Journal of the Statistical and Social Inquiry Society of Ireland
Vol. XXXIX

Investigating the Public-Private Wage Gap in Ireland Using Data from the National Employment Survey 2007

Patrick Foley & Fiona O'Callaghan[1]

Central Statistics Office

(*read before the Society, 5 November 2009)*

***Abstract*:** This paper uses data from the 2007 National Employment Survey to analyse the public-private sector wage gap in Ireland. The purpose of the paper is to highlight the issues around attempting to establish a definitive measurement of the public-private sector wage differential. A number of different approaches are used to estimate the public-private wage gap, and the limitations of these various methods are highlighted. A range of plausible estimates of this wage gap are presented for 2007. In particular, this report investigates the effect of weighting on the estimates as well as exploring the impact that the inclusion of size of enterprise has on the results. These results show that, in general, there is a premium associated with working in the public sector when average estimates are calculated. When quantile regression is used to consider how the premium varies across the distribution of earnings, it was found that the premium was highest at the lower end of the earnings distribution and lowest at the top end. However, in most instances modelled, the public-private wage gap became a discount at the higher end of the earnings distribution. Generally, the exclusion of size of enterprise from the model increased the public-private wage gap by approximately 5 percentage points. Conducting the analysis on weighted data rather than unweighted data tended to increase the size of the wage gap, and this was particularly true for male employees. In this paper, we also analyse a sample that excluded personal and protective services employees. The exclusion of this particular sub-sample had the effect of considerably reducing the public-private wage gap, most noticeably for males. This impact was evident in both the weighted and unweighted analyses. In light of the difficulties highlighted in this report in comparing public and private sector wages, it should be noted that any type of regression or related analysis that attempts to directly compare earnings across the public and private sector is prone to oversimplification.

Keywords: weighted regression, survey weights, quantile regression, public sector wage differential
JEL Classifications: C21, C83, J31

1. INTRODUCTION

The National Employment Survey (NES) 2007 publication reported that, on average, public sector weekly earnings were 32.6% higher than the private sector, €847.17 in the public sector, compared with €639.05 in the private sector. However, these figures did not take into account the differences in characteristics of employees in both sectors. Sector of employment is not the only determinant of earnings; in this study, both the attributes of the employees (e.g. educational attainment, experience, hours worked etc.) and the characteristics of their employment (e.g. size of organisation) were used to further explore the wage differential between the two sectors.

[1] Corresponding author, e-mail: fiona.ocallaghan@cso.ie. The authors wish to acknowledge helpful comments from their colleagues in the CSO and also from Dr. Anthony Murphy, Department of Economics, University of Oxford & Hertford College. The views expressed and any errors or omission are those of the authors.

Using NES 2007 data, this paper highlights the issues and difficulties that exist in attempting to estimate the public-private pay differential using regression analysis. The hypothesis that there exists a constant wage differential in the Irish context, as implied by ordinary least square squares (OLS), Blinder-Oaxaca, and the propensity score matching (PSM) methods, between public and private pay is clearly rejected using quantile regression.

Furthermore, this paper illustrates that estimates of the wage differential can vary dramatically as a result of (i) the decisions taken by the researcher(s) to adjust for the inherent bias in the sample, (ii) the inclusion and specification of size of enterprise variable as an explanatory variable, (iii) the method of analysis adopted and (iv) the particular sub-sample of the NES data being analysed.

Taking all of this into account, we advise caution in attempting to estimate a single definitive "answer" for the average public-private sector wage gap and suggest a range, while not clear cut, might provide a more accurate picture. In this report, we present a range of plausible estimates that result from adopting a variety of approaches that are common in the international literature on this topic.

The next section outlines the context for this paper in the international and national literature. This is followed in section 3 by a description of the data. Section 4 outlines the methodology adopted, and section 5 presents the results. Finally, section 6 summarises the findings of the study.

2. LITERATURE REVIEW

Gregory and Borland (1999) provide a comprehensive review of the early international literature on the estimation of the public-private wage differential. They emphasise that the public sector pay gap estimates are sensitive to model specification and the specific subset of employees analysed.

In a European context, Lucifora and Meurs (2006) show that the hypothesis of a constant wage differential as implied by the OLS, Blinder-Oaxaca and PSM methods is rejected. Using quantile regression they show that the public-private wage gap is sensitive to the choice of quantile and that the pattern of the premia varies with gender and skill. They conclude that in France, Britain and Italy the public sector is found to pay more to low skilled workers than in the private sector while the reverse is true of higher skilled workers. The effects are more evident for females.

Lucifora and Meurs decompose the public-private gap into a part explained by the characteristics of the employee and into an unexplained part. They find that symmetrically, the unexplained part of the wage differential between the public and private sector decreases and becomes close to zero at the highest quantiles suggesting that differences in unobserved characteristics are more important at lower quantiles.

The issue of the public-private pay differential in an Irish context has been explored in a number of recent papers. Boyle, McElligot and O'Leary (2004) used an unweighted sub-sample from the European Community Household Panel Survey (ECHP) to conduct an analysis of the Irish public-private pay differential between 1994 and 2001. Using the ECHP household survey data and adopting OLS techniques, they estimated that public sector workers earned an unexplained premium of 13% over employees in the private sector. Boyle et al. pointed out that the premium was considerably higher than that estimated for other countries. Their quantile regression results showed the estimated premium was sensitive to the choice of quantile. An unusual outcome of their analysis in the international context was that female public sector workers enjoyed the same public sector pay premium as their male counterparts while the direct opposite is found in the international literature.

Murphy and Ernst & Young (2007) used the much larger 2003 National Employment Survey (NES) dataset to investigate the Irish public-private pay differential. The study used OLS, Blinder-Oaxaca, PSM and quantile regression approaches to estimate the public sector pay differential. They found that the OLS consistently under-estimated the public-sector pay gap whilst their preferred estimates from the Blinder-Oaxaca and PSM approaches were quite similar. Their study, largely based on unweighted NES data showed that the public sector pay differential was significantly different for males and females with females enjoying a significantly higher premium. The overall average estimated pay differential for all public sector employees was 10%. The estimated pay differential for public sector males was 6% and for public sector females was 15%. Quantile regression analysis estimated from the

10^{th} to the 90^{th} percentile at increments of 10% confirmed that the estimated pay differential varied across the earnings distribution and was an actual discount at the 90^{th} percentile.

Kelly, McGuinness and O'Connell, P. (2009a) used the NES data from March 2003 and March 2006 to estimate how the public-sector pay gap had changed over that period. Their analyses based on OLS regression estimated that the public-private pay differential had more than doubled from 10% in 2003 to 22% in 2006 and the earnings gap for males had increased even more dramatically in that period from 5% in 2003 to 23% in 2006. In this study the Blinder-Oaxaca estimates of the average pay-gap in 2006 of 18.4% was actually smaller than the OLS estimate. Kelly et al. (2009a) also conducted quantile regression analysis for the March 2006 NES data. For all employees the quantile analysis showed similar patterns to that found in Murphy et al. (2007). However when the results were broken down by gender female public sector workers experienced a smaller public-pay differential than their male counterparts at the majority of the percentiles estimated.

Kelly, McGuinness and O'Connell, P. (2009b) used NES data from March 2003 and October 2006 to conduct a sub-sectoral analysis of changes to the public-sector pay gap over that period. In this analysis the public sector pay premium was estimated as 14% in 2003 and 26% in 2006. This study adopts a PSM approach similar to that adopted by Murphy, Ernst & Young (2007). The study categorises the public sector into a number of sub-sectors and the analysis shows that there was substantial variation in the pay premium across the public sector. Their estimates show that the estimated average pay-gap was lowest in the Civil Service and Local Authorities (9.6% and 11.8%) and highest in Education and Security Services (52.6% and 30.7%). This analysis finds that the OLS and PSM estimates are similar. This finding is in contrast to the findings of Murphy, Ernst & Young (2007) who found that the OLS tended to under-estimate the pay gap relative to the Blinder-Oaxaca and PSM approaches.

The CSO (2009) study analysed data from the October 2007 NES. The CSO's Blinder-Oaxaca analysis of full-time permanent employees aged 25-59 yielded an average public sector wage differential of 12.6%, with a premium of 10.4% for males and 15.1% for females. Further analysis of the differential at differing points throughout the earnings distribution showed that the premium was largest at the lower end of the earnings distribution and generally decreased as earnings increased.

The recent Irish literature addressing the issue of the public-private pay differential has consistently shown that the average public sector pay differential in Ireland is positive and is large when compared to that estimated in other OECD countries using similar techniques. The estimated pay gap is not constant and varies throughout the earnings distribution, with public sectors workers at lower percentiles receiving much larger premia than those at higher percentiles.

However, the results from the recent Irish literature are less clear on a number of issues. Firstly the size of the average public sector premia varies according to the analysis. This is primarily due to the choice of the explanatory variables, the specification of particular variables, the particular sub-sample analysed, the methodological approach adopted and the decision to use either weighted or unweighted data.

The magnitude of the reported premia also varies simply because most studies report the estimated pay-gap in terms of the natural log of weekly earnings. Most models estimate the public-private pay-gap in terms of the log of weekly earnings; therefore to express the pay-gap in terms of weekly earnings the anti-log of the estimated regression coefficient must be calculated. For example, in Table A1 of the Appendix, the OLS regression yielded public sector coefficient of 0.096, this can be interpreted as a 9.6% premium on the log of weekly earnings, or as a 10.0% premium on weekly earnings. This point is trivial when one estimates a pay-gap of less than 10% however when the estimated pay-gap is 26% on the log of weekly earnings this is actually estimating a pay gap of approximately 30% in terms of weekly earnings. This study and the CSO (2009) study reported the estimated differential in terms of weekly earnings, i.e. on the level scale as opposed to the log scale.

The second major difference in the reported results in the Irish literature on public-private pay differential relates to the issue of the measurement of the pay gap for male and female workers separately. In line with the international literature, Murphy and Ernst & Young (2007) and the CSO (2009) estimated a higher average public sector pay average premium for female public sector workers compared to the estimates for male public sector workers. Using a much smaller sample Boyle et al.

estimated a higher pay gap for female public sector workers but they find that the difference is not statistically significant. Kelly et al. (2009a) on the other hand find that in 2006 male public sector workers receive a higher premium than their female counterparts.

The explanation for the differences in the estimates of the public-private pay-gap for male and female workers relates primarily to the use of either weighted or unweighted data. The public-private pay gap estimates using the weighted data suggest that the female and male public sector workers receive a similar premium over their counterparts in the private sector, see Kelly et al. (2009a). In contrast, analyses using the unweighted data suggest that the estimated public pay differential is much higher for female public sector workers, see Murphy and Ernst & Young (2007) and the CSO (2009).

The NES sample, like most structural business datasets, is inherently biased towards larger firms. There are potential biases within the data, arising from non-response and which are liable to vary by gender and across sectors. Some researchers use weighted data to control for the inherent bias in their sample. However the use of regression analysis on weighted data is a contested issue. Other strategies are available to control for bias in a sample. There are numerous problems associated with constructing accurate survey weights for use in regression analysis, see Gelman (2007), Fazio, Lam, and Ritchie (2006), Deaton (1997) and Winship and Radbill (1994).

In general, survey weights are constructed for the purposes of conducting specific, univariate analysis of the survey variables for which the survey instrument was designed. The function of these survey weights is to provide accurate population estimates of simple descriptive statistics such as cross-tabulations and means. Preserving complex econometric relationships within the population is not factored into the construction of the structural business survey weights.

Fazio et al. (2006) point out that complex sampling design can distort the information contained in the finite population. They suggest that weights are more important for ensuring the unbiasedness of simple marginal statistics like means and tabulations. On the other hand, more complex statistics that depend on the correlations between variables may remain approximately unbiased even if unweighted. They also point out that in theory, applying weights to a model is safe as the impact of unnecessary weighting is to reduce the efficiency of the model, not to bias it. However, the fact that weighting does have the largest impact on variables where there is known to be significant measurement error does raise concerns.

