Preliminary Draft, Comments Most Welcome.
Not to be Quoted Without the Authors' Permission.

Are Public Sector CEOs Different? Leadership Wages and Performance in Schools

Timothy Besley* and Stephen Machin**

October 2006

* *Department of Economics and STICERD, London School of Economics*

** *Department of Economics, University College London and Centre for the Economics of Education, London School of Economics*

Acknowledgements

We are extremely grateful to Joan Wilson for her excellent research assistance and to the Department for Education and Skills for funding this project (under Strand 2, on teachers, in the Centre for the Economics of Education projects from 2002-2004) and for providing access to the administrative data we utilise.

1. Introduction

The past twenty years has seen a big increase in the study of corporate governance and the role of Chief Executive Officers (CEOs) in the management of firms. It is now well understood that even when firms are subject to competitive pressures, the structure of governance matters to firm performance. There can be little doubt that leadership in the private sector matters and this shows up in the way in which CEOs are rewarded.

A striking feature of the existing literature is the development of a wide variety of data sets that are used to look at the way CEOs are remunerated and determinants of CEO turnover. This has been connected to the performance of the organizations for which they work, with an intention of developing tests of simple incentive models that have predictions on the way in which executives are rewarded. The evidence base resulting from this work now illuminates academic and policy literatures on corporate governance - see, for example, Murphy (1999), Hall and Leibman (1998) and Bertrand and Schoar (2003).

Issues of leadership are equally important in the public sector, not least because market pressures are usually much weaker. Yet there is almost no literature amongst economists on the study of public sector leaders. It is our aim to begin to rectify this by looking at the usefulness of economic approaches in thinking about the role of leadership in public sector institutions. Indeed, there are a number of places in the public sector where the role played by leaders is potentially critical. The issue of leadership is particularly important in the education sector where key figures such as head teachers and vice chancellors perform a leadership function analogous to the CEO in a private corporation.

Looking at the roles of key players in the education sector ties in well with the recent expansion of work on the economics of education (see Machin and Vignoles, 2005). Some of this work is concerned with the way in which the development of education systems and educational policy in the last twenty years or so, including aspects of increased choice and accountability, has impacted upon the efficiency and equity of the education sector. The role of potentially influential people holding key positions of responsibility has not been addressed, however, and this seems to be an important omission from a wider research perspective.

One key motivating factor for this work is a concern with understanding whether there is a substantial difference between the role of leaders and key personnel in the public and private sectors. There is certainly much popular discussion of the notion that those attracted to work in the public sector are somewhat different to their private sector counterparts. Besley and Ghatak (2005) have argued that this can be thought of in terms of the idea of motivated agents - the fact that individuals who typically work in the public sector will do so to achieve non-pecuniary benefits. While some of the private sector CEO literature refers to things like "empire building" as a motive for management, there is a broad consensus that views CEOs as being money minded and disliking effort. In principle, motivated agents will behave quite differently. They are likely to respond less to monetary incentives and will be motivated more by exercising control over the organization's "mission".

We would like to take these ideas forward by trying to confront them with data and by trying to test specific predictions that result from such models. This paper offers a first step in this direction. In the next section of the paper we develop and build up a

fairly simple, incentives based framework for thinking about and understanding the role played by head teachers in schools. Section 3 of the paper describes the data we use, drawing on administrative data on head teachers in secondary schools in England. Section 4 presents results that try to evaluate the key predictions of the theoretical model presented earlier. In Section 5 we offer some conclusions, including some pointers as to where we think research should be heading so as to generate a better understanding and intuition of how public sector leaders operate and how they can impact on the institutions they head up.

