

## **The Public-Private Wage Gap: Issues of Specification**

### **Vote of thanks proposed by Seamus McGuinness, ESRI, at the Statistical & Social Inquiry Society of Ireland, 5 November 2009**

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<sup>1</sup> and based on very small samples of data to those that rely on multivariate estimation using much larger samples.<sup>2</sup> 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.

---

<sup>1</sup> The 2002 Benchmarking study relied on data from 3,991 public sector workers of whom 347 were interviewed in depth.

<sup>2</sup> See Kelly et al (2009) and Murphy et al (2008)?

### 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<sup>3</sup>. 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 wage equation<sup>4</sup>. 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<sup>5</sup>. 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<sup>6</sup>.

### 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

---

<sup>3</sup> 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.

<sup>4</sup> Result available from the author.

<sup>5</sup> This analysis is based on the number of employee's variable contained within the NES.

<sup>6</sup> 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.

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 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.

**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)**

	<b>Mean Size</b>	<b>Number of Entity Returns*</b>
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, are 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

	<b>NES</b>	<b>NES</b>	<b>QNHS</b>
	<b>Unweighted</b>	<b>Weighted</b>	<b>Weighted</b>
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