University of Dundee

University of Dundee

Dundee Discussion Papers in Economics 259

Chatterji, Monojit; Devlin, Stephen

Publication date:
2011

Dundee Discussion Papers in Economics

How Heavy Is A Job?: A Critical Survey of Job Evaluation as a Payment Device.

Monojit Chatterji & Stephen Devlin

How Heavy Is A Job?: A Critical Survey of Job Evaluation as a Payment Device.

Monojit Chatterji*

University of Dundee

Dundee DD1 4HN

Stephen Devlin

University of Edinburgh

Edinburgh EH8 9YL

August, 2011

Abstract. *This paper provides a summary of the purpose, practice and history of job evaluation in its primary role as a compensation system. It critically reviews the implicit assumptions made in the construction and application of the procedure with specific reference to PwC's evaluation of the salary structure of the Welsh Assembly in 2004.*

Keywords: job evaluation, compensation systems

JEL Classification: J31, J33

* corresponding author, m.chatterji@dundee.ac.uk

How Heavy Is A Job?: A Critical Survey of Job Evaluation as a Payment Device.

Public concern about high-end salaries has never been greater, both in respect of the private and the public sector. Recent examples of this concern are:

'*Bankers' pay soars 6% despite taxpayer anger*',   The Times headline, Aug 18th 2009

'*The justice of public-sector pay is a neuralgic issue at the best of times—which these are not.*', The Economist, Mar 17th 2011

'*Revealed: The new public service Fat Cats and why they're immune from the cuts*', Daily Mail headline, Mar 13th 2011

'*Ireland's 'morbidly obese cats' and runaway public sector pay*', Guardian headline, Dec 6th 2010

With salaries subjected to scrutiny more than ever, it is increasingly important that the process by which they are determined be understood and justifiable. Both public and private organisations now routinely rely on so-called "job evaluation" as a means of constructing an appropriate pay-scale and as such it is ever more necessary that we appreciate how this system works and that we recognise its limits. Only with such an understanding of the way in which salaries are set can we hope to have a meaningful discussion of their economic function. This paper aims to expound the details of job evaluation both in theory and in practice, and critically assess its shortcomings. In Section 1 below we describe the job evaluation system and in Section 2 we briefly outline the history and the usage of the system in both the private and the public sector. In Section 3 we theoretically analyse the often unstated but nonetheless implicit assumptions made by practitioners of the art of job evaluation. Section 4 applies the analysis of Section 3 to review a particular and important

case study, namely The Senior Salaries Review of the Welsh Assembly 2004. Section 5 concludes.

1. What is Job Evaluation?

Job evaluation is a system, employed within an organisation, which attempts to define the value or importance of each job relative to all others within the internal structural hierarchy. It is used as a tool for determining relative wage rates without bias and is undertaken or commissioned by an organisation primarily in an effort to improve the equity of the internal pay structure. It should be noted however that, in practice, the process actually determines *absolute* wages as well as a relative wage structure.

The initial implementation of the process will usually encompass most jobs within the organisation and the resulting structure will be periodically reviewed. The process begins with the write-up of a formal job description for each of the roles to be evaluated, with input from job-holders and a third party (usually a manager). This set of job descriptions is submitted to the evaluating panel which may be a private evaluation service provider or a committee of internal staff. The panel will have a predetermined system for evaluating and categorising each job.

This system consists of the following components:

1. a defined list of compensable factors under which the jobs will be evaluated and a set of progressive levels within each factor (e.g. knowledge, skills, working environment, responsibility)
2. a point value for each level of each factor – which can be construed as a "factor weight"
3. a conversion process by which the individual "factor weights" identified in 2 above are converted by some process of aggregation into a single overall job weight – which is the end result of the whole exercise

Using their chosen system the panel will produce a set of job weights for all roles in the company. The primary use of this set is to determine nominal salaries. Jobs may be grouped

into discrete salary bands which cover a specific range of job weights or conversion to salaries may be a continuous process. In either case the standard procedure is to observe job weights and corresponding salaries from the *external* labour market and use them as a basis for determining the appropriate *internal* salaries. Clearly, this is only justifiable if the external job weights used for comparison have been calculated using the exact same system used for the evaluation in question.

