empirical Economics* 26, 87-113.

Budria, S. (2006). Schooling and the Distribution of Wages in the European Private and Public Sectors, MPRA papers, 90.

Christofides, L., and P. Pashardes (2002). "Self/Paid-Employment, Public/Private Sector Selection, and Wage Differentials." *Labour Economics* 9, 737-762.

Disney, R. (2007). Public-Private Sector Differentials Around the World: Methods and Evidence. Mimeo, University of Nottingham, School of Economics.

Dustmann, C., and A. van Soest (1998). "Public and Private Sector Wages of Male Workers in Germany." *European Economic Review* 42, 1417-1441.

Gannon, R., and B. Nolan (2004). "Inter-Industry Wage Differentials in Ireland." *The Economic and Social Review*, 35, 157-182.

García-Pérez, I., and J. Jimeno (2005). Public Sector Wage Gaps in Spanish Regions. Spanish Bank Working Papers No. 0526.

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

Heitmueller, A. (2006). "Public-Private Sector Pay Differentials in a Developed Scotland". *Journal of Applied Economics* 9(2), 295-323.

Jann, B. (2008). "The Blinder-Oaxaca Decomposition for Linear regression Models." *The Stata Journal* 8, 453-749.

Kanellopoulos, C. (1997). "Public-Private Wage Differentials in Greece." *Applied Economics* 29, 1023-1032.

Koenker, R., and G. Basset (1978). "Regression Quantiles." *Econometrica* 46, 33-50.

Koenker, R. (2001). "Linear Hypothesis: Regression (Quantile)." *International Encyclopedia of Social & Behavioral Sciences*, 8893–8899.

Krueger, A., and L. Summers (1988). "Efficiency Wages and Inter-Industry Wage Structure." *Econometrica* 56, 259-93.

Lassibille, G. (1998). "Wage Gaps Between the Public and Private Sectors in Spain." *Economics of Education Review* 17, 83-92.

Lewis, G., and Frank, S. (2002). "Who Wants to Work for the Government?" *Public Administration Review* 62, 395-404.

Lucifora, C. (1993). "Inter-Industry and Occupational Wage Differentials in Italy." *Applied Economics,* 25 , 1113-24.

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

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

Mincer, J. (1974). "Schooling, Experience and Earnings." NBER and Columbia University Press, New York.

Mueller, R. (1998). "Public-Private Wage Differentials in Canada: Evidence from Quantile Regressions." *Economics Letters* 60, 229-235.

Napari, S. (2008). Essays on Gender Wage Gap in Finland. Helsinki School of Economics. Acta Universitatis Oeconomicae Helsingiensis A-328.

Neuman, S., and R. Oaxaca (2004). "Wage Decompositions with Selectivity-Corrected Wage Equations: A Methodological Note." *Journal of Economic Inequality* 2, 3-10.

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

Papapetrou, E. (2006). "The Unequal Distribution of the Public-Private Sector Wage Gap in Greece: Evidence from Quantile Regression." *Applied Economics Letters* 13, 205-210.

Prescott, D., and B. Wandschneider (1999). "Public/Private Wage Differentials in Canada-Evidence from the 1991 and 1982 Surveys of Consumer Finance." *Applied Economics* 31, 723-731.

Shapiro, D., and M. Stelcner. (1989). "Canadian Public-Private Sector Earnings Differentials, 1970-1980." *Industrial Relations*, 28, 72-81.

Tansel, A. (2005). "Public-Private Employment Choice, Wage Differentials, and Gender in Turkey." *Economic Development & Cultural Change* 53, 453-477.