Fazio et al. (2006) also highlight that most business surveys disproportionately sample too many large firms and as a result an unweighted regression will be driven by the data from these large firms, while a weighted estimate will be driven largely by data from small firms. In the Irish context most private sector firms are small firms, see Table A1 in the appendix.

An alternative strategy to applying weights to the raw data was outlined by Fazio et al. (2006) - that is the use of conditioning variables in the regression model. Using conditioning variables has the advantage that they may also have an economic interpretation. However, if you use weighting or conditioning variables, there is always the issue of over compensation.

Gelman (2007) states the issue a little more strongly when he declares that "survey weighting is a mess". He highlights the fact that survey weights are not equal to the inverse probability of selection and this point is salient when considering the use of survey weights when conducting regression analysis on the NES 2007 sample.

Angrist and Pischke (2009) also address the issue of weighting. They point out that weighting by the inverse probability will generate consistent population estimates, but they also warn that incorrect or poorly estimated weights can do more harm than good.

3. DATA

3.1 The National Employment Survey

The NES 2007 was a major workplace survey conducted by the CSO. The survey covered both the public and private sectors, the only excluded sectors being agriculture, forestry and fishing. Information was collected in a linked and integrated way from a sample of employers and employees. Only employers with more than three employees were surveyed. Employers were required to have been

trading in the reference month of October in 2007. Sampled employees were required to have been employed in the reference month of October in 2007.

The NES sample of employers was selected from the CSO Central Business Register. The sample was selected based on the proportion of companies in each economic sector (NACE Rev 1.1 two digit division) and in each size class. The employers were asked to select a systematic sample of employees from their payrolls. The table below outlines the number of employers and employees sampled for each size class of business unit:

Table 1: The number of employers and employees sampled for each size class of business unit

Size of Enterprise (No. of Employees)	No of employers sampled	No. Employees sampled
3-9	1 in 20	All
10-19	1 in 10	All
20 - 49	1 in 7	1 in 2
50 - 99	1 in 4	1 in 3
100 - 249	1 in 2	1 in 7
250 - 999	All	1 in 10
1000 +	All	1 in 20

The responding employers returned the employer questionnaire that contained a list of the names of sampled employees to the CSO who were then contacted and asked to return a questionnaire directly to the CSO.

Overall, the number of respondent employees was equivalent to 3.5% of all relevant employees. The respondent enterprises represented approximately 6% of all enterprises which employed approximately 56% of all employees. The data provided from employers and employees were then weighted to compensate for differing sampling fractions, non-response and to gross up to the overall population. Non-response rates were higher in the smaller size classes. The construction of the weighting system used for the NES 2007 involved a number of stages. Two separate weights were calculated – one to gross up to the number of employees in the enterprise, and another to gross up to the population of employers. These weights were then combined to yield an overall weight to make an initial estimate of the overall population of employees. This weight was further calibrated so that the NES population estimates matched the QNHS estimates. This calibration was based on NACE sector, education group, public-private sector, occupation, gender, full-time/part-time status and age group.

3.2 Univariate descriptive statistics for permanent full-time employees aged 25-59

In line with Kelly et al (2009a and b), and Murphy and Ernst & Young (2007), we restrict the sample of employees to a cohort consisting of permanent full-time employees aged 25 – 59.

As well as earnings data, the NES 2007 contained a wide range of data on the background characteristics of each individual employee. When we look in detail at these characteristics, we see that the profiles of the public sector and private sector employees differed in a number of ways. An analysis of educational qualifications in the public and private sectors showed that 39.3% of public sector employees had a third level degree or higher qualification compared with 23.5% in the private sector. In the public sector 16% of employees had a primary or lower secondary qualification while in the private sector, this figure was 23.2%.

There was also a noticeable difference in the structure of employment in the two sectors. In the private sector, 14.7% of employees were Managers compared with 4.3% in the public sector. Over 27% of public sector workers described themselves as Professional, compared with 8.4% in the private sector. In contrast, only 0.1% of public sector employees were categorised in Sales occupations whereas this figure was 9.2% in the private sector.

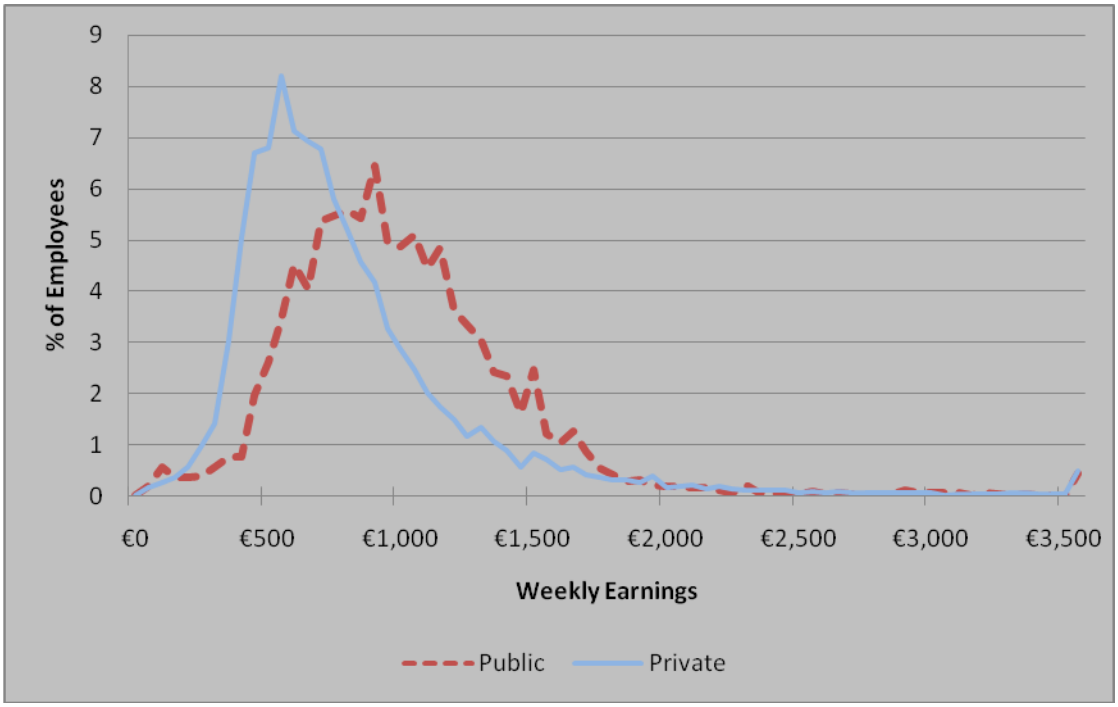
For the group of permanent full-time employees aged 25-59 the average public service weekly earnings were over 24% higher than the private sector. The corresponding premium for hourly earnings was almost 36%. The male premium for hourly earnings was 31.2% and for females it was 48.5%. The hourly earnings for males were €22.88 and €21.20 for females. The gender gap was smaller in the public service where the hourly earnings were on average 4.3% higher for males than for females, compared with 18% higher in the private sector.

Public service employees tended to be older that those in the private sector; the average age in the public sector was 40.7 years, compared with 37.7 in the private sector. Employees in the public sector had spent on average 13.3 years with their current employer. This figure was 9.4 years for the private sector. Similarly, public sector employees had more overall experience than those in the private sector, with 18.9 years in total paid employment compared with 16.3 for the private sector.

Employees in the private sector worked a longer week, on average, than those in the public sector. The average number of hours worked per week in the public sector was 36.1 compared with 38.7 in the private sector.

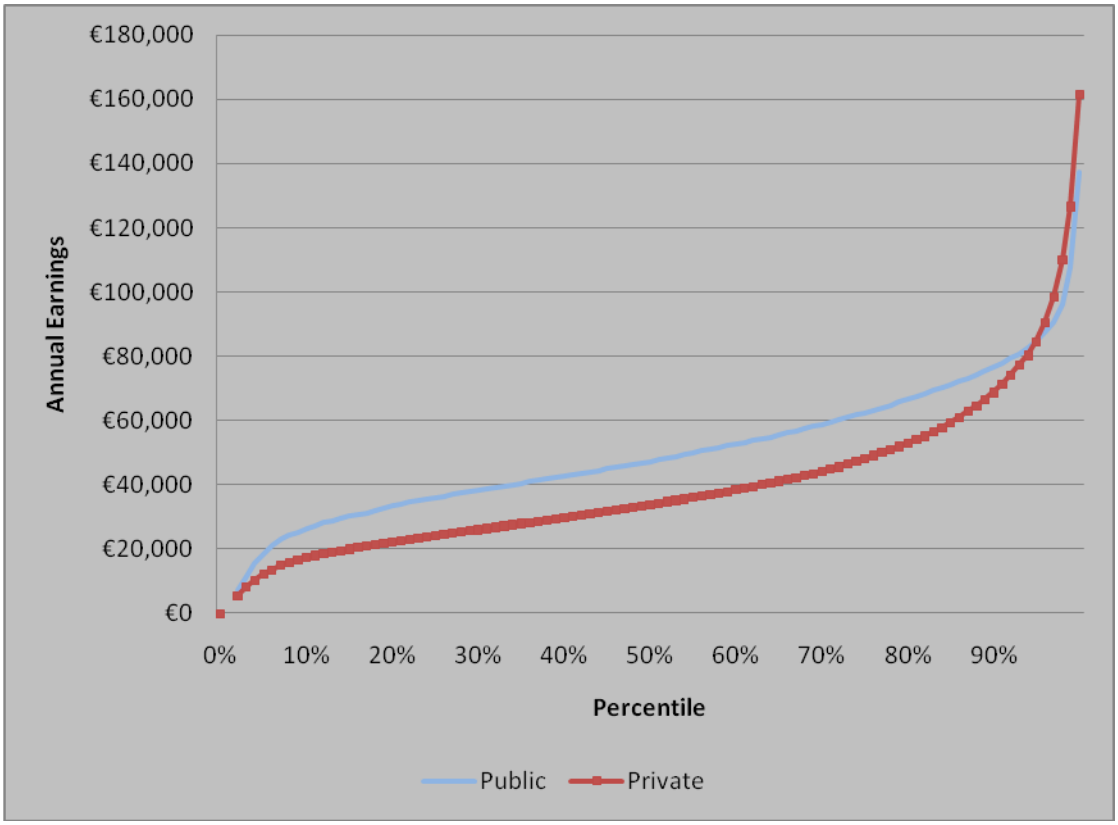
Figure 1 shows, the distribution of weekly earnings in both the public and private sectors for permanent full-time employees aged 25-59. It is clear from the graph that the earnings distribution for the private sector was more positively skewed than that for the public sector. There was a higher concentration of employees from the private sector at the lower end of the earnings distribution.

Figure 1. Distribution of Weekly Earnings for permanent full-time employees aged 25-59



The fact that the distribution of earnings is quite different across the two sectors means that it is important to look at the earnings differential across the entire distribution, and not just at the mean. The following graph presents the annual earnings percentiles for the public and private sectors. We can see that the differential between the two sectors starts to diminish in the higher percentiles, and actually turns into a negative value from the 95th percentile on.

Figure 2. Percentiles of Annual Earnings for permanent full-time employees aged 25-59



4. METHODOLOGY

Simple comparisons of weekly earnings, as presented in Section 3, do not take account of the differing composition of private sector and public sector employees with regard to education, gender, experience etc. It is important to control for all of these characteristics when drawing comparisons between public and private sector pay. This report presents a typical multivariate model, which controls for relevant factors such as age, gender, and education. This analysis did not attempt to match individuals across the two sectors, or to control for differences in job "types" etc. between the sectors as this information needed to conduct such an analysis is not available in the NES 2007 dataset.