2. History and Incidence of Job Evaluation

Job evaluation is arguably over 120 years old, with the original prototypes emerging in 1871 through Ulysses S. Grant's Civil Service Commission and in the Principles of Scientific Management developed by Frederick Taylor in the 1880s (Figart 2000). The aim of these initial attempts at job evaluation was to 'rationalize personnel policy'.

In the 1920s, points and factor comparisons were incorporated into job evaluations to form the first of the modern systems (Benge, Burk and Hay 1941). But it was not until World War II that separate pay scales for men and women began to be replaced with a single scale. By this time the ideology of equal pay for equal work regardless of gender was gathering momentum, aided by the increasing influence of unions, and the culmination saw that idea enshrined in the Equal Pay Act of 1963 (USA) and 1970 (UK). Since then the use of job evaluation has grown rapidly as a convenient way to fulfil the legal requirements of equal pay legislation.

Initially job evaluation was most prevalent in the public sector, especially among unionised industries in which employees had more power to push for equitable pay. The use of job evaluation in the public sector is now practically ubiquitous.

Some examples of specific schemes in the UK are: the Single Status Agreement which covers most local council workers (cooks, street-cleaners, pharmacists, clerks, etc.); the Framework Agreement which is a similar programme for those employed in the higher education sector; and the Agenda for Change ('one of the most complex and lengthy pay negotiations in the world' Bevan et al. 2004) which ensures equitable pay throughout the NHS with the

exception of doctors, dentists and some senior management. Perkins and White (2010) discuss the context, implementation and outcomes of these three initiatives.

Since the Equal Pay Act of 1970 all employers must be able to prove that they do not discriminate on the grounds of gender or any other factor irrelevant to the job or else risk facing legal action. As such, the use of job evaluation has proliferated as a means to combat equal pay claims and is now in widespread usage in the private as well as the public sector.

The Hay Group website claims that '[t]housands of organisations – including more than half of the world's largest companies – rely on Hay Group's job evaluation methodologies.'[1] Since the Hay Group is only one of many companies offering job evaluation services, the total extent of the use of such services is presumably much greater. Indeed, an estimate from as far back as 1957 claims that two thirds of the American labour force was covered by some kind of scheme and more recently the Hay Group has estimated that between 82% and 96% of organisations evaluate jobs.

In the 1960s and 1970s, there was even suggestion of a national job evaluation scheme in the United Kingdom (Craig 1977; Bowey and Eccles 1975). Although this was always an extremely unlikely outcome considering the perceived failure of such an initiative in the Netherlands a decade or so earlier (Oettinger 1964), the fact that it was even a possibility indicates the influence and extent of job evaluation in the UK.

3. Theoretical Issues

A list of jobs are identified and indexed by $\{j = 1, 2, \ldots J\}$. The jobs are broken down into a set of factors indexed by $\{i = 1, 2, \ldots k\}$. A point score, a_{ij}, is assigned to each factor i of each job j. These are the individual factor weights. Assuming that this process can be undertaken accurately, the job weight, $S_j = f(a_{1j}, a_{2j}, \ldots a_{kj})$, is a monotonic increasing function of all factors. This is the basic premise and most general form of a job evaluation system. Essentially, any job weight function $f(\ldots)$ consists of an aggregation rule whereby for every job j, the individual factor weights a_{ij} are aggregated to produce a single value job weight S_j for that job. A relative wage structure is based on the principle that $w_{S1} > w_{S2}$ iff $S_1 > S_2$. In

---

[1] http://www.haygroup.com/uk/services/index.aspx?id=2424

other words, a given job 1 should be paid a higher wage than another job 2 if and only if the weight of job 1 is higher than that of job 2.