2. Theoretical Issues

In this section, we will discuss the underlying theoretical reasons for a link between school performance and teacher pay. The standard model for looking at this in the existing literature is the career concerns model of Holmstrom (1999). He posited a situation in which the effort of an employee depends on the effort-wage relationship determined in market equilibrium. Employees who put in more effort "learn more" about their abilities and if successful get higher paying jobs. As emphasized by Gibbons and Murphy (1992), this model predicts a "life-cycle" relationship between pay increases and performance. An employee towards the end of their career is less likely to put in effort as signalling ability has a payoff over a shorter time horizon. The basic career concerns model does not offer any predictions about the relationship between turnover and performance - firms can offer lower wages to poorly performing employees rather than dismissing their workers. The career concerns model has become the industry standard in thinking about CEO pay and performance. Many empirical studies are structured around explaining its implications (see Murphy, 1999).

The approach taken here is somewhat different, emphasizing the role of matching in creating a relationship between wages and performance. Schools of different qualities match with head teachers of different qualities. There is a market price for high quality heads which schools must pay. If a school performs well then this will be a signal of head teacher quality necessitating a wage premium relative to untried head teachers.

We develop a model in which all incentives are implicit as in the pure career concerns framework. The model has overlapping generations of teachers and a fixed stock of schools. Time is infinite and denoted by $t = 1,...,T$. Teaching careers last two periods and are denoted by $s \in \{1,2\}$. In period one of a career, head teachers are matched randomly with schools and are tried out. The market then gets to observe the performance of the head. In period two the head-teacher is either confirmed in the job or fired given the market determined wage of high quality heads. We will assume that wages are determined in a competitive labour market.

There is a continuum of schools, indexed by z, uniformly distributed on $[0,1]$. The productivity of school z in period t is denoted by

$$\pi(z,\delta)$$

where $\delta \in \{0,1\}$ is a measure of head teacher quality where '0' is low quality and '1' is high. We assume that $\pi(z,\delta)$ is increasing in z and δ with $\pi_{z\delta} > 0$. The last of these assumptions says that head teacher quality and school quality are complements.

There is a measure of head teachers of size $n > 1$. Each head teacher has an outside opportunity of u. The fact that $n > 1$ implies that the wages of incoming heads will always be bid down to u in a competitive labour market.

There are three kinds of head teachers in every generation:

- A fraction γ of head teachers is good, always setting $\delta = 1$.
- A fraction β of head teachers are opportunists, setting $\delta = 1$ if it is in their interest to do so and $\delta = 0$ otherwise. Specifically, they face a cost of $c \in [0, C]$ of putting in effort in period of life. This cost is uniformly distributed across the opportunists.
- A fraction $(1 - \beta - \gamma)$ of head teachers are bad, never achieving $\delta = 1$.

The good head teachers are like motivated agents in the sense of Besley and Ghatak (2005) - putting in effort regardless of whether it is in their interest to do so. The opportunists are introduced to capture crudely the effect of the standard career concerns model in which an agent may slack towards the end of their career. One fundamental difference between this model and the career concerns model of Holmstrom (1999) is that head teachers are assumed to know their type. A head-teaching career is a pair of teaching decisions $\{\delta_1, \delta_2\}$, one in each phase of life.

We assume that:

(1) $$\gamma > (\gamma + \beta)^2$$

which says that there are few enough opportunists in the population.

There are two periods to teaching careers. In period one, head teachers are randomly assigned to school vacancies. At the end of period one, their performance is observed and they are either rehired or fired. At the end of period two, they retire. The model is solved backwards. In the second period, all good head teachers set $\delta_2 = 1$. All opportunists who are re-hired set $\delta_2 = 0$. The period two hiring decision is made understanding this.

We will look for a stationery equilibrium in which a fraction $\phi (\leq \beta)$ of

opportunistic head teachers put in effort in period one of their careers. The wage for a head teacher with $\delta_1 = 1$ is w. The positive complementarity between z and δ in our model imply that the best schools will hire such teachers. Thus, the fraction of schools that hires such teachers is:

$$1 - \hat{z} = \gamma + \varphi.$$

That is, all the good head teachers are hired in good schools along with the opportunists who set $\delta_1 = 1$.

