DISCUSSION PAPER SERIES

IZA DP No. 3065

The Determinants of Performance Appraisal Systems: A Note (Do Brown and Heywood's Results for Australia Hold Up for Britain?)

John T. Addison
Clive R. Belfield

September 2007

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

The Determinants of Performance Appraisal Systems: A Note (Do Brown and Heywood's Results for Australia Hold Up for Britain?)

John T. Addison
Queen's University Belfast and IZA

Clive R. Belfield
Queens College, City University of New York

Discussion Paper No. 3065
September 2007

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3065
September 2007

ABSTRACT

The Determinants of Performance Appraisal Systems: A Note
(Do Brown and Heywood's Results for Australia Hold Up for Britain?)

This paper offers a replication for Britain of Brown and Heywood's analysis of the determinants of performance appraisal in Australia. Although there are some important limiting differences between our two datasets – the AWIRS and the WERS – we reach one central point of agreement and one intriguing shared insight. First, performance appraisal is negatively associated with tenure: where employers cannot rely on the carrot of deferred pay or the stick of dismissal to motivate workers they will tend to rely more on monitoring, ceteris paribus. Alternatively put, when the probability of job separation is greater, the influence of deferred compensation diminishes. Second, there is also some suggestion in the data that employer monitoring and performance pay may be complementary. However, consonant with the disparate results from the wider literature, there is more modest agreement on the contribution of specific HRM practices, and still less on the role of job control. Finally, there is no carry over to Britain of the structural determinants identified by Brown and Heywood.

JEL Classification: J5, L23, M5

Keywords: performance appraisal, monitoring, deferred compensation, performance pay, HRM practices, worker tenure, unionism

Corresponding author:

John T. Addison
Queen's University Management School
Queen's University
Belfast BT7 1NN
Northern Ireland
United Kingdom
E-mail: j.addison@qub.ac.uk

1. Introduction

In one of the most frequently accessed papers in this *Journal*, Brown and Heywood (2005) examine the determinants of performance appraisal systems in Australia using the 1995 cross section from the AWIRS.[1] Performance appraisal is defined as "formal appraisal of non-managerial workers at least once a year", and it is reported that this monitoring of employees is associated with shorter-tenure workers and workers who have greater influence over productivity. It is also found that certain human resource management practices are also associated with increased use of performance appraisal, and conversely for union density.

In the present paper, we investigate whether or not the same empirical regularities can be observed for Britain, using the 2004 Workplace Employment Relations Survey. Our results are mixed. Like Brown and Heywood, we find that the use of performance appraisal is associated with similar worker characteristics. We also report some support, albeit weaker, for their human resource variables. But there is little carry over to Britain of the job control arguments, while the structural factors found important by Brown and Heywood either play no role or are even contraindicated in the British case. Nevertheless, the crucial and anticipated result is that performance appraisal is negatively associated with tenure. A further important common finding is the intriguing result that monitoring and (a measure of) performance pay may be complementary.

2. Theoretical Motivation

Brown and Heywood offer four broad determinants of performance appraisal systems. Of these the most important is *workforce characteristics*. All the variables subsumed under this heading reflect expected tenure. The longer is expected tenure, so the argument runs, the less the role of current remuneration and the greater the importance of deferred pay. *Vulgo*: the need for extensive monitoring is reduced.[2] Three proxies for expected tenure are used: actual tenure (strictly the proportion of workers with more than 5 years tenure within the establishment), the share of female workers, and labour turnover. A fourth variable is the share of casual workers but this is an inverse proxy on the grounds that their increased use implies greater expected tenure for the core labour force.

The remaining broad determinants of performance appraisal are *job control*, *human resource management (HRM) practices*, and *structural factors*. As far as job

control is concerned, the authors rely on two arguments. First, the greater the degree of control exercised by workers over the job, the larger the scope for performance appraisal, provided that *the results of these choices are not immediately obvious*. (Brown and Heywood, 2005: 663; emphasis added). Second, where managers form a large share of the workforce, they are more likely to use performance appraisal, especially, where they are trained in HRM management.

Turning to HRM practices, the authors identify worker training, job redesign, and joint consultative committees. These are seen as commonly bundled with performance appraisal, respectively generating increased losses in productivity from reduced effort (or increasing the need for evaluation), increasing the benefits of performance appraisal stemming from differences in the length of time that a worker occupies a specific job, and via organizational change promoted by the disclosure of private information by workers. In short, greater opportunities for productivity improvement increase the need for performance appraisal. The authors also link individual performance related pay to performance appraisal, arguing that the benefits of the latter are likely to be greatest in circumstances where performance can be measured. (We return to the latter, potentially controversial argument below.)