4.1 Basic weekly earnings regression

An ordinary least square (OLS) regression was used to model the natural log of weekly earnings on a set of explanatory variables. The semi-log hedonic earnings equation may be represented as follows:

$$\ln w_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_p x_{pi} + \varepsilon_i$$

Where w_i is the weekly earnings of individual *i*, and $x_{1i}, x_{2i}, \ldots x_{pi}$ are a set of *p* explanatory variables, capturing individual and work-place characteristics. The intercept term is denoted by α, and the ε_i term is the random error term. This standard OLS model is widely used in the analysis of gender and public-private wage gaps in both the national and international literature. The approach adopted in this report is similar to that used in Belman and Heywood (2004) and used the following explanatory variables:

(i) occupation, (ii) educational attainment, (iii) full-time status, (iv) gender, (v) public or private sector, (vi) nationality, (vii) membership of a trade union, (viii) membership of a professional body, (ix) age, (x) age-squared[2], (xi) size of enterprise[3], (xii) permanent/non-permanent job status, (xiii) length of service with current employer, (xiv) total length in employment, (xv) log[4] of overtime hours (38+) worked and (xvi) log of hours worked, (xvii) shift work, (xviii) supervisory status. It should be noted that the included variables will not capture all the variation in earnings

The approach is sometimes referred to as a hybrid approach (Belman and Heywood (1996), Bender and Elliott (2007)) in that it accounts both for differences in the human capital of the employees in the two sectors, and for differences in the characteristics of the workplace. Models both including and excluding size of enterprise as an explanatory variable were considered in this analysis.

Boyle et al. (2004) include firm size as an explanatory variable in their analysis of European Community Household Panel (ECHP) 1994-2001 data. Murphy and Ernst & Young (2007) analyse models that include and exclude size separately. The CSO (2009) also report results that include and exclude size as an explanatory variable. Kelly et al. (2009a) and Kelly et al. (2009b) exclude size as an explanatory variable.

The *a priori* expectation is that including the size of enterprise as an explanatory variable will decrease the public sector premium; public sector organisations are generally large organisations, and there is evidence that, in general, workers in large organisations are paid more. For a more detailed discussion on some of the issues around the use of size as an explanatory variable see Boyle et al. (2004), Chatterji and Mumford (2007), Cai and Liu (2008) and Kelly et al. (2009b).

4.2 The Blinder-Oaxaca Decomposition

The public-private sector wage differential calculated using the OLS regression method, described above, is limited in the information it provides about the differential. While it takes account of individual characteristics, it assumes that the return on these characteristics is the same for both the public and private sectors.

In the Blinder-Oaxaca[5] method, two separate OLS equations are calculated for the public and private sectors. Using the estimated parameter from the two models, the differential can be decomposed into the part that can be explained by the differential attributes of individuals and the characteristics of their workplace in the public and private sectors with the remainder representing the unexplained part of the differential. The Blinder-Oaxaca decomposition of earnings is often represented in the literature as follows:

$$\ln \overline{w}^{pub} - \ln \overline{w}^{pri} = \Sigma_{\mathrm{p}}\left(\bar{x}_p^{pub} - \bar{x}_p^{pri}\right)\hat{\beta}_p^{pri} + \Sigma_{\mathrm{p}}\bar{x}_p^{pub}\left(\hat{\beta}_p^{pub} - \hat{\beta}_p^{pri}\right) + \left(\hat{\alpha}^{pub} - \hat{\alpha}^{pri}\right)$$

where, $\Sigma_{\mathrm{p}}\left(\bar{x}_p^{pub} - \bar{x}_p^{pri}\right)\hat{\beta}_p^{pri}$ represents the difference in the log of earnings that is explained by the explanatory variables and $\Sigma_{\mathrm{p}}\bar{x}_p^{pub}\left(\hat{\beta}_p^{pub} - \hat{\beta}_p^{pri}\right) + \left(\hat{\alpha}^{pub} - \hat{\alpha}^{pri}\right)$ estimates the unexplained earnings gap, i.e. the public–private wage differential.

The Blinder-Oaxaca decomposition is currently considered the preferred method of calculating the public–private wage differential in the literature. In keeping with Murphy and Ernst & Young (2007) and Kelly et al (2009a,b), the reference category[6] used for the Blinder-Oaxaca decompositions was the private sector.

[2] We used age-squared as an explanatory variable to capture the non-linear relationship between earnings and age. Similar results are obtained using experience and experience squared.

[3] Size of enterprise is categorised into a banded variable in keeping with the stratification adopted in the sampling design of the NES 2007.

[4] In line with Murphy and Ernst & Young (2007).

[5] Blinder (1973), Oaxaca (1973).

[6] The Blinder-Oaxaca decomposition is not unique and the choice of reference group affects the results. Results were also calculated using the public sector as the reference group but these results negates the effect that size of enterprise has as an explanatory variable.

4.3 Quantile Regression

OLS regression is limited in the information that it can provide about earnings as it only estimates *average* (the conditional mean of) earnings corresponding to the various explanatory variables. Quantile regression is used when an estimate of the effect of changes in the various explanatory variables on the quantiles or percentiles of a population is required, rather than simply estimating the effect at the conditional mean. It is widely used in the literature on the public-private sector wage gap as it allows us to examine how the public sector differential varies across the earnings distribution.

In this report we add value to the recent literature by estimating the public sector premium at each percentile along the distribution of earnings rather than just at the deciles. The advantage of this more detailed analysis is that we can identify more clearly any turning points in the growth/decline in the pay differential along the distribution, as well as being able to identify any point along the distribution where the public sector premium may become a discount.

Using unweighted data we estimate the quantile regressions for all permanent full-time employees aged 25-59, as well as for male and female permanent full-time employees aged 25-59 separately. We also estimate the quantile regression both including and excluding size of enterprise as an explanatory variable in order to estimate the impact of the exclusion of this variable on the estimated pay differential over the full earnings distribution.

4.4 Propensity Score Matching

Angrist and Pischke (2009) point to a number of concerns around the use of PSM. The primary concern as outlined by Angrist and Pischke is that: "…there are many details to be filled in when implementing propensity score matching … (and) these details are not yet standardized." They develop this concern to say that researchers using the same data and the same explanatory variables may end up reaching different conclusions. This point may be evident In the Irish literature on public-private pay gap measurement. Murphy and Ernst & Young (2007) found that the PSM estimates of the public sector wage gap matched the estimates obtained by their Blinder-Oaxaca decomposition whilst their OLS estimate was lower. In contrast, Kelly et al. (2009b) found very little difference between their PSM and OLS estimates.

Due to the lack of systematic guidelines on the selection of a comparison group, this study does not use propensity score matching to estimate the public-private pay gap. We have found that the estimated premium using propensity score matching is highly sensitive to the criteria used to discard sub-samples without a common support.

5. RESULTS

5.1 OLS and Blinder-Oaxaca Results[7]

The first set of results presented here give estimates of the average public sector wage differential for permanent full-time employees aged 25-59. In order to highlight the issues surrounding the use of weights in multivariate analyses, we provide estimates based on unweighted and weighted data. Two sets of weighted results are presented; (i) Weight 1 is the weight used to gross the NES sample up to the overall population of employees and accounts for non-response as well as the probability of being sampled, (ii) Weight 2 is the calibrated weight used for univariate analysis in the NES. See Table 2 below.

[7] All of the results presented in this report relate to the percentage premium based on weekly earnings. The premia were calculated as the antilog of the estimated coefficient minus 1. Only the estimated premia are presented in this paper. Detailed regression results for the unweighted data are available in the Appendix. Full details of all other regression results are available from the authors on request.

Table 2 - Estimated Average Public Sector Premium for Permanent Full-Time Employees Aged 25-59 (weekly earnings) – including size of enterprise as an explanatory variable

	All		Males		Females	
	OLS	Blinder-Oaxaca	OLS	Blinder-Oaxaca	OLS	Blinder-Oaxaca
Unweighted	10.0%	11.5%	7.2%	9.0%	12.7%	13.9%
Weight 1	12.6%	13.0%	9.8%	11.2%	13.2%	14.5%
Weight 2	14.5%	15.9%	13.7%	16.6%	13.9%	14.9%

The results of the Blinder-Oaxaca decomposition yielded a public sector premium for permanent full-time employees aged 25-59 ranging from 11.5% based on unweighted analysis to 15.9% based on an analysis using Weight 2. The OLS regression analyses produced similar results.

When the differential was calculated separately for males and for females we found that when unweighted analysis was used, the premium for females was higher than that for males (a 7.2% premium for males compared with a 12.7% premium for females). This is consistent with what is found in the international literature. Similar trends were found when the analysis was weighted using Weight 1. Interestingly however, we found that when the analysis was weighted using Weight 2, the difference in the premium between males and females almost disappears using OLS, and actually results in a higher premium for males than females using the Blinder-Oaxaca decomposition. It should be noted that non-response is higher for males in this survey than for females and that the male response rates are particularly poor in some sectors of the private sector, e.g. construction.

All of the analyses were repeated using a model specification that excluded size of enterprise as an explanatory variable. By excluding size of enterprise from the model specification, the average estimated premium increased by approximately 5 percentage points. See Table 3 below.

Table 3 - Estimated Average Public Sector Premium for Permanent Full-Time Employees Aged 25-59 (weekly earnings) – excluding size of enterprise as an explanatory variable

	All		Males		Females	
	OLS	Blinder-Oaxaca	OLS	Blinder-Oaxaca	OLS	Blinder-Oaxaca
Unweighted	15.1%	16.0%	12.6%	13.7%	18.0%	19.2%
Weight 1	17.4%	17.7%	15.9%	16.5%	19.3%	20.0%
Weight 2	20.1%	20.5%	19.2%	21.3%	21.4%	21.2%

In keeping with Murphy and Ernst & Young (2007), we estimated the public private pay differential using OLS on the core sample excluding personal and protective services. This occupation group includes Gardaí, Prison Officers and some elements of the Defence Forces as well as private sector employees engaged in personal and protective services. Intuitively, this sector is one of the occupation groups where comparison between public and private employees is of questionable value (See appendix for a more detailed list of the occupations included in this classification group). The disparity in the earnings distributions between the public and private sectors was more pronounced in this group than in any of the other occupation groups.

It was found that when this particular sector was excluded from the analysis, the size of the public sector premium decreased, especially for males. The public sector pay differential for females relative to males is, in this instance, in line with the findings in the international literature, even for the weighted data. See Tables 4 and 5 below.

Table 4 - Estimated Average Public Sector Premium for Permanent Full-Time Employees Aged 25-59 (weekly earnings) – excluding Protective Services (including size)

	All	Males	Females
Unweighted	7.1%	2.7%	11.5%
Weight 1	7.3%	3.2%	11.2%
Weight 2	8.4%	5.0%	10.1%

Table 5 - Estimated Average Public Sector Premium for Permanent Full-Time Employees Aged 25-59 (weekly earnings) – Excluding Protective Services (excluding size)

	All	Males	Females
Unweighted	11.9%	7.7%	16.8%
Weight 1	12.8%	9.0%	17.5%
Weight 2	13.9%	10.2%	17.8%

Again the difference between the estimates of the average premia for the regressions including and excluding size is consistently about 5 percentage points.

5.2 Quantile Regression Results

Figure 3 shows the estimated premia at each percentile in the earnings distribution based on quantile regression analysis for permanent full-time employees aged 25-59. This analysis is again based on a regression model that included the size of enterprise as a banded variable and the data is unweighted. The OLS estimate of the average premium is also included as reference and to highlight the limitation of this estimate. The estimated premium peaks at the 3rd percentile at just over 20% before then declining monotonically along the earnings distribution. The estimated pay-gap becomes a discount at the 89th percentile.[8] At the 99th percentile the discount is 16.6% and is statistically significant at the 1% level.[9]

Figure 4 shows the estimated premia at each percentile in the earnings distribution for *male* employees. The estimated premium peaks at the 3rd percentile at just over 20% before then consistently decreasing along the earnings distribution. The estimated pay-gap becomes a discount at the 80th percentile. At the 99th percentile the discount is 25.5%.

Similarly, Figure 5 shows the estimated premia at each percentile in the earnings distribution for *female* employees. It is clear that the overall trend in the differential across the earnings distribution is flatter than that for males. The estimated differential peaks at the 4th percentile at just over 21%, it then reduces to become a discount at the 97th percentile.