The equilibrium wage of a good quality period one head is determined by the demand for quality by the school located at $\hat{z}$. That school compares the benefit from hiring a good head at w with hiring an untried head. Let μ be the probability that an experienced head teacher with $\delta_1 = 1$ will perform in period two of their career. Assuming that the alternative is to hire an untried head teacher, the equilibrium wage must solve:

$$\mu\pi(1-(\gamma+\phi),1)+(1-\mu)\pi(1-(\gamma+\phi),0)-w=(\gamma+\phi)\pi(1-(\gamma+\phi),1)+(1-\gamma-\phi)\pi(1-(\gamma+\phi),0)-u.$$

Then:

$$w = u + (\mu - \gamma - \phi)\left[\pi(1-(\gamma+\phi),1) - \pi(1-(\gamma+\phi),0)\right]. \tag{2}$$

Thus the gain to setting $\delta_1 = 1$ depends on the gain in reputation $(\mu - \gamma - \phi)$ and the gain in productivity of the marginal school $\left[\pi(1-(\gamma+\phi),1) - \pi(1-(\gamma+\phi),0)\right]$.

To complete the description of the market equilibrium, we need to determine ϕ and derive the beliefs of schools: μ. The latter are determined using Bayes rule. Hence,

$$\mu = \frac{\gamma}{\gamma + \phi}. \tag{3}$$

As long as $\mu > (\gamma + \phi)$, there is a wage premium. This is implied by (1).

To determine the proportion of opportunists who put in effort, observe that the wage gain from setting $\delta_1 = 1$ is $w - u$. Then, the fraction of opportunists who set $\delta_1 = 1$ is:

$$\varphi = \beta \min\left\{\frac{w - u}{C}, 1\right\}. \tag{4}$$

Then, a stationery competitive equilibrium in the market for head teachers is a pair (w, φ) which solves (4) and:

$$w = u + \left(\frac{\gamma - (\gamma + \varphi)^2}{\gamma + \varphi}\right)\left[\pi(1 - (\gamma + \varphi), 1) - \pi(1 - (\gamma + \varphi), 0)\right].$$

Thus, the wage premium depends on how many opportunists there are in the population. An increase in γ holding $\gamma + \varphi$ constant raises wages. This thought experiment can be interpreted as increasing the fraction of good head teachers while reducing the fraction of opportunists.

The model can be used to generate some predictions for the market for head teachers.

- **Prediction 1:** There is a positive correlation between head teacher wages and school quality. This is due to the assumed positive complementarity between school and teacher quality.
- **Prediction 2:** Head teachers with poor track records are fired rather than being offered wage cuts.
- **Prediction 3:** Performance of head teachers deteriorates towards the end of their career if $\beta > 0$. Otherwise performance is constant across careers.

Predictions 2 and 3 potentially differentiate the model put forward here from the standard

career concerns model. In this model, if everyone were opportunistic, there cannot be a situation in which $\delta_1 = 1$ induces a future wage increase.

3. Data Description

Data Sources

The data we use is a very rich administrative data source on teachers in English schools that contribute to the Teachers Pension Scheme, known as the Database of Teacher Records (DTR). We look at head teachers in English secondary schools (attended by children aged 11 through 16) over the years 1995-2002. There are approximately 4,000 observations on head teachers per DTR year. We concentrate on full and part time head teachers currently in service, but also include those that may have left the teaching service during the year, either to transfer to another occupation or for retirement.

To test our key hypotheses of interest we need to match data on head teachers to school level data. To do so we establish a school identifier based on DTR records of the school establishment number and Local Education Authority code for each year. This identifier is used to match in annual school performance tables statistics, such that our dataset encompasses an array of key indicators of school characteristics. These include measures of pupil performance at the end of secondary school, total pupil numbers and school types (e.g. whether community, foundation or voluntary).