Finally, the structural factors identified by Brown and Heywood concern the size of the organization (number of employees, number of HRM professionals, and the share of labour costs in total costs) and unionism. The authors predict a positive association between the use of performance appraisal and establishment size by virtue of the longer chains of command involved and on economy of scale grounds. A direct association is also expected for HRM professionals, where the appointment of such professionals is linked to a standardization and formalization of industrial relations. And, on balance, they anticipate a negative association between performance appraisal and worker representation in unions because of potential union resistance to tailoring rewards to individual appraisals.

The key tenure variable(s) arguments apart, controversy attaches to a number of Brown and Heywood's priors. For example, theory suggests that performance appraisal is likely to be introduced where it is difficult or costly to monitor effort (i.e. inputs), whereas the use of individual performance related pay – one of the HRM arguments recall – might seem to imply that inputs can be monitored.[3] Another contentious issue is the *net* productivity of individual practices, as well as the question of their bundling (admitted by the authors to be contested in the literature (see Wood,

1999; Addison, 2005). Given the uncertainty in the literature, it would not be surprising if one set of practices deemed important in one jurisdiction/dataset were found inconsequential in another. Again, workers are pre presumably given substantial decision making autonomy when the worker is capable of making better decisions than a supervisor or manager, which may in turn mean that reward mechanisms such as profit related pay and employee share ownership may substitute for monitoring.

3. Data

Our data are taken from the 2004 Workplace Employment Relations Survey. WERS 2004 is the fifth in a series of workplace surveys conducted since 1980.[4] In this paper, we use the cross-section survey of managers. Survey weights are constructed to relate the sample to the population of British workplaces and to control for response bias.

There are a number of differences between the AWIRS and the British WERS. In the first place, the performance appraisal question in the WERS identifies whether or not the non-managerial groups in question have their performance formally appraised (albeit without specifying a time interval) and if so the extent of that appraisal. Information on the extent of coverage is provided in six categories: 100%, 80-99%, 60-79%, 40-59%, 20-39%, and 1-19%. Some 24% of the raw sample report that none of their employees conduct performance appraisal, while 61% percent claim that all their workers are covered. The remaining 15 percent are thus spread across the five other bands, with roughly 3% in each. The (main) AWIRS by contrast simply inquires of the full sample whether or not there is such formal appraisal *at least once a year*. In practice, 70% of the establishments in sample undertake performance appraisal. Only for a sub-sample of the 1995 cross-section – 559 out of 1,642 establishments – is information provided in the AWIRS on the extent of performance appraisal. The categories here are ‘none’, ‘some’, most’, and ‘all’ (respectively set at 0, 1, 2, and 3 in Brown and Heywood’s subsequent ordered probit exercise).

There are two consequences of this difference in the two datasets. First, to achieve rough correspondence between studies our definition of the presence or otherwise of formal performance appraisal in the WERS will be guided by the actual mean value recorded in the AWIRS. Therefore, we take as cut-off for the presence of performance appraisal those circumstances in which 60 percent or more of an

establishment's non-managerial workforce is reported as being appraised; correspondingly, all other situations are treated as an absence of performance appraisal.[5] Second, because of the tiny spread of coverage rates across the WERS intervals, we only comment on the results of fitting ordered probit equations to the data in passing.

(Table 1 near here)

The definitions and frequencies of performance appraisal and all the other WERS variables are given in Table 1. We measure high tenure by the proportion of workers over the age of 50 years, whereas in the AWIRS this is a measured by the proportion of workers having remained more than 5 years with the establishment. There are no real differences between the two studies, however, as regards the share of casual and female workers or in the labour turnover measure. The same is also true of the two *job control* variables. Reflecting the frequencies, however, our dummies assume the value of one where the employer-assessed degree of employee influence is "a lot" rather than "some to a lot" in the case of the Brown and Heywood measures.

As far as the various HRM practices are concerned, the worker involvement measure (viz. presence of a joint consultative committee (JCC)) is the same across data sets so that the differences pertain to "individual performance related pay", "train workers", and "job redesign" in that order. First, we substitute (the presence of) "profit related pay" for individual performance related pay not only because we lack that exact measure/definition in the WERS but also because individual payment by results schemes proper imply that effort input is not an issue. That said, we shall report in passing the results of substituting individual merit pay and individual merit pay and/or payment by results for our preferred variable. Next, we measure job redesign by changes in workplace procedures experienced at the workplace rather than by "job redesign programmes introduced by management in the last year". And our training dummy takes the value of one where there is provision of more than 2 days of training for the establishment's most experienced workers as opposed to the AWIRS definition of "formal training of workers designed to develop their skills."