Quantile regression analysis was also performed on a model specification excluding size of enterprise. Figure 6 compares these with analyses that included size of enterprise. Overall the effect of excluding the size of enterprise from the model is to increase the size of the premium by just over 5 percentage points. The difference between the two models is at its highest at the 1st percentile (7.6 percentage points) and is at its lowest at the 95th percentile (2.5 percentage points). The results show in general that the trend in the pay-gap across the earnings distribution is similar in both instances, but the difference is higher at the lower end of the distribution and the impact of the variable's inclusion diminishes along the earnings distribution.

[8] See Table A5 in the Appendix for the earnings distribution in terms of annual earnings.
[9] In general, all of the estimates are statistically significant at the 5% level. The only exceptions to this are at those percentiles where the estimated premia are close to zero.

Figure 3. Public sector premia (%) across weekly earnings distribution for All Employees – including size as an explanatory variable

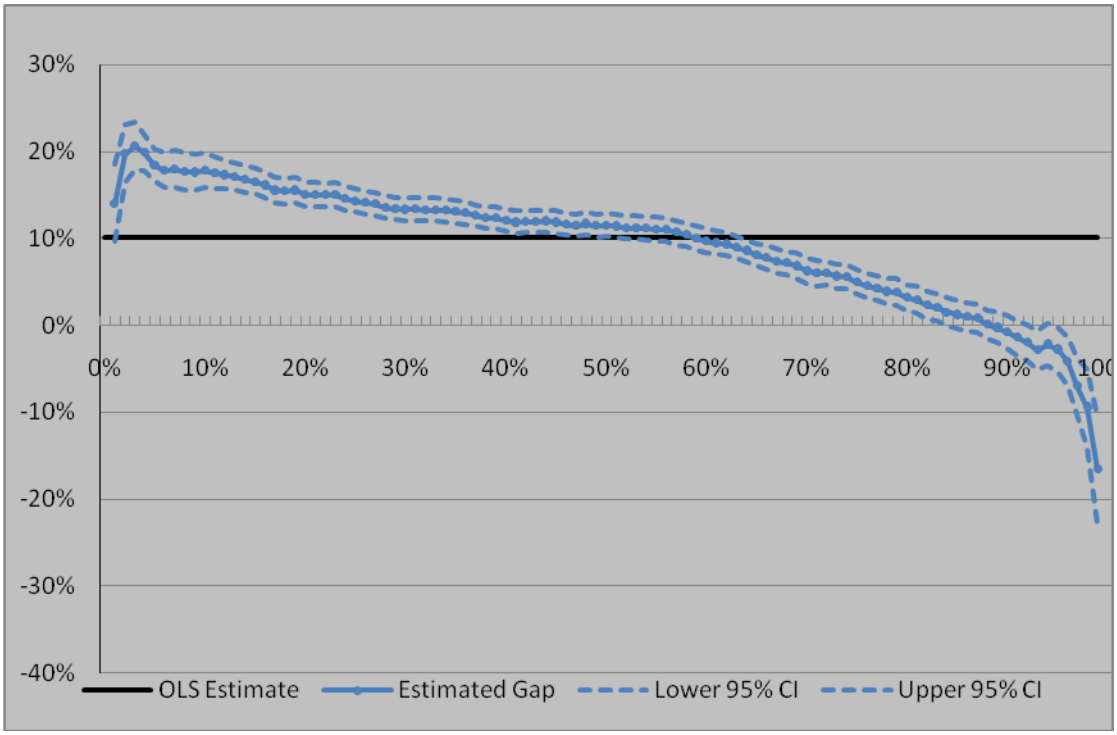


Figure 4. Public sector premia (%) across weekly earnings distribution for Male Employees – including size as an explanatory variable

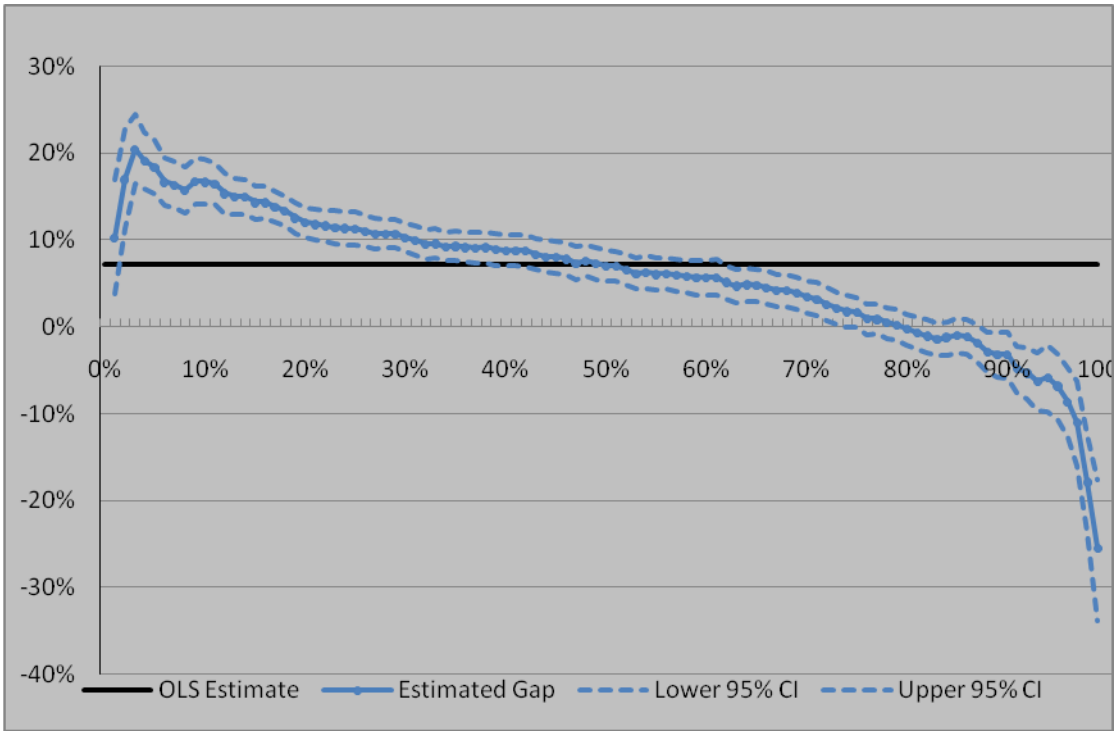


Figure 5. Public sector premia (%) across weekly earnings distribution for Female Employees – including size as an explanatory variable

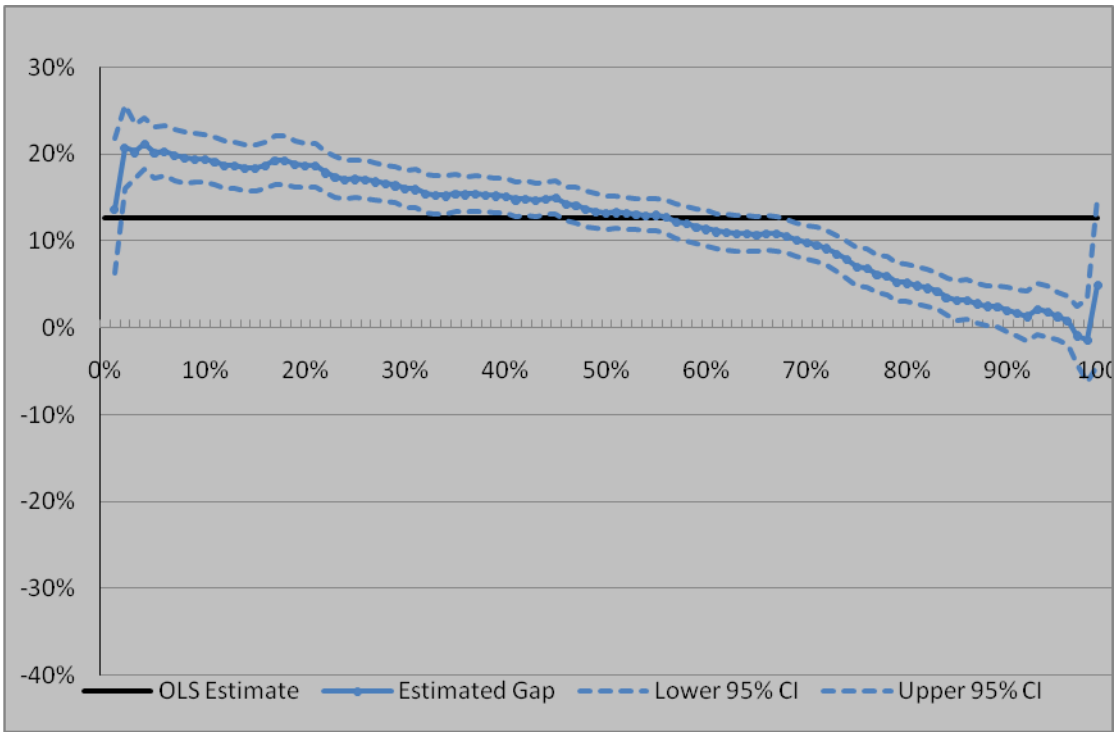
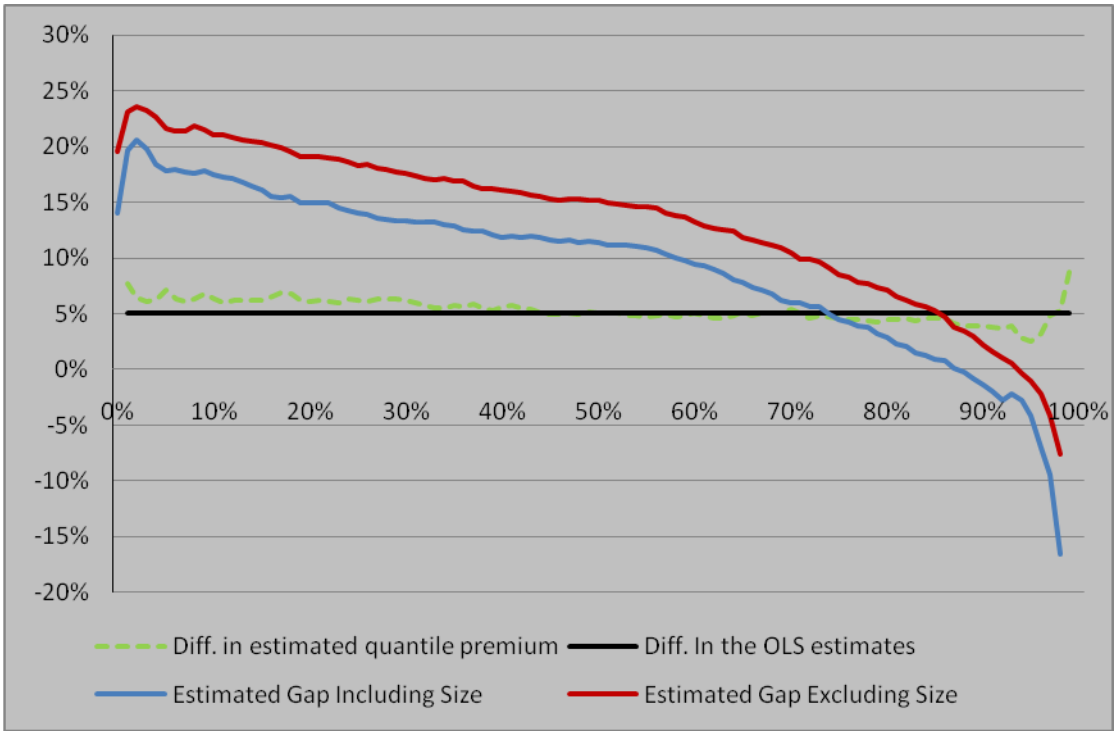


Figure 6. Public sector premia (%) across weekly earnings distribution for All Employees – including and excluding size as an explanatory variable



6. SUMMARY AND CONCLUSIONS

In the international literature, there is no general consensus on how to calculate a definitive measure of the public-private sector wage gap. There are a number of accepted approaches that one can take from a methodological point of view (OLS, Blinder-Oaxaca decomposition, quantile regression, propensity score matching, etc.). However there are also competing arguments with regard to the specification of the models used with these approaches. To further compound these difficulties, there is also an issue regarding the use of survey weights in these types of multivariate analyses of structural business microdata, with an inherent bias towards larger firms with potential non-response bias.