In selecting only head teachers from the DTR dataset we found some cases where schools had more than one head teacher observation for the year.[1] This required us to

[1] This could be due to updating and recording problems within the DTR where head teachers might not be accurately traced as they change schools within the LEA.

choose the head teacher we perceived was most likely to have been in leadership in that year for that school, such that we had only one head teacher observation per year. Manual selection of the head teacher observation took place for 2,430 schools and was based on a range of indicators: salary information (highest salary), tenure (longest tenure), and last recorded service date.[2]

The dataset includes generated variables to indicate head teacher turnover and employment activity, where the former includes turnover both within and between the years and the latter establishes the prior employment status of those head teachers that are not in our sample from the beginning (1994). Through this we are able to establish promotional activity - that is whether a current head teacher was promoted from within the same school or another school and from which teaching level they were promoted – and which head teachers are entirely new to the profession.

Descriptive Statistics

Some descriptive statistics are given in Table 1. Annual means are given for three variables: head teacher salary; the headline figure of secondary school performance usually considered, namely the proportion of students obtaining 5 or more grade A*-C examination results in the final year 11 exams (the General Certificate of Secondary Education, GCSE); the proportion of head teacher turnovers.

The Table shows the average head teacher annual salary to be around £58,000 in 2002, having risen from just over £41,000 in the first year we consider, 1995. The headline school achievement figure, the proportion of children getting 5 or more A*-C GCSEs, rises from .41 in 1995 up to .49 by 2002. And each year, around one in ten head

[2] Note that the DTR dataset does include a withdrawal indicator to reflect whether a teacher is no longer in the pension scheme. This variable is not used to establish the true head teacher as withdrawal from the

teachers leave their job, although there are some year-to-year fluctuations in the turnover rate.

4. Results

Basic Empirical Models Testing Prediction 1

In this section we use the DTR data to consider the predictions arising from the model presented in Section 2. We consider Prediction 1, namely a positive correlation between head teacher wages and school performance, by estimating a pay-performance equation like those estimated in the sizable literature on CEO pay in the private sector (e.g. Jensen and Murphy's, 1990, of US chief executives). The estimating equation relates the log of head teacher salary to school performance between 1995 and 2002. The starting point is simply a regression on the pooled data of the following form for head teacher i in school s in year t:

$$\ln(S_{ist}) = a_1 + b_1 P_{ist} + \varepsilon_{ist}$$

where S is salary, P is school performance and ε is an error term which for the moment comprises a white noise error, u_{ist}, and a set of year dummies, T_t, such that $\varepsilon_{ist} = u_{ist} + T_t$. The parameter of interest is b_1 which, in parallel to the private sector CEO literature, measures the sensitivity of head teacher salaries to school performance.

Specification (1) of Table 2 reports estimates of this basic pooled model. A positive and significant estimate of b_1 is uncovered. At least at first glance, it does seem that salaries of head teachers are associated with higher observed school performance. In other words, leaders of better performing schools are paid more.

---

pension scheme may not equate with withdrawal from teaching service.

This said, we clearly want to refine this empirical model in several dimensions. First, as also recognised in the extensive CEO literature (see Murphy's, 1999, survey), some characteristics of workplaces are correlated with both salaries and with performance. The clearest, and most robust, finding in the CEO work is that firm size is a significant determinant of CEO compensation and that controlling for size does affect the estimated pay-performance sensitivity. As part of our interest is to see how well the private sector approach applies to the public sector, we have added school size to the model:

$$\ln(S_{it}) = a_2 + b_2 P_{it} + c_2 \ln(N_{it}) + \varepsilon_{it}$$

where N is the number of pupils in the school.

As specification (2) of Table 2 shows size is a strong determinant of head teacher salaries. We should not be surprised about this as head teacher salaries are directly linked to the size of school they head up. But it is interesting that the association with school performance, whilst considerably reduced in magnitude, remains positive and significant. Over and above school size we are able to uncover a significant association between head teacher salaries and pupil performance.

We can also use the nature of our data, where we observe head teachers and schools over time, to control for time-invariant teacher and/or school characteristics by incorporating teacher and/or school fixed effects. We can thus amend the error structure of the estimating equation to have following components: $\varepsilon_{ist} = a_i + a_s + u_{ist} + T_t$, where a_{i} and a_{s} are teacher and school fixed effects.