Finally, turning to the *structural factors* there are no differences in the establishment size and union variables (while labour costs are defined in bands with the omitted category corresponding to the Brown and Heywood measure). But the HRM professionals variable in our case is simply the share of persons trained in

human resource management, whereas in the AWIRS it is defined as the number of managers with industrial relations and human resource management titles.[6]

4. Findings: Similarities and Differences

Results of fitting the basic probit equation to the British data are given in Table 2. Recall that, unlike Brown and Heywood, our dependent variable is not the presence or otherwise of any performance appraisal among non-managerial workers but rather a measure based on a coverage rate of 60% or more. The basis of this cut-off is that 70% of Australian establishments offered formal appraisal of non-managerial employees and 66% of their British counterparts had coverage rates above 60%.

(Table 2 near here)

The most striking result is the strong performance of the workforce characteristic variables. Like Brown and Heywood, we find that workers with longer tenure are less likely to be subject to performance appraisal and that female workers are more likely to be monitored in this manner (although the latter association is not always well determined). Interestingly, our turnover variable achieves more traction than in Brown and Heywood: in two specifications higher turnover rates are accompanied by greater use of performance appraisal. The only source of real disagreement concerns the role of casual work: we find that the greater the share of workers on fixed-term contracts, the more likely is widespread performance appraisal, whereas Brown and Heywood anticipate the opposite result arguing that atypical work for some makers for greater employment continuity for others, namely, the core labour force. Our result seemingly brings out the more direct relation. Overall, then, we obtain rather strong evidence to support Brown and Heywood's proposition that there is greater need to monitor those with less tenure because one cannot rely on deferred compensation – or the threat of dismissal – to do the trick. *Vulgo*: employers with shorter expected tenure workers will be more likely to engage in extensive monitoring.

That being said, the remaining evidence is less supportive of the relevance for Britain of the other arguments put forward by Brown and Heywood. This is particularly evident for the job control variables. Only one of these arguments – percent managers – evinces a well-determined positive association with the use of performance appraisal. Unlike Brown and Heywood, then, we do not detect any association between either HRM training or employee influence over the pace of

work/task allocation and performance appraisal. In the latter case, whatever the abilities of workers in these areas, there is no suggestion in the British data that this influence over productivity encourages the use of monitoring. We earlier suggested that this might reflect the fact that workers are more knowledgeable than supervisors.

However, rather more in line with Brown and Heywood we do discern some positive association between HRM practices and performance appraisal. Specifically, the coefficient estimates for employee involvement (via JCCs) and worker training – although not job redesign – are positive and statistically significant. However, we should be cautious in concluding that the HRM practices identified by Brown and Heywood necessarily form part of a bundle with performance appraisal. This is not only because unlike these authors we are unable to divine any association between job redesign and performance appraisal but also because of the parsimonious range of HRM practices identified.

But note that that our replacement argument for individual performance related pay – namely, profit related pay, which substituted for individual performance related pay – was positively associated with performance appraisal. We did not anticipate this association but it does rather suggest that reward systems (if not necessarily individual performance related pay) may be complementary with monitoring individuals.

Given the uncertainty as to what constitutes “individual performance related pay”, we experimented with two other payment measures contained in the WERS. The first of these is merit pay, defined as pay related to a subjective assessment of individual performance by a supervisor or manager. When this dichotomous variable was substituted for the profit related pay dummy its coefficient was again positive but never statistically significant. When we added in payment by results (defined as any method of payment where pay is determined by the amount done or its value, exclusive of profit related pay) to create a composite individual measure, the coefficient estimate abruptly changed sign and was statistically significant in one specification. In each case, the coefficient estimates for the other variables were to all intents and purposes unaffected. A tentative conclusion would be that one does not need monitoring where input can be measured and that payment by results proper should bear a negative relation to the presence of performance appraisal. For its part, merit pay is more analogous to profit related pay, even if it its association with performance appraisal proved more elusive. But the real bottom line must be that our

understanding of the monitoring of inputs on the one hand and outputs on the other remains rudimentary at this stage.

The role of the structural factors is as follows. First, *pace* Brown and Heywood, there was no association between establishment size and the use of performance appraisal and no suggestion of any role for labour costs. Also, again in contrast to Brown and Heywood, union presence was positively associated with the deployment of performance appraisal. We would interpret the strength of this latter result (and the weakness of the union density result) as perhaps implying that unionism increases the need for employer monitoring but that union strength may be sufficient to enable unions to resist this pressure.