Taking all of this into account, we advise caution in attempting to estimate one definitive "answer" for the average premium. Using the Blinder-Oaxaca decomposition method, we found estimates of the public-private pay gap in October 2007 ranging from 11.5% using unweighted data, to 15.9% using weighted data for all employees. We found that excluding size of enterprise from the model had the effect of increasing the magnitude of the average premium by approximately 5 percentage points.

It is widely accepted that the hypothesis of a constant wage premium, as implied by any of the analyses that focus on average differentials, can be rejected. It is therefore important to look at the pay gap across the entire distribution of earnings. Using quantile regression at each percentile, we found that the premium was highest at the lower end of the earnings distribution, and generally decreased to become a discount at the top end.

Furthermore, in the Irish context, there are a number of occupations within the public sector that really have no comparable occupation in the private sector, and vice versa (most noticeably Gardaí, Prison Officers, and members of the Defence Forces). To highlight the consequences of ignoring this lack of comparability in some occupation sectors, this study also estimated the public-private wage gap on a sub-sample excluding personal and protective services employees. The impact of excluding this sector was to dramatically reduce the average public-private wage gap, especially for males (from 7.2% to 2.7% using OLS). The size of this reduction was even larger when the analysis was conducted using weighted data.

The purpose of this paper is to raise awareness of all of the issues mentioned above and we present them in an open-ended and transparent fashion in the hope of generating constructive debate. There is obviously scope for future work in this area. In particular, we think that much greater debate and awareness needs to be given to the issue of handling structural business survey data. The specification of models based on biased samples needs to be addressed, either by using conditioning variables, or the use of appropriate survey weights. To this end, it is essential that researchers are fully aware of how survey weights that are included in micro datasets are constructed and on the appropriate use of these weights in multivariate analysis. Finally, if we accept the importance of quantile regression as a method of estimating the public-private pay gap, then the logical next step is to estimate quantile regressions using the Blinder-Oaxaca decompositions and similar techniques.

BIBLIOGRAPHY

Angrist, J. and Pischke, J. S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, New Jersey.

Belman, D. and Heywood, J. S. (1996). "The Structure of Compensation in the Public Sector." *Public Sector Employment*, (ed.) Gunderson, Hyatt, Belman. Madison, WI: Industrial Relations Research Association.

Belman, D. and Heywood, J. S. (2004). "Public-Sector Wage Comparability: The Role of Earnings Dispersion." *Public Finance Review*, Vol. 32 No. 6, 567-587.

Bender, K. A. and Elliott, R. F. (2002). "The Role of Job Attributes in Understanding the Public-Private Wage Differential." *Industrial Relations*, Vol. 41, No. 3. 407-421

Blinder, A. (1973). "Wage Discrimination: Reduced Form and Structural Estimates", *Journal of Human Resources*, 8, 436-55.

Boyle, G., McElligot, R., & O'Leary, J. (2004). "Public-Private Wage Differentials in Ireland, 1994-2001." *ESRI Quarterly Economic Commentary*, Special Article. Summer 2004. Dublin: The Economic and Social Research Institute.

Cai, L. & Liu, A. (2008). "Public-Private Wage Gap in Australia: Variation Along the Distribution", *Centre for Economic Policy Research Discussion Paper* No. 581 (June) Canberra, Australia: The Australian National University.

Central Statistics Office (2009), National Employment Survey 2007. http://www.cso.ie

Central Statistics Office (2009), National Employment Survey 2007 Supplementary Analysis. http://www.cso.ie

Chatterji, M. and Mumford, K. (2007). "The Public-Private Sector Wage Differential for Full-Time Male Employees in Britain: A Preliminary Analysis." *IZA Discussion Paper* no. 2871 (May). Bonn, Germany: The Institute for the Study of Labour.

Deaton, A. (1997). *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. World Bank / John Hopkins University Press. Baltimore, MD.

Fazio, G., Lam, K. H. and Ritchie, F. (2006) "Sample Bias in Microeconometric Analyses of Official Microdata." Office for National Statistics, UK, URN 06/737.
http://www.berr.gov.uk/files/file21911.pdf

Gelman, A. (2007) "Struggles with Survey Weighting and Regression Modeling." *Statistical Science*, Vol. 22, No. 2, 153-164.

Gregory, R. & Borland, B. (1999). "Recent Developments in Public Sector Labor Markets", Asenfelter, O. and Card, D. (eds.), *Handbook of Labour Economics*, Volume 3, Elsevier Science: Amsterdam.

Kelly, E., McGuinness, S., & O'Connell, P. (2009a). "Benchmarking, Social Partnership and Higher Remuneration: Wage Setting Institutions and the Public-Private Sector Wage Gap in Ireland". *Economic and Social Review*, Vol. 40, No. 3, Autumn, 2009, pp. 339-370.

Kelly, E., McGuinness, S., & O'Connell, P. (2009b). "The Public-Private Sector Pay Gap in Ireland: What Lies Beneath?". ESRI Working Paper No. 321 (October). Dublin, Ireland: The Economic and Social Research Institute.

Lucifora, C. & Meurs, D. (2006). "The Public Sector Pay Gap in France, Great Britain and Italy", *Review of Income and Wealth*, Vol. 52, No. 1, pp. 43-59.

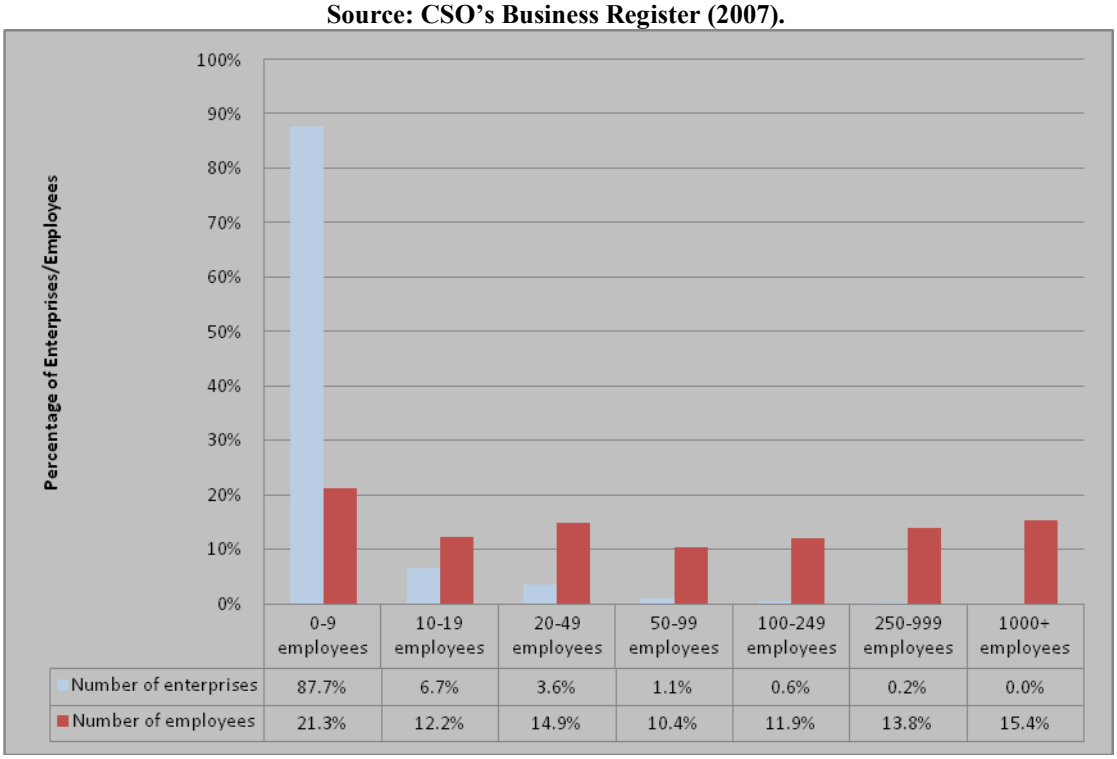
Murphy, A. and Ernst & Young (2007). "An Econometric Study of Earnings Based on National Employment Survey 2003 Data". (Report to Public Service Benchmarking Body)
http://www.benchmarking.gov.ie/Documents/Econometric%20Study%20of%20Earnings%202007.pdf

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

Winship, C. and Radbill, L. (1994). "Sampling Weights and Regression Analysis". *Sociological Methods and Research* 23(2):230-257.

APPENDIX

Figure A1 Distribution of Private Sector Enterprises and Employees by size of Enterprise. Source: CSO's Business Register (2007).



	0-9 employees	10-19 employees	20-49 employees	50-99 employees	100-249 employees	250-999 employees	1000+ employees
Number of enterprises	87.7%	6.7%	3.6%	1.1%	0.6%	0.2%	0.0%
Number of employees	21.3%	12.2%	14.9%	10.4%	11.9%	13.8%	15.4%

Variable Definitions

Earnings

This is defined as gross earnings (before the deduction of tax, PRSI, superannuation) payable by organisations to its employees. It includes normal wages, salaries and overtime, taxable allowances, regular bonuses and commissions, holiday and sick pay. It does not include irregular bonuses and commissions, employer's PRSI, redundancy payments and back pay.

Hours

This is defined as total paid contracted hours plus paid overtime hours. It includes paid leave and excludes unpaid leave and unpaid overtime.

Nationality

Irish: Republic of Ireland.

EU15 excluding Ireland: Great Britain and Northern Ireland, Austria, Belgium, Denmark, Finland, France, Germany, Greece, Netherlands, Italy, Luxembourg, Portugal, Spain, and Sweden.

Accession States EU15 to EU27: Bulgaria, Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Romania, Slovakia, and Slovenia.

Other nationalities: All other nationalities not included in the above three groupings as well as those who could not be coded (the uncoded employees represented approximately 1.2% of all employees).

Public Sector

The Public Sector includes:

- Civil Service; Defence Forces; Garda Síochána; Local Authorities; Education (excluding private institutions); Regional Bodies; Health (excluding private institutions); Semi-State Bodies (excluding their subsidiary companies)

Occupations classified as personal and protective services workers

511 Travel attendants and related workers
- 5111 Travel attendants and travel stewards
- 5112 Transport conductors
- 5113 Travel guides

512 Housekeeping and restaurant services workers
- 5121 Housekeepers and related workers
- 5122 Cooks
- 5123 Waiters, waitresses and bartenders

513 Personal care and related workers
- 5131 Child-care workers
- 5132 Institution-based personal care workers
- 5133 Home-based personal care workers
- 5139 Personal care and related workers not elsewhere classified

514 Other personal services workers
- 5141 Hairdressers, barbers, beauticians and related workers
- 5142 Companions and valets
- 5143 Undertakers and embalmers
- 5149 Other personal services workers not elsewhere classified

516 Protective services workers
- 5161 Fire-fighters
- 5162 Police officers
- 5163 Prison guards
- 5169 Protective services workers not elsewhere classified

Figure A2 Distribution of Weekly Earnings for Personal and Protective Services Employees by sector