It is justifiable and reasonable to assume that the job weight function should be monotonic and increasing in all of its elements: a higher value for any individual factor weight should undeniably be associated with a higher overall job weight and hence higher pay, *ceteris paribus*. However, all job evaluation systems make an assumption when they choose one specific job weight function out of all possible functions. Frequently the choice is made without any theoretical motivation. Most commonly, as in the Senior Salaries Review of the Welsh Assembly 2004, the following simplistic job weight function is assumed:

$$S_j = f(a_{1j}, a_{2j}, ..., a_{kj}) = \sum_{i=1}^{k} a_{ij}$$

The use of such an additive function is highly restrictive and incorporates a number of strong assumptions. In particular, it excludes the possibility of an interaction between different factors. For example, we might expect heavy lifting (physical exertion) to vary in difficulty depending on the average ambient temperature (working environment) but this identification does not allow for that possibility. The additive function also makes the strong assumption that factors are entirely separable – is it really feasible that "skills" and "knowledge" have no overlapping areas? The identification also fails to explicitly model non-linearities in the contribution of individual factor weights to job weight.

Although most systems, like the Welsh Assembly Review, have a different number of levels in each individual factor and different points allocated to the equivalent levels across factors and therefore contain an element of weighting in the summation of attributes, this is still only one arbitrary identification plucked out of the infinite potential identifications. Two jobs in completely different areas with entirely dissimilar attributes are deemed to be of equal value purely because their attribute scores add up to the same total. This is not an assumption which should be taken lightly or left unjustified. If a different job weight function were chosen (for example, the sum of the square roots of attribute scores) this result would not hold. Consider the following example of a 5-factor job weight model.

There are two jobs being evaluated under five individual factors: responsibility, knowledge, skills, communication and intellectual demand. An expert panel judges that the factor weights for each job are as follows:

	Responsibility	Knowledge	Skills	Communication	Intellectual Demand
Job 1	16	0	0	0	0
Job 2	0	4	4	4	4

Using a simple additive function, the job weight of Job 1 is 16+0+0+0+0 = 16 and the job weight of Job 2 is 0+4+4+4+4 = 16. The two jobs have been evaluated as being equivalent in weight and would be paid the same wage. Now consider a different but equally valid aggregating function: the sum of the square roots of individual factor weights. The job weight of Job 1 is now 4+0+0+0+0 = 4 and the job weight of Job 2 is 0+2+2+2+2 = 8. Job 2 is now twice as "heavy" as Job 1 and would be paid a higher wage (though not necessarily double). This example illustrates the dependency of job weight relativities on the specific job weight function. Noting that there is a negative payment associated with the factor 'responsibility' in the second example above, it also shows how patterns may develop in which certain individual factors come to be associated with wages in a non-positive way.

This type of anomaly is corroborated in a study examining the effects on gender bias of implementing a job evaluation procedure in Washington state, O'Neill et al. (1989) find that the factor 'mental demands' is actually negatively related to pay and that 'accountability' is unrelated. As they conclude,

> the points for mental demands and accountability are summed together with the other categories as though they make equal (positive) contributions to the "worth" of the occupation. It would appear that there is much that is capricious in the process of assigning points to characteristics.

However, the fundamental issue goes beyond the valid criticism made by O'Neill et al. (1989, op. cit.). As our example above demonstrates, even if there is no capriciousness in determining the individual factor weights, there still remains the issue of whether they all should enter equally weighted in a separable additive function to determine the overall job weight. Studies by Lawshe and Satter (1944), Howard and Schutz (1952) and Weiner (1991) further suggest that factors themselves have been poorly specified.

Although the results of a job evaluation are the set of job weights and as such only imply an ordinal set of wage rates, in practice job weights are used to determine absolute wages. A common method for determining the wage of job j is to search the external market for a job j* such that $S_j = S_{j*}$ and then simply set $w_j = w_{j*}$. More realistically, job j will be compared to a group of jobs with similar weights, for example, in the way that PwC established a comparator group for jobs in the Welsh Assembly. This was done by choosing comparator groups from their proprietary pay database on the basis that they are 'broadly similar' in terms of job weight. The implication is that jobs of equal weight merit an equal salary and this rests on the assumption that there is a simple relationship whereby one job weight is associated with one and only one salary. Again this is no insignificant assumption: there exists no other determinant of salary. One arbitrary administrative metric is sufficient to determine an appropriate wage.