With this kind of error structure we can estimate more sophisticated models which look at the relationship between changes in salary and changes in school performance for

teachers and/or schools over time. These net out the unobserved heterogeneity linked to particular, time-invariant, characteristics of teachers or schools. Of course, the models are only identified if some head teachers leave or join schools, but our earlier descriptive statistics revealed that we observe head teacher turnovers in our data.

Specifications (3) and (4) of Table 2 respectively add a full set of school fixed effects to the pooled models given in (1) and (2). Thus the analysis is now on the relationship between changes in salaries and performance within schools over time. Prediction 1 - a positive correlation between salary and is performance – is re-confirmed in the school fixed effects models, where the pay-performance sensitivity is significant and positive in specifications (3) and (4).

Specification (5) in Table 2 refines the school performance measure, by also adding in a measure of poor performance, namely the proportion of children getting no GCSE passes. Whilst attracting a negative coefficient, this is statistically insignificant, and so does not add any additional explanatory power. Moreover the estimated coefficient on the headline measure remains significant and positive (and of very similar magnitude).

In the final three specifications of Table 2 we additionally include the head teacher fixed effects thereby estimating the most general model our data permits. Prediction 1 is still borne out by these demanding models. Head teacher salaries and school performance do seem to be complements and the notion of a pay-performance sensitivity that is stressed in the private sector literature on CEOs does seem to translate over to the public sector setting.

The interpretation of this positive association does need some discussion:

whether it reflects sorting of more able head teachers into better performing schools, or whether better performing head teachers can be persuaded not to leave via movements in their salaries. To get a handle on this is to look at whether there is a pay-performance relation on entry to a new position as compared to pay-performance relations for incumbents. Table 3 shows that, once one conditions on school size, that there is no pay-performance relation for new hires, and thus that the pay-performance relation looks more in line with the retention hypothesis.

Turnover Models: Predictions 2 and 3

Given that we have uncovered results in line with Prediction 1 from the incentives based theoretical model, we can now consider Predictions 2 and 3. Like the work on private sector CEOs that tests the implications of principal-agent models, Prediction 2 asserted that one should see head teachers in poorly performing schools being more likely to leave. This is considered in columns (1) and (2) of Table 4 where the probability of head teacher turnover is related to school performance in the previous year. The models show a strongly significant negative coefficient on lagged school performance that is in line with Prediction 2. This remains the case controlling for age of teachers as reported in column (2). Thus, head teachers in schools with lower academic achievement are more likely to leave. This is also consistent with the salary results above suggesting that incentives operate to keep higher paid teachers in place in better performing schools.

Prediction 3 is to do with whether career concerns type notions can be applied to head teachers. So far we have looked at this in terms of differential turnovers linked to performance that vary across the individual head teacher's career. We have done so by estimating turnover models for three age groups: those less than 45, those between 45

and less than 55, and the 55 and over group. The idea of opportunistic behaviour resulting in career concerns having an impact (i.e. effort dwindling in the later years) is suggested by the results, in that turnover seems more sensitive to performance for the oldest (55 and over) group.

5. Conclusions

The recent corporate governance literature puts weight on the role of leadership in shaping corporate strategy. However, the literature to date has not considered whether public sector CEOs are different. In this paper, we have used data on head teachers in England to investigate these issues.

Our results show that pay is sensitive to school performance for this group. This is consistent with a model in which there is a competitive market for high quality head teachers where schools and teachers match. This creates a wage premium to being a high quality head teacher whether or not head teachers respond to incentives by putting in effort. However, in our results on turnover we find some evidence that effort considerations may be important, with turnover of head teachers occurring later in their careers being more strongly linked to lower levels of prior pupil achievement.

The results suggest the importance of recognizing that labour markets for head teachers, and for leaders in the public sector in general, are important. They can create a wage-performance gradient without the use of explicit incentives. Much recent discussion of teacher incentives seems to be premised on a view that teachers were not previously subject to wage incentives. However, as these results emphasise, these discussions must be placed in a market context.

References

Bertrand Marianne, and Antoinette Schoar, [2003], "Managing with Style: The Effect of Managers on Firm Policies," Quarterly Journal of Economics, 118(4), 1169-1207.