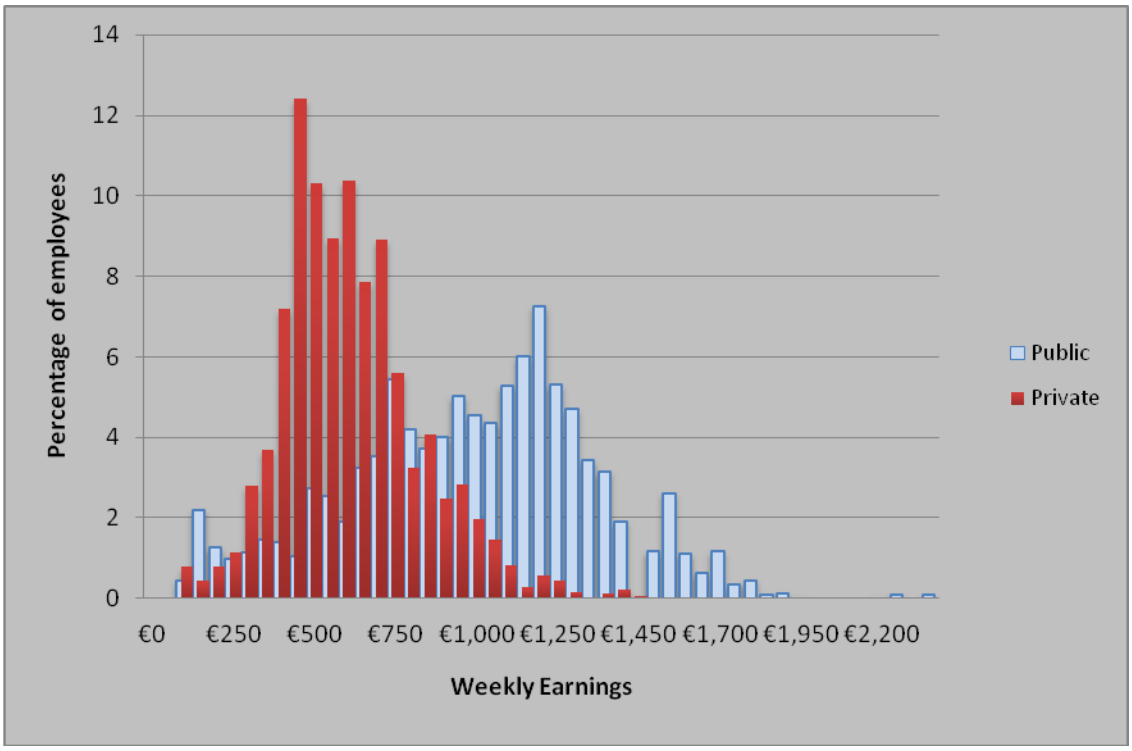


Table A1: OLS model estimates on log weekly earnings: Including size of enterprise as an explanatory variable

Permanent Full-Time Employees Aged 25-59 (Unweighted Results)						
	Males & Females		**Males**		**Females**	
Parameter	**Estimate**	**t Value**	**Estimate**	**t Value**	**Estimate**	**t Value**
Intercept	1.803	36.8	1.639	23.6	2.158	31.4
Occupation						
Manager and administrators	0.494	52.0	0.488	40.9	0.505	31.4
Professional	0.406	47.0	0.381	33.0	0.446	33.2
Associate professional and technical	0.229	24.9	0.200	16.9	0.282	19.1
Clerical and secretarial	0.098	11.6	0.017	1.4	0.157	12.4
Craft and related trades	0.186	19.5	0.176	16.2	0.094	3.2
Personal and protective services	0.072	6.9	0.069	5.0	0.092	5.6
Sales	0.066	6.2	0.096	6.8	0.052	3.3
Plant and machine operatives	0.046	5.4	0.031	3.0	0.051	3.0
Education attained						
Third level degree or above	0.322	43.8	0.306	31.2	0.328	28.5
Third level non degree	0.180	24.7	0.175	18.0	0.180	15.8
Post leaving certificate	0.109	14.3	0.119	12.8	0.067	5.1
Higher secondary	0.091	14.0	0.090	11.0	0.089	8.2
Male	0.160	36.7				
Public sector	0.096	16.1	0.069	8.3	0.119	14.3
Nationality						
EU15 excluding Ireland	-0.081	-8.5	-0.088	-6.9	-0.070	-5.0
EU Accession states	-0.214	-21.9	-0.211	-16.1	-0.202	-14.0
Other nationality	-0.126	-13.3	-0.163	-12.1	-0.080	-6.2
Trade Union Member	0.026	5.4	0.021	3.2	0.021	3.0
Member of a Professional Body	0.075	14.0	0.083	11.2	0.069	9.2
Age	0.041	21.6	0.050	19.1	0.031	11.1
Age^2	-0.487	-21.5	-0.563	-18.4	-0.390	-11.6
Size of Enterprise						
1000+ employees	0.235	28.2	0.262	22.9	0.210	17.6
250 - 999 employees	0.242	30.6	0.265	24.9	0.218	18.7
100 - 249 employees	0.201	22.9	0.226	19.3	0.176	13.4
50 - 99 employees	0.213	22.5	0.228	18.4	0.194	13.3
20 - 49 employees	0.120	14.2	0.135	12.1	0.099	7.7
10 - 19 employees	0.086	10.0	0.088	7.7	0.079	6.2
Length of service with current employer	0.005	19.1	0.004	10.9	0.009	18.5
Total time in all paid employment	0.005	16.6	0.005	10.4	0.005	9.9
Ln Overtime Hours	-0.018	-20.1	-0.020	-16.5	-0.016	-13.0
Ln Hours	0.893	101.7	0.933	71.2	0.848	74.1
Shift Work	-0.028	-5.5	-0.011	-1.6	-0.054	-6.9
Supervisor	0.082	19.0	0.092	15.7	0.077	12.1
n	35,047		20,352		14,695	
R-Square	0.563		0.517		0.602	

Table A2: OLS model estimates on log weekly earnings: Excluding size of enterprise as an explanatory variable

Permanent Full-Time Employees Aged 25-59 (Unweighted Results)

Parameter	Males & Females		Males		Females	
	Estimate	t Value	Estimate	t Value	Estimate	t Value
Intercept	1.863	37.6	1.653	23.4	2.267	32.6
Occupation						
Manager and administrators	0.488	50.5	0.481	39.6	0.500	30.7
Professional	0.428	48.8	0.414	35.4	0.454	33.4
Associate professional and technical	0.252	27.0	0.232	19.4	0.292	19.5
Clerical and secretarial	0.106	12.4	0.050	3.9	0.149	11.7
Craft and related trades	0.164	17.0	0.156	14.1	0.079	2.6
Personal and protective services	0.084	7.9	0.084	6.0	0.097	5.8
Sales	0.065	6.1	0.102	7.0	0.041	2.6
Plant and machine operatives	0.066	7.5	0.050	4.9	0.082	4.8
Education attained						
Third level degree or above	0.346	46.3	0.340	34.2	0.336	28.7
Third level non degree	0.195	26.3	0.196	19.7	0.184	15.9
Post leaving certificate	0.113	14.6	0.129	13.5	0.062	4.6
Higher secondary	0.097	14.7	0.099	11.8	0.088	8.0
Male	0.162	36.6				
Public sector	0.141	25.8	0.119	15.4	0.165	21.5
Nationality						
EU15 excluding Ireland	-0.065	-6.7	-0.072	-5.5	-0.054	-3.8
EU Accession states	-0.207	-20.8	-0.208	-15.6	-0.190	-12.9
Other nationality	-0.123	-12.8	-0.156	-11.4	-0.077	-5.8
Trade Union Member	0.065	13.7	0.065	10.1	0.055	8.0
Member of a Professional Body	0.067	12.3	0.074	9.8	0.061	8.0
Age	0.042	21.7	0.051	19.5	0.031	10.9
Age^2	-0.503	-21.8	-0.590	-18.9	-0.397	-11.6
Length of service with current employer	0.005	18.2	0.004	10.0	0.008	18.1
Total time in all paid employment	0.006	16.9	0.005	11.0	0.005	9.7
Ln Overtime Hours	-0.018	-20.0	-0.020	-16.2	-0.017	-13.0
Ln Hours	0.908	101.9	0.958	71.8	0.857	73.9
Shift Work	-0.008	-1.6	0.016	2.5	-0.044	-5.6
Supervisor	0.082	18.5	0.088	14.7	0.080	12.4
n	35,047		20,352		14,695	
R-Square	0.547		0.497		0.589	

Table A3: OLS model estimates on log weekly earnings: Including size of enterprise as an explanatory variable

Permanent Full-Time Employees Aged 25-59 (Unweighted Results)

Excluding Personal and Protective Services

	Males & Females		Males		Females	
Parameter	**Estimate**	**t Value**	**Estimate**	**t Value**	**Estimate**	**t Value**
Intercept	1.902	37.2	1.718	23.6	2.277	31.9
Occupation						
Manager and administrators	0.485	50.6	0.479	39.7	0.497	30.8
Professional	0.402	46.1	0.378	32.4	0.440	32.6
Associate professional and technical	0.231	25.0	0.207	17.4	0.282	19.1
Clerical and secretarial	0.095	11.2	0.020	1.5	0.155	12.3
Craft and related trades	0.184	19.2	0.172	15.8	0.095	3.2
Personal and protective services	**	**	**	**	**	**
Sales	0.059	5.6	0.089	6.2	0.049	3.1
Plant and machine operatives	0.047	5.5	0.028	2.8	0.056	3.3
Education attained						
Third level degree or above	0.326	42.6	0.307	30.1	0.333	27.7
Third level non degree	0.174	22.9	0.165	16.2	0.176	14.8
Post leaving certificate	0.110	13.9	0.116	12.0	0.071	5.1
Higher secondary	0.090	13.2	0.088	10.3	0.087	7.7
Male	0.157	34.9				
Public sector	0.069	11.1	0.027	3.0	0.109	12.6
Nationality						
EU15 excluding Ireland	-0.083	-8.5	-0.089	-6.8	-0.073	-5.0
EU Accession states	-0.204	-20.0	-0.201	-14.9	-0.189	-12.4
Other nationality	-0.111	-11.2	-0.142	-10.0	-0.071	-5.3
Trade Union Member	0.038	7.6	0.040	5.8	0.024	3.3
Member of a Professional Body	0.077	14.2	0.083	10.9	0.075	9.9
Age	0.042	21.6	0.051	19.1	0.032	11.1
Age^2	-0.493	-21.1	-0.566	-17.9	-0.399	-11.5
Size of Enterprise						
1000+ employees	0.234	27.6	0.256	21.9	0.215	17.7
250 - 999 employees	0.245	30.5	0.267	24.7	0.221	18.7
100 - 249 employees	0.206	23.0	0.230	19.1	0.184	13.8
50 - 99 employees	0.219	22.6	0.231	18.3	0.203	13.5
20 - 49 employees	0.120	13.9	0.133	11.8	0.100	7.6
10 - 19 employees	0.084	9.7	0.083	7.1	0.081	6.2
Length of service with current employer	0.005	16.9	0.003	8.9	0.008	17.4
Total time in all paid employment	0.006	16.3	0.004	9.5	0.005	10.3
Ln Overtime Hours	-0.018	-19.7	-0.020	-16.4	-0.016	-12.4
Ln Hours	0.857	91.8	0.902	64.8	0.807	66.4
Shift Work	-0.042	-7.9	-0.021	-3.1	-0.073	-9.0
Supervisor	0.091	20.3	0.104	17.1	0.084	12.9
n	33,136		19,204		13,932	
R-Square	0.553		0.511		0.593	

Table A4: OLS model estimates on log weekly earnings: Excluding size of enterprise as an explanatory variable

Permanent Full-Time Employees Aged 25-59 (Unweighted Results)