Furthermore, this assumption implies that one absolute pay *structure* is appropriate in all areas of the market. In other words, it assumes not only that $w_j = w_{j*}$ when $S_j = S_{j*}$ but also, by implication, that $w_j/w_k = w_{j*}/w_{k*}$ (where $S_k = S_{k*}$). It is implausible that the wage differential, which can be an important tool for managing incentives and achieving objectives, should optimally be constant over all firms and industries. A 2006 Confederation of British Industry report on public sector compensation schemes (CBI 2006) showed an appreciation of this idea, stating that 'the government needs to take a robust line to defend management's right to treat groups and individuals differently to meet organisational needs and on the basis of market evidence'. Job evaluation systems, which have been designed to treat all individuals and sectors equally, are incompatible with this need at a very basic level.

At a more fundamental level there is no clear agreement regarding what job evaluation actually measures. The standard claim is that it measures 'job worth', but what really is this? The value of a job is very difficult to define and the absence of any natural units of measurement makes it an inherently vague and ambiguous concept, without uniquely well defined upper and lower bounds. It is even harder to separate the value of a job from the value of a job-holder. Yet this is what job evaluation attempts to do. As such, there is a great deal of debate as to whether or not independent job evaluators are measuring the same thing and as to the consistency across such evaluators (Arnault et al. 2001; Madigan and Hoover 1986; Schwab 1980; Snelgar 1983; Doverspike 1983; Cunningham and Graham 1993).

A further criticism is that the inherent subjectivity and capriciousness of the job evaluation procedure (and the job evaluator), as well as measurement error problems, leaves the door wide open for inconsistent results. However, the evidence on this matter is mixed (Smith et al. 1989; Weiner 1991; Chenf et al. 1999).

Finally, a serious concern is the role that job evaluation systems play in affording legal protection to firms (Gilbert 2005). One of the most commonly cited advantages of job evaluation is the automatic invulnerability of the firm to equal pay claims. The desirability of this defence is open to doubt, at least from an employee's perspective, and especially so considering the continuing debate over the persistent existence of gender bias (Weiner 1991; Acker 1987; Schwab and Wichern 1983; Treiman and Hartmann 1981; Grams and Schwab 1985; Arvey 1986). Perkins and White (2010) argue that '[e]qual pay considerations, rather than productivity and cost savings, appear to have dominated' in the development of major UK public sector wage reforms, an idea paralleled by Fores and Heath (1970) who criticise the proposed implementation of job evaluation in the UK Civil Service by arguing that the proposal was motivated by staffing problems rather than pay problems. Quaid (1993) goes further, arguing that job evaluation is a 'myth' and that its real purpose is to legitimise the inequality inherent in salary structures by '[removing] the pay determination process from everyday discourse and [placing] it into the realm of the 'scientific' or the mystical'.

4. Case Study: The Senior Salaries Review of the Welsh Assembly 2004

The Senior Salaries Review Body commissioned PwC to conduct a review of Assembly Members salaries in 2004 (SSRB 2004). Job roles in the Welsh Assembly were assessed by PwC using their Monks Six Factor system. The six factors evaluated are: job complexity, intellectual demand, interpersonal communication, skills, knowledge, and nature of responsibility. The greatest point score variation was within 'job complexity': the highest was 248 for First Minister and the lowest was 61 for Assembly Member. The least point score variation was within 'intellectual demand': the highest was 52 for First Minister and the lowest was 47 for Assembly Member. Total points are defined as the sum of the component factor point scores. As argued above, this is an arbitrary choice of identification.