As a final exercise, we attempted to replicate the ordered probit exercise conducted by Brown and Heywood. Note that despite the fact that information on coverage is available for only one third of the 1995 wave of the AWIRS, Brown and Heywood were able to confirm most of their previous results. The major exception was structural factors, where none of the coefficient estimates was now statistically significant. Our problem, it will be recalled, is the compressed frequency distribution of performance appraisal among plants. Not surprisingly, our results – not reported here but available from the authors upon request – were poor. Only in the cases of the share of female workers and the number of persons trained in HRM practices were the coefficient estimates statistically significant (and of the expected sign).

5. Interpretation

The evidence assembled from the AWIRS and the WERS indicates that the occurrence of performance appraisal can be attributed in no small part to worker characteristics; in particular, to the expected tenure of workers. For low tenure individuals, deferred rewards cannot be expected to bind workers and firms. As a result, firms employing substantial numbers of low tenure workers are cet. par. more likely to rely on monitoring. This is an important result, even if the British evidence points to a broad threshold effect rather than a continuous association between share of lower tenure workers and extent of monitoring more evident in (a subset of) the Australian data.

Beyond the tenure-performance appraisal nexus, commonalities between the two datasets are largely limited to HRM practices. Few of the job control variables identified in the Australian study as determinants of performance appraisal proved

statistically significant in the WERS, although the role of worker control over the pace of work and job allocation is probably best approached using objective measures rather than management perceptions. For their part, the structural factors were with one exception statistically insignificant in the British case. The exception was union presence where interestingly the direction of association was opposite in sign to that reported in the Australian case.

Of the better-performing HRM arguments, perhaps the most intriguing result was the suggestion in the WERS that monitoring and reward systems might be complements rather than substitutes after all. This is a new finding and one also reported for Australia by Brown and Heywood, whose measure (based on incentive pay) admittedly differs from our own.[7] Evidently, the devil is in the detail regarding this complementarity of reward system and performance appraisal.

Apart from the pressing need to further explore this latter, Tayloresque relation, future research on the determinants of performance appraisal has to address the largely unexplored contributions of the skills and technology used in production (as well as type of management) on the one hand and firm performance outcomes on the other. Brown and Heywood also say as much.

Endnotes

1. The authors also supplement this analysis with an investigation of changes in the use of performance appraisal between 1990 and 1995 for a subset of establishments. Their fixed effect logit estimates are broadly supportive of their cross-section results.

2. That said, the authors recognize that, to the extent that the appraisal mechanism is "to engage workers in organizational goals and guide worker development" (p. 662), the direction of causality may be reversed.

3. Most obviously so if piece rates included in the measure of individual performance related pay.

4. Full details of WERS 2004 are contained in Kersley et al. (2006).

5. As a practical matter, redefining the dependent variable as the presence of any performance appraisal did not change the results reported below.

6. In line with Brown and Heywood, we use sector dummies and distinguish between commercial and all establishments. We do not, however, use occupational dummies.

7. Bryson and Freeman (2007) also report a positive association between monitoring and performance pay in the WERS 2004.

References

Addison, J. T. (2005). 'The determinants of firm performance: unions, works councils, and employee involvement/high performance work practices'. *Scottish Journal of Political Economy*, 52: 406-450.

Brown, M. and Heywood, J.S. (2005). 'Performance appraisal systems: determinants and change'. *British Journal of Industrial Relations*, 43: 659-679.

Bryson, A. and Freeman, R.B. (2007). *Doing the Right Thing? Does Fair Share Capitalism Improve Workplace Performance*? Employment Relations Research Series No. 82. London: Employment Market Analysis and Research, DTI.

Kersley, B., Alpin C., Forth J., Bryson, A, Bewley H., Dix G., and Oxenbridge, S. (2006). *Inside the Workplace: Findings from the 2004 Workplace Employment Relations Survey*. London: Routledge.

Wood. S. J. (1999). 'Getting the measure of the transformed high performance organization'. *British Journal of Industrial Relations*, 37: 391-417.