Excluding Personal and Protective Services

	Males & Females		Males		Females	
Parameter	**Estimate**	**t Value**	**Estimate**	**t Value**	**Estimate**	**t Value**
Intercept	1.958	37.7	1.724	23.3	2.388	33.1
Occupation						
Manager and administrators	0.478	49.0	0.471	38.3	0.492	30.1
Professional	0.423	47.7	0.410	34.6	0.449	32.8
Associate professional and technical	0.254	27.1	0.238	19.7	0.291	19.4
Clerical and secretarial	0.104	12.0	0.052	4.0	0.148	11.5
Craft and related trades	0.162	16.7	0.152	13.6	0.080	2.6
Personal and protective services	**	**	**	**	**	**
Sales	0.058	5.4	0.094	6.5	0.037	2.3
Plant and machine operatives	0.067	7.7	0.048	4.6	0.088	5.1
Education attained						
Third level degree or above	0.352	45.3	0.344	33.3	0.342	28.0
Third level non degree	0.191	24.7	0.188	18.2	0.181	15.0
Post leaving certificate	0.114	14.2	0.127	12.8	0.064	4.5
Higher secondary	0.097	14.0	0.099	11.3	0.086	7.5
Male	0.160	34.9				
Public sector	0.113	19.5	0.074	9.0	0.155	19.4
Nationality						
EU15 excluding Ireland	-0.066	-6.6	-0.072	-5.4	-0.055	-3.8
EU Accession states	-0.197	-19.0	-0.199	-14.4	-0.176	-11.4
Other nationality	-0.109	-10.8	-0.135	-9.4	-0.069	-5.0
Trade Union Member	0.079	15.8	0.085	12.5	0.059	8.3
Member of a Professional Body	0.068	12.3	0.073	9.5	0.067	8.7
Age	0.043	21.6	0.053	19.4	0.032	10.9
Age^2	-0.510	-21.5	-0.595	-18.4	-0.406	-11.5
Length of service with current employer	0.005	15.9	0.003	7.9	0.008	17.1
Total time in all paid employment	0.006	16.8	0.005	10.4	0.005	10.1
Ln Overtime Hours	-0.018	-19.6	-0.021	-16.2	-0.016	-12.4
Ln Hours	0.874	92.1	0.927	65.5	0.817	66.3
Shift Work	-0.021	-3.9	0.007	1.1	-0.063	-7.6
Supervisor	0.090	19.8	0.099	16.1	0.087	13.2
n	33,136		19,204		13,932	
R-Square	0.536		0.490		0.579	

Table A5: Distribution of Annual Earnings* of Permanent Full-Time Employees Aged 25-59

Percentile	Males & Females	Males	Females	Percentile	Males & Females	Males	Females
1%	€8,942	€11,379	€6,968	**51%**	€37,661	€40,112	€33,930
2%	€12,480	€15,327	€10,154	**52%**	€38,074	€40,612	€34,242
3%	€14,963	€17,403	€12,493	**53%**	€38,635	€41,120	€34,759
4%	€16,289	€18,418	€14,183	**54%**	€39,092	€41,600	€35,011
5%	€17,511	€19,500	€15,067	**55%**	€39,608	€42,008	€35,503
6%	€18,200	€20,280	€16,120	**56%**	€40,004	€42,512	€36,007
7%	€18,967	€20,917	€16,848	**57%**	€40,521	€43,016	€36,400
8%	€19,624	€21,580	€17,537	**58%**	€41,015	€43,524	€36,855
9%	€20,176	€22,090	€18,075	**59%**	€41,590	€44,044	€37,387
10%	€20,787	€22,620	€18,642	**60%**	€42,008	€44,538	€37,891
11%	€21,196	€23,166	€19,084	**61%**	€42,601	€45,009	€38,480
12%	€21,710	€23,660	€19,539	**62%**	€43,173	€45,427	€38,987
13%	€22,152	€24,185	€19,877	**63%**	€43,719	€45,929	€39,512
14%	€22,620	€24,814	€20,280	**64%**	€44,304	€46,521	€40,004
15%	€23,075	€25,345	€20,670	**65%**	€44,814	€47,190	€40,534
16%	€23,465	€25,857	€21,008	**66%**	€45,331	€47,829	€41,106
17%	€23,993	€26,208	€21,385	**67%**	€45,877	€48,433	€41,792
18%	€24,375	€26,689	€21,840	**68%**	€46,488	€49,140	€42,404
19%	€24,921	€27,040	€22,191	**69%**	€47,169	€49,855	€43,069
20%	€25,220	€27,425	€22,547	**70%**	€47,866	€50,310	€43,784
21%	€25,740	€27,833	€22,932	**71%**	€48,516	€51,168	€44,481
22%	€26,026	€28,249	€23,284	**72%**	€49,348	€52,000	€45,009
23%	€26,481	€28,600	€23,621	**73%**	€50,002	€52,702	€45,721
24%	€26,909	€29,029	€24,055	**74%**	€50,700	€53,560	€46,365
25%	€27,196	€29,497	€24,323	**75%**	€51,636	€54,496	€47,049
26%	€27,573	€29,991	€24,700	**76%**	€52,354	€55,235	€47,879
27%	€27,989	€30,264	€24,999	**77%**	€53,183	€56,291	€48,633
28%	€28,379	€30,706	€25,311	**78%**	€54,171	€57,215	€49,513
29%	€28,665	€31,174	€25,753	**79%**	€55,016	€58,149	€50,219
30%	€29,120	€31,525	€26,000	**80%**	€56,108	€59,176	€51,214
31%	€29,550	€31,980	€26,416	**81%**	€57,200	€60,190	€52,195
32%	€29,991	€32,335	€26,717	**82%**	€58,175	€61,425	€53,110
33%	€30,225	€32,695	€27,041	**83%**	€59,319	€62,556	€54,145
34%	€30,654	€33,046	€27,456	**84%**	€60,362	€63,973	€55,211
35%	€31,083	€33,540	€27,794	**85%**	€61,906	€65,390	€56,459
36%	€31,421	€33,943	€28,085	**86%**	€63,310	€66,794	€57,731
37%	€31,876	€34,327	€28,483	**87%**	€65,000	€68,453	€58,734
38%	€32,240	€34,814	€28,756	**88%**	€66,466	€70,109	€60,034
39%	€32,630	€35,100	€29,178	**89%**	€68,209	€72,158	€62,053
40%	€33,006	€35,563	€29,588	**90%**	€70,117	€74,018	€63,742
41%	€33,514	€35,932	€29,982	**91%**	€72,350	€76,371	€65,520
42%	€33,955	€36,358	€30,160	**92%**	€74,477	€79,222	€67,717
43%	€34,320	€36,712	€30,537	**93%**	€77,324	€82,359	€70,009
44%	€34,819	€37,115	€30,888	**94%**	€80,535	€85,841	€72,518
45%	€35,074	€37,557	€31,242	**95%**	€84,331	€90,317	€75,005
46%	€35,568	€37,986	€31,655	**96%**	€89,586	€96,343	€79,299
47%	€35,984	€38,350	€32,010	**97%**	€97,747	€104,840	€83,572
48%	€36,390	€38,909	€32,435	**98%**	€109,863	€120,023	€92,766
49%	€36,777	€39,312	€32,890	**99%**	€142,863	€152,945	€114,046
50%	€37,255	€39,858	€33,423				

* These figures are approximate estimates of annual earnings and were calculated by multiplying weekly earnings by 52.

FIRST VOTE OF THANKS PROPOSED BY SEAMUS MCGUINNESS, ECONOMIC AND SOCIAL RESEARCH INSTITUTE

In the context of a highly open economy, whose growth performance is driven by competitiveness, the assessment of the public-private sector pay gap is a central issue for public policy. For this reason it is important that we get as accurate a picture as is possible in terms of both the magnitude and distribution of any difference across sectors, which is described in the literature as a wage premium. Approaches to the measurement of the public sector pay premium vary from the job evaluation approach, which was adopted by the Public Service Benchmarking Body[1] and based on very small samples of data to those that rely on multivariate estimation using much larger samples.[2] The latter technique is adopted in the study presented here by the CSO statisticians. The obvious advantage of the multivariate approach based on a large and rich dataset is that it allows us to control simultaneously for a range of primary factors that determines an individual's earnings.

I am happy to propose a vote of thanks for the study discussed here, which presents a rigorous interrogation of a large and rich dataset that is designed to assess separately the impact of the (i) estimation method, (ii) specification, (iii) weighting strategy and (iv) sample choice on the estimated wage gap. The paper demonstrates that the estimated public-private sector pay premium is indeed sensitive to estimation approach, specification, sample restrictions and weighting choices made. For example, the premium was much higher when organisation size was excluded from the model and was marginally higher when the data were weighted to the population of employees and when the estimated premium was generated by the Oaxaca-Binder decomposition. The paper also confirms the finding of previous research that the public sector pay premium tends to be higher among lower skilled workers.

The paper makes it very clear that one needs to consider very carefully the question of model specification before embarking on a study of the public-private pay gap, as the specification influences the size of the gap. However, the paper raises serious questions for consideration regarding what the correct measurement approach actually is in this context. It is not correct to say that each and every specification presented in this study is equally plausible and, therefore, that each generated estimate is equally valid. It is important that analysts come to some consensus as to the most appropriate approach to estimating the public-private sector wage gap in Ireland, as the generation of a wide range of estimates could serve to confuse the policy debate. I will briefly consider the questions of the most appropriate model specification and weighting strategy in an attempt to shed some further light on these issues.

Model specification

The two features that are crucial here are whether or not to include Trade Union membership and Organisational Size. We consider each of these in turn.

Trade Union Membership

The authors confirm that the estimated public sector wage premium is sensitive to the inclusion of an organisation size control. However, a control for trade-union membership is also included within the specifications and, as Table 1 below demonstrates, the estimated premium is also sensitive to the inclusion of this control. The inclusion of the trade-union variable reduces the pay premium by around 5 percentage points (from 25.4 to 20.9 per cent); however, questions can be raised regarding the legitimacy of including this variable within the model. A primary concern with this control relates to the fact that trade-union membership in the public sector is often a consequence of public sector employment and, is therefore, highly collinear.[3] Furthermore, trade-union membership is generally included within wage equations to reflect the impact of local union level bargaining on wages; however, it is not clear that this bargaining effect operates in a parallel fashion across the public and private sectors in Ireland. For example, within the private sector we observe a standard wage premium of just below 7 per cent associated with trade-union membership, which is significant at the 99 per cent level. By contrast, the coefficient on the trade-union variable is not significant within a public sector

[1] The 2002 Benchmarking study relied on data from 3,991 public sector workers of whom 347 were interviewed in depth.

[2] See Kelly et al (2009) and Murphy et al (2008).

[3] Kelly *et al* (2009) report that trade union membership within the public sector was just below 80 per cent compared to 30 per cent in the private sector.

wage equation.[4] This is unsurprising given that since 1987 wages in the public sector have been set primarily through the national wage agreement and are therefore largely independent of trade-union density. Thus in addition to concerns relating to colinearity, the disparities in the nature of collective bargaining regimes between the public and private sectors in Ireland suggest that it may not be appropriate to include a trade-union control in a model of the Irish public-private sector pay gap.

Organisational Size

With respect to organisational size, the variable included within the models estimated in this paper relates to a binary control for an organisation containing more than 250 persons[5]. Within the data approximately 33 per cent of private-sector workers fall into this category compared to over 98 per cent of public-sector employees. The impact on the estimated premium in relation to organisational size is very large at around 10 percentage points (from 25.4 to 15.4 per cent) and is well in excess of the 2.5 percentage point reduction reported by Boyle *et al* (2004). Kelly *et al* (2009) argue that the asymmetric nature of the public and private sector organisation size distributions implies that organisation size cannot be considered a fixed characteristic and, on these grounds, it should be excluded from the model of the public-private sector pay gap. Kelly *et al* (2009) also question the logic of applying a private organisation size premium derived primarily from private sector productivity related economies of scale to virtually all public sector employees.[6]

Combined Effect of Trade Union Membership and Organisational Size

When both a trade union membership and organisational size control are included with a specification, the estimated wage premium falls by almost 50 per cent from 25.4 per cent to 13.1 per cent. Clearly, such an effect is non-trivial and has the potential to fundamentally alter the policy implications arising from an analysis of this type.

Table 1: Estimates of the Public-Private sector Pay Gap using the NES 2006

Specification*	Premium
Excluding Size and TU	25.4
Including TU only	20.9
Including Size only	15.4
Including Size and TU	13.1

* Models are estimated using weighted data and include comprehensive controls for a range of human capital and organisational controls.