In order to analyse the wages of the members of the Welsh Assembly, PwC chose a group of 'representative comparators' from the private sector (chosen from their proprietary pay database on the basis that they are 'broadly similar' in terms of job weight) and expressed the salary under review as a percentage of the median salary of this comparator group. For example, Assembly Members were compared to 'Departmental Heads and Senior Professionals (HR, legal departments) in Service Sector Subsidiary Companies'. They found that the salary of Assembly Members was 87% of the median salary for this group. The sample size is only 20. The same process was undertaken for a public sector comparator group with the result that AM's salaries were 97% of the median salary of that group. They also made a job weight comparison between Welsh Assembly Members and Members of Parliament at Westminster, finding that the total job weights were 277 and 307 respectively. Primarily based on the fact that this job weight differential has decreased, the report recommended an increase in the basic wage of Welsh Assembly Members. Clearly, by employing this type of comparative analytics PwC implies that it is appropriate to compare the wages of two jobs with the same job weight, regardless of how different the jobs may appear to be.

The PwC evaluation makes all of the assumptions criticised above. It produces a measure of the absolute and relative importance of jobs within the Welsh Assembly which depends crucially on the arbitrary job weight function which has been chosen. There is no justification for this choice. Furthermore, the report makes a number of vague assertions regarding the evaluations which lack sufficient justification, for example, the claim that workload should not significantly affect total job weight.

The report also makes the assumption that the average wage structure encountered in the external market is optimal in the context of the Welsh Assembly, therefore failing to account for the possibility that the comparator group is not necessarily a suitable or appropriate benchmark.

5. Conclusion

Both the construction and application of a job evaluation system relies on unjustified (even unacknowledged) assumptions. In particular, two relationships are assumed. First, individual

factor weights contribute additively to an overall job weight (rather than any other equally reasonable function). Second, there is a simple relationship between job weight and salary which implies that the same salary is broadly applicable to all jobs of the same weight. Given the potential for these assumed relationships to affect the outcomes of the process, their significance should not be ignored.

The critical role which job evaluation now plays in the labour market has arisen as a gradual response to the perceived injustices of the market. While eradication of bias in salary structures is very much a desirable goal, one must wonder whether the current means of achieving it are the most appropriate. Job evaluation suffers from deep-rooted flaws and an abstraction from economic considerations, such as recruitment and retention factors. Given the costs of job evaluation exercises, their somewhat arbitrary methods are at the very least worrying, and at worst potentially damaging in that, used in isolation, they may well produce compensation outcomes that do not reflect genuine labour market considerations at all.

References

Arnault, E. J., Gordon, L., Joines, D. H. and Phillips, G. M. (2001). 'An Experimental Study of Job Evaluation and Comparable Worth'. *Industrial and Labor Relations Review*, 54 (4): 806-815

Acker, J. (1987). 'Sex Bias in Job Evaluation: A Comparable Worth Issue' in Bose, C. E., and Spitze, G. D. (ed.), *Ingredients for Women's Employment Policy*: 183-196, Albany: State University of New York Press.

Arvey, R. D. (1986), 'Sex Bias in Job Evaluation Procedures', *Personnel Psychology*, 39 (2): 315–335

Benge, E. J., Burk, S. L. H. and Hay, E. N. (1941). *Manual of Job Evaluation*. New York: Harper.

Bevan, S., Horman, L. and Turner, N. (2004). *Case Study Report Prepared for the Public Services Forum*. London: The Work Foundation.

Bowey, A. M. and Eccles, A. J. (1975). 'Relativities, National Job Evaluation and Collective Bargaining', *Journal of Management Studies*, 12: 83–94.

CBI (2006). *For What It's Worth: Managing Public Sector Reward 2008-11*. London: Confederation of British Industry.

Chenf, S. N., Orazem, P. F., Mattila, J. P. and Greig, J. J. (1999). 'Measurement Error in Job Evaluation and the Gender Wage Gap'. *Economic Inquiry*, 37: 181–194.