TABLE 1
Descriptions and Frequencies of Variables

Variable	**Definition**	Frequency
Performance appraisal	Non-managerial employees have their performance formally appraised (>60%) =1, 0 otherwise	0.66
High tenure	Proportion of workers aged over 50	0.18
Percent casual	Proportion of workers on fixed term contracts	0.05
Percent female	Proportion of female workers	0.42
Turnover rate	Proportion resigned, dismissed, and made redundant in previous year	0.23
HRM training	Managers are trained in employee relations = 1, 0 otherwise	0.44
Influence over pace of work	Individual workers have a lot of control over the pace at which they work = 1, 0 otherwise	0.60
Influence over task allocation	Individual workers have a lot of involvement in decisions over how their work is organized = 1, 0 otherwise	0.66
Percent manager	Proportion of workers who are managers	0.12
Receipt of PRP	Do any of the employees at this workplace receive profit-related payments or profit-related bonuses =1, 0 otherwise	0.44
JCC	Presence of joint consultative committee =1, 0 otherwise	0.30
Job redesign	Workplace has experienced changes in workplace procedures = 1, 0 otherwise	0.52
Train workers	Receipt of greater than 2 days of training per year for the most experienced workers = 1, 0 otherwise	0.77
Size	Number of workers (log)	4.26
HRM professionals	Share of persons trained in HR	0.05
Labour costs 1	Labor costs are 0-25% of total costs = 1, 0 otherwise	0.26
Labour costs 2	Labor costs are 25-50% of total costs =1, 0 otherwise	0.33
Labour costs 3	Labor costs are 50-75% of total costs =1, 0 otherwise	0.18
Union presence	Presence of recognized union at workplace = 1, 0 otherwise	0.34
Union density	Percent workers covered by union	0.18
Internal labour market	Workers hired either exclusively or mainly through internal vacancies = 1, 0 otherwise	0.15
Commercial	Workplace trades with other organizations =1, 0 otherwise	0.77
n		*1591*

TABLE 2
Determinants of Performance Appraisal (Probit Estimation)

	(1)	(2)	(3)	(4)
Variable	**Full sample**	**Commercial sector**	**Commercial sector**	**Commercial sector**
High tenure	-1.4714	-1.0823	-1.0045	-0.9554
	(0.3721)***	(0.4711)**	(0.4904)**	(0.4914)*
Percent casual	0.4354	1.0664	1.0879	1.0879
	(0.3216)	(0.4698)**	(0.4958)**	(0.4957)**
Percent female	0.7926	0.5706	0.3993	0.4102
	(0.1943)***	(0.2543)**	(0.2703)	(0.2722)
Turnover rate	0.2106	0.5302	0.6753	0.6241
	(0.1943)	(0.3242)	(0.3414)**	(0.3402)*
HRM training	-0.1144	-0.1218	-0.1773	-0.1676
	(0.1360)	(0.1670)	(0.1705)	(0.1701)
Influence over pace of work	-0.0165	-0.0463	-0.0215	-0.0166
	(0.1242)	(0.1592)	(0.1620)	(0.1620)
Influence over task allocation	0.1328	0.1892	0.1250	0.1469
	(0.1270)	(0.1555)	(0.1550)	(0.1555)
Percent manager	0.7391	1.6006	1.6949	1.6843
	(0.4633)	(0.7540)**	(0.7806)**	(0.7856)**
Individual PRP	0.2426	0.2768	0.2540	0.2583
	(0.1185)**	(0.1484)*	(0.1547)	(0.1553)*
JCC	0.3453	0.4537	0.4761	0.5338
	(0.1795)*	(0.2245)**	(0.2118)**	(0.2089)**
Job redesign	0.1704	-0.0292	-0.1227	-0.0928
	(0.1151)	(0.1453)	(0.1495)	(0.1502)
Train workers	0.7748	0.9071	0.8828	0.8637
	(0.1249)***	(0.1684)***	(0.1744)***	(0.1750)***
Size	0.0564	-0.0122	0.0024	0.0282
	(0.0588)	(0.0765)	(0.0780)	(0.0780)
HRM professionals	0.5828	0.1660	0.2802	0.3180
	(0.3030)*	(0.3497)	(0.3273)	0.3337)
Labour costs 0-25%	-0.2514	-0.2973	-0.1467	-0.1465
	(0.1677)	(0.2033)	(0.2196)	(0.2197)
Labour costs 25-50%	-0.0789	-0.1605	-0.1384	-0.1622
	(0.1527)	(0.1934)	(0.2091)	(0.2115)
Labour costs 50-75%	-0.1628	-0.2693	-0.3336	-0.3210
	(0.1831)	(0.2513)	(0.2637)	(0.2642)
Union presence	0.5258	0.5875	0.4050	
	(0.1409)***	(0.1529)***	(0.1921)**	
Union density				0.0567
				(0.3180)
Commercial sector	0.4406			
	(0.1211)***			
n	1591	1049	1049	1049

*Statistically significant at 10%.
** Statistically significant at 5%.
*** Statistically significant at 1%.
Notes: Robust standard errors in parentheses. Columns 3 and 4 include sector dummies.