However, the organisational size distributions implied by the data seem unusual as one cannot imagine that virtually all schools, healthcare establishments, social security, etc employ in excess of 250 persons. This is confirmed in Table 2 when we compare the size distributions from the October 2006 National Employment Survey (NES) with those from the 2001 European Community Household Panel Survey (ECHP) as reported by Boyle *et al* (2004). While the private sector data broadly align across both datasets the NES measurement approach is clearly very different with respect to the public sector and suggests further exploration to uncover the difference.

[4] Result available from the author.
[5] This analysis is based on the number of employee's variable contained within the NES.
[6] As with the trade-union variable, when the wage models are estimated separately for each grouping we observe a statistically significant and positive organisation size effect in the private sector equation only.

Table 2: Comparison of Organisation Size Distributions using the NES & ECHP

	NES 2006		ECHP 2001	
	Public	**Private**	**Public**	**Private**
3 to 4	0	4.4	6.1	12.1
5 to 19	0.2	28.5	19.4	26.2
20 to 49	0.3	14.8	20.6	21.8
50 to 99	0.4	10	13.3	11.7
100 to 499	4.8	18.8	24.7	19.6
500+	94.2	23.5	14.5	7.6

The explanation is that within the NES there is only one return for the primary and secondary education sectors, for each division of the civil service, and one return for the army, guards and prison officers (Table 3). Thus the NES data are not in fact capturing the size of primary schools, secondary schools and garda stations; instead they appear to be capturing the total number of primary school teachers, secondary school teachers and gardai employed within the public sector. These aggregates do not relate to any organisational size measure, as it is commonly understood, and, therefore, it is likely that wage models estimated with this variable will contain substantial measurement error given that the large organisational size premium will undoubtedly be applied to many public sector workers located in small schools, garda stations and civil service offices. In fact, if the ECHP size distribution approximates reality, as one would expect it to do given how it is constructed, then the NES data will incorrectly classify at least 60 per cent of the public sector jobs in terms of organisational size. Clearly this is not a position that one could easily defend and, on the grounds that the organisational size variable information is collected very differently across the public and the private sectors, it is not appropriate to include this variable in any wage models that estimated with NES dataset.

Table 3: Mean Organisational Size and CBR information by Public Sector Component (NES October 2006)

	Mean Size	**Number of Entity Returns***
Civil Service	3893	22
VECS and Institutes of Technology (ITs)	1136	36
University Sector	3077	6
Primary Schools	34084	1
Secondary Schools	17168	1
Garda	12954	1
Prison Officers	3219	1
Army	7141	3
Health	8327	35
Non Commercial Semi-states	912	25
Commercial Semi-states	5987	13
Local Authorities	1096	26

Note: * This comes from CBR information, which is a unique Business ID

Weighting and Sample Decisions

The authors demonstrate that the estimated wage premium will vary depending on whether the data are unweighted, weighted to account for sample attrition or weighted to be representative of the population of employees in employment. However, given that the question relates to the population of employees in employment it seems that the use of population weighted data are of extreme importance within this context. Presumably, this is the only population in respect of which we are interested in establishing the magnitude of any public-private sector pay premium. Furthermore, while the issue of weighting may be more trivial when using datasets designed to reflect the structure of the working population, it is a much more serious issue in this instance where we are using a population designed to reflect the structure of firms in Ireland in order to make inferences regarding the population of employees. In order to allay some of the concerns raised by the authors, as suggested by Fazio (2006), the regression analysis should, at the very least, be conducted using weight-conditioned variables. Failure to make any adjustment for the for the structure of the data will generate an inaccurate estimate.

The point regarding weighting is illustrated in Table 4 below by comparing the distribution of key characteristics from both a weighted and unweighted NES sample (October 2006) with those taken from the a sample of prime age employees from 2006 (Q2) Quarterly National Household Survey (QNHS), which is representative of the general Irish population in that year. While the weighted NES corresponds well to the QNHS distribution, it is obvious that graduates and professionals are heavily over-represented in the unweighted NES sample, while persons holding upper secondary qualifications and belonging to Craft and Protective, Construction and Personal Services occupations are under-represented. Clearly the results generated by an unweighted sample, or one weighted to account for non-response, will not be consistent with the key population of employees in employment. Consequently I would suggest that there is very little justification for using the un-weighted approach given the objectives at hand.

Finally, the authors demonstrate that the wage premium falls substantially when employees from Personal and Protective Service occupations are excluded from the data. The authors state that this restriction in some way controls for a lack of comparability between public and private sector jobs in these specific sub-sectors. However, the rationale for doing so is unclear within the current context as the techniques adopted in the paper are designed to provide a like-with-like comparison based primarily on individual human capital characteristics and are, as such, not designed to match across occupations. Furthermore, we could think of many possible such exclusions. Indeed, it is arguable that the problem surrounding a lack of comparability is more of an issue when looking across from the private to the public sector given the absence of activities such as construction and manufacturing within the public sector. Realistically such issues can only be addressed within a matching framework.

Table 4: Comparison of Weighted and Unweighted NES Data with QNHS Data

	NES	NES	QNHS
	Unweighted	**Weighted**	**Weighted**
Male	50.3	52.5	52.7
Primary or less	6.6	11	8.1
Lower Secondary	12.1	14.4	15.2
Upper Secondary	23.7	29.9	28.8
Post Secondary	11.2	11	10.5
Third-level No Degree	16.4	11.5	11.7
Third-level Degree	30	22.2	22.5
Other / Not Stated	-	-	3.1
Managers & Administrators	10.6	9.9	10.4
Professionals	22.8	12.6	12.2
Associate Professionals & Technical	10.2	9.3	8.9
Clerical & Secretarial	17.8	14.1	14.4
Craft & Related	6.6	12.6	12.7
Personal & Protective Services	7.3	12.7	12.3
Sales	6.8	9.6	9.8
Plant & Machine Operatives	8.5	8.6	8.5
Other Services	9.4	10.6	10.8
Agriculture	-	-	1.4
Mining and Manufacturing	16.1	15.3	15.2
Electricity	0.6	0.6	0.6
Construction	6.5	12.7	11.8
Wholesale & Retail	16.1	14.7	14.7
Hotels & Restaurants	4.7	6.5	6
Transport	4.8	5.5	5.7
Financial Mediation	6.1	5.1	4.9
Business Services	12.3	8.9	8.9
Public Administration	8.4	6	6.2
Education	8.4	7.9	7.7
Health	11.7	11.8	11.2
Other Services	4.4	5.1	5.9

Summary

The central contribution of the paper presented here is that it emphasises the importance of model specification in the context of attempts to measure the public-private sector wage gap. The analysis demonstrates that the estimated premium will vary substantially depending on the specification, estimation technique, sampling restrictions and weighting strategy adopted. Within this discussion I have attempted to draw further light on the issue by considering what the most appropriate approach to estimation might be. Based on the evidence presented here, I would argue that a population weighted estimate based on a specification that excludes controls for trade-union membership and organisational size represents the most appropriate approach to measuring the public-private sector pay gap in Ireland when using NES data.

SECOND VOTE OF THANKS PROPOSED BY JIM O'LEARY, NATIONAL UNIVERSITY OF IRELAND, MAYNOOTH

As one who co-authored the first paper published on public-private sector earnings differentials in Ireland, I have a kind of proprietorial feeling towards research on this topic and am especially happy to second the vote of thanks this evening.

Since the institution of the benchmarking process in 2000, the public-private wage gap has become a matter of great interest and not a little controversy. Both the interest and the controversy were quickened by the outcome of the first round of benchmarking and the opacity that characterised that process. Our 2004 paper, motivated in part by a sense of undischarged public duty on my part, since I had resigned from the Public Service Benchmarking Body several months before it published its first report, was an attempt to provide illumination where darkness seemed to be the official policy.

A central finding of our research was that, controlling for a set of personal attributes and workplace features, public service workers on average earned 13% more than their private sector counterparts in 2001. Another important conclusion of our research was that this premium was not significantly different from the premium we estimated to have obtained in 1994. All of this, despite the widespread perception that the public sector had 'fallen behind' the private sector during the ever-tightening labour market conditions of the Celtic Tiger. We had much else to report besides, including the outcome of quantile regressions which indicated that the public sector premium was highest towards the bottom of the income distribution and tapered away at the upper end.

I should say that, though I confidently expected that our research would establish that a public sector premium existed, I was taken aback by the size of the premia that emerged from our work. Bear in mind that figures in the range 4-8%, at the mean, were typical of other jurisdictions for which similar analysis had been carried out. My fear was that our estimates were so large that they would undermine the credibility of the research. Conscious of that risk, we erred on the side of inclusiveness in deciding on our explanatory variables (for example, by including size of workplace despite the theoretical case for excluding it), and we were extremely careful to test our results for a range of potential estimation biases and satisfy ourselves that our results were robust.

Since 2004, there has been a flurry of other papers published on the topic, including a series of papers from the ESRI team of which Seamus McGuinness is now a member - a series that started with the O'Connell and Russell paper of 2006 which looked specifically at public-private wage differentials amongst graduates. Particularly worthy of reference is Tony Murphy's work, conducted as part of the second benchmarking exercise and published in full with the 2007 Public Service Benchmarking Body Report. I single out this paper for special mention, not only because its author is a very distinguished econometrician, but also because I believe his research results (he confirmed the existence of a large public service premium) effectively shaped the outcome of the second round of benchmarking.

My judgement is that, in analytical terms, the principal increment of value added by the papers published since our 2004 paper, including that presented this evening, has been that of updating our work with reference to more recent (and much larger) samples. Our analysis was carried out on an ECHP sample which by 2001, the last year we covered, had dwindled to just over 4,000 of which our sub-sample numbered less than 1500. By contrast, the Foley and O'Callaghan analysis, for example, has been carried out on a 2007 NES sample with 35,000 households/individuals. One area where the superiority of the large sample is especially evident is in relation to the quantile regressions. The sample with which we were working was too small to permit meaningful quantile regressions to be carried out at the 95^{th} or 99^{th} percentiles. That aside, I don't believe that the more recent papers have broken fresh ground in terms of methodology, and I include this evening's offering in that assessment although its authors seem to be claiming otherwise.

Some of the specific comments I had intended to make about the technical aspects of Foley and O'Callaghan's work have already been made, and with a good deal more acuity and authority than I could muster, by Seamus McGuinness. I'm not going to go over the same ground and risk looking like his pale shadow. The one exception that I'm prepared to make concerns Foley and O'Callaghan's decision to analyse a sub-sample that excludes employees in personal and protective services on the grounds that the dominant occupational groups here (Gardai, prison officers and members of the defence forces) are public service workers for whom there are no private sector comparators.

A similar argument could be advanced in respect of nurses, teachers, university lecturers and others. Indeed, depending on how strictly one defines the notion of comparability, a similar argument could be advanced in respect of almost all public service workers. At the end of the day, if one regards security of tenure, performance unrelated pay and/or a defined benefit pension arrangement as fundamental defining characteristics of a job, then the generality of public service employees are in jobs that are not comparable to the jobs held by the generality of private sector workers.

This in turn means that getting exercised about the precise size of the public service premium or premia, as estimated according to the type of methodology represented by this evening's paper, is something of a waste of energy. What matters most here is the order of magnitude, not highly calibrated sensitivity to changes in the specification of an econometric equation. When one distils all the research that has been carried out over the past five years or so, it seems to me that there is one clear and consistent conclusion: on average and on a like-for-like basis (or as close to a like-for-like basis as it is feasible to construct given available data), public service workers were paid substantially more than private sector employees in every year for which analysis has been carried out up to 2007.

The really interesting questions that this result prompts are questions that we have scarcely started to think about, much less draw policy implications from. One of them is: why does a substantial public sector premium persist? Is it required to compensate for unobserved or impossible-to-measure features of public service jobs? Or, has it something to do with barriers to entry created by recruitment and promotion practices within the public service? Can we take the persistence of a large public sector premium as estimated therefore, to be evidence of a dysfunctional labour market? If so, should policymakers not be trying to eliminate the sources of such dysfunctionality?

I think it's time the debate about public service pay moved on from the attempted quantification of differentials to the consideration of these questions.