Besley, Timothy and Maitreesh Ghatak, [2005], "Competition and Incentives with Motivated Agents," forthcoming in the American Economic Review.

Gibbons, Robert and Kevin J. Murphy, [1992], "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence," Journal of Political Economy, 100(3), 468-505.

Hall, Brian and Jeff Leibman, [1998], "Are CEOs really paid like bureaucrats?", Quarterly Journal of Economics, 113(3), 653-691.

Holmstrom, Bengt, [1999], "Managerial Incentive Problems: A Dynamic Perspective", Review of Economic Studies, 66(1), 169-82.

Jensen, Michael and Kevin Murphy, [1990], "Performance Pay and Top-Management Incentives", Journal of Political Economy, 98(2), 225-64.

Machin. Stephen and Anna Vignoles, [2005], "What's the Good of Education? The Economics of Education in the United Kingdom", book manuscript, forthcoming Princeton University Press

Murphy, Kevin J., [1999], "Executive Compensation," in Orley Ashenfelter and David Card (eds), Handbook of Labor Economics, Amsterdam: North-Holland.

Table 1:
Descriptive Statistics

	1995	1996	1997	1998	1999	2000	2001	2002
Average salary	41371	42667	44655	46209	48222	51612	54851	57919
Proportion Getting 5 or More A-C* GCSE Grades	.409	.423	.427	.440	.458	.470	.478	.493
Turnover proportion	.082	.095	.093	.130	.095	.097	.134	.128

Notes: Salary data (in £'s) from Database of Teacher Records; Proportion Getting 5 or More A-C* GCSE Grades from School Performance Tables; Turnovers calculated from Database of Teacher Records. Salary and GCSE grades data based on sample size of 22587, corresponding to 3001 schools with at least five continuous time series observations and 4767 headteachers between 1995 and 2002. Turnovers based on sample size of 22312, corresponding to 3001 schools and 4753 headteachers between 1995 and 2002. For the 22587 sample the balance is: 5 observations – 79 schools; 6 observations – 307 schools; 7 observations – 570 schools; 8 observations – 2045 schools.

Table 2:
Basic Pay-Performance Regressions For Secondary School Headteachers, 1995-2002

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Proportion Getting 5 or More A-C* GCSE Grades	.127 (.009)	.036 (.007)	.070 (.010)	.054 (.009)	.053 (.009)	.040 (.006)	.028 (.005)	.027 (.005)
Log(Number of Pupils in School)		.222 (.004)		.147 (.010)	.147 (.010)		.144 (.006)	.144 (.006)
Proportion Getting No GCSE Passes					-.015 (.011)			-.015 (.008)
School Fixed Effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Headteacher Fixed Effects	No	No	No	No	No	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	.469	.704	.898	.902	.902	.965	.967	.967
Sample Size	22587	22587	22587	22587	22587	22587	22587	22587

Notes: standard errors (clustered on teacher id) in parentheses.

Table 3:
New Headteacher (First Year in Post) Salary Equations

	(1)	(2)
Proportion Getting 5 or More A-C* GCSE Grades	.073 (.013)	-.031 (.011)
Log(Number of Pupils in School)		.199 (.005)
Year Dummies	Yes	Yes
R-Squared	.475	.680
Sample Size	2369	2369

Notes: standard errors in parentheses.

Table 4:
Headteacher Turnovers

	(1)	(2)	(3)	(4)	(5)
	All	All	Age<45	Age≥45 and Age<50	Age≥55
Proportion Getting 5 or More A-C* GCSE Grades (t-1)	-.092 (.010)	-.123 (.011)	-.131 (.026)	-.112 (.012)	-.186 (.037)
Age of Teacher		.011 (.001)	.004 (.003)	.002 (.001)	.036 (.003)
Year Dummies	Yes	Yes	Yes	Yes	Yes
Sample Size	22312	22312	2491	14628	5193

Notes: marginal effects from probit models; standard errors (clustered on teacher id) in parentheses.