Craig, C. (1977). 'Towards national job evaluation? Trends and attitudes in Britain and in the Netherlands'. *Industrial Relations Journal*, 8: 23–36.

Cunningham, J. B. and Graham, S. (1993). 'Assessing the Reliability of Four Job Evaluation Plans'. *Canadian Journal of Administrative Sciences*, 10: 31–47.

Doverspike, D., Carlisi, A. M., Barrett, G. V. and Alexander, R. A. (1983). 'Generalizability Analysis of a Point-Method Job Evaluation Instrument'. *Journal of Applied Psychology*, 68 (3): 476-483.

Figart, D. M., (2000). 'Equal Pay for Equal Work: the Role of Job Evaluation in an Evolving Social Norm'. *Journal of Economic Issues*, 34 (1): 1-19.

Fores, M. and Heath, J. (1970). 'The Fulton Report: Job Evaluation and the Pay Structure'. *Public Administration*, 48: 15–22.

Gilbert, K. (2005). 'The Role of Job Evaluation in Determining Equal Value in Tribunals: Tool, Weapon or Cloaking Device?'. *Employee Relations*, 27 (1): 7-19.

Grams, R. and Schwab, D. P. (1985). 'An Investigation of Systematic Gender-Related Error in Job Evaluation', *Academy of Management Journal*, 28 (2): 279-290.

Howard, A. H. and Schutz, H. G. (1952). 'A Factor Analysis of a Salary Job Evaluation Plan'. *Journal of Applied Psychology*, 36 (4): 243-246.

Lawshe, C. H. Jr. and Satter, G. A. (1944). 'Studies in Job Evaluation. I. Factor Analyses of Point Ratings for Hourly-Paid Jobs in Three Industrial Plants'. *Journal of Applied Psychology*, 28 (3): 189-198.

Madigan, R. M. and Hoover, D. J. (1986). 'Effects of Alternative Job Evaluation Methods on Decisions Involving Pay Equity'. *The Academy of Management Journal*, 29 (1): 84-100.

O'Neill, J., Brien, M. and Cunningham, J. (1989). 'Effects of Comparable Worth Policy: Evidence from Washington State'. *The American Economic Review*, 79 (2): 305-309.

Oettinger, M. P. (1964). 'Nation-wide Job Evaluation in the Netherlands'. Industrial *Relations: A Journal of Economy and Society*, 4 (1): 45–59.

Perkins, S. J. and White, G. (2010). 'Modernising pay in the UK public services: trends and implications'. *Human Resource Management Journal*, 20 (3): 244–257.

Quaid, M. (1993). 'Job Evaluation as Institutional Myth'. *Journal of Management Studies*, 30 (2): 239-260.

Schwab, D. P. (1980). 'Job evaluation and pay setting: Concepts and practices'. in E. R. Livernash (ed.), *Comparable worth: Issues and alternatives:* 49-78. Washington, D. C.: Equal Employment Advisory Council.

Schwab, D. P., and Wichern, D. W. (1983). 'Systematic bias in job evaluation and market wages: Implications for the comparable worth debate', *Journal of Applied Psychology*, 68 (1): 60-69.

Smith, B. N., Hornsby, J. S., Benson, P. G. and Wesolowski, M. (1989). 'What is in a Name: the Impact of Job Titles on Job Evaluation Results'. *Journal of Business and Psychology*, 3 (3): 341-351.

Snelgar, R. J. (1983). 'The Comparability of Job Evaluation Methods in Supplying Approximately Similar Classifications in Rating One Job Series'. *Personnel Psychology*, 36 (2): 371-380.

SSRB (2004). *National Assembly for Wales Review of Pay and Allowances 2004*. London: Senior Salaries Review Body

Treiman, D. J., & Hartmann, H. J. (ed.) (1981). *Women, work, and wages: Equal pay for jobs of equal value*, Washington, D.C.: National Academy Press.

Weiner, N. J. (1991). 'Job Evaluation Systems: A Critique'. *Human Resource Management Review*, 1 (2): 119-132.