

Work in progress

**Promotion to hospital consultant in NHS Scotland:
does effort matter?**

Kostas Mavromaras and Anthony Scott

Health Economics Research Unit

University of Aberdeen

Abstract

In an internal labour market such as the NHS, and with salaried payment of hospital doctors, incentives for effort are determined by structure of medical careers. The aim of this paper is to investigate the factors that influence promotions of medical staff to consultant in the Scottish NHS. The paper tests the hypothesis that effort in previous periods, measured using whole time equivalents (WTEs), influences the incidence of promotion in the current period. The paper uses a unique panel data set which contains individual level information on all NHS hospital doctors in Scotland from 1991 to 2000. A panel model of promotion is estimated that controls for unobserved heterogeneity at specialty level. Females are less likely to be promoted than males. The probability of promotion increased throughout the 1990s. Effort does not influence promotion to consultant. This may be due to a pre-determined career structure where bureaucratic rules dominate, rather than individual heterogeneity. Specialties are found to exhibit considerable differences in their rate of promotion over and above the differences explained by the characteristics of the doctors in them. The results have implications for the structure of rewards within medical careers.

Acknowledgements

Thanks to the Information and Statistics Division of the Scottish Executive Health Department for providing the data. The Health Economics Research Unit is funded by the Chief Scientist Office of the Scottish Executive Health Department. The views are those of the authors.

1. Introduction

The National Health Service is the largest employer in Great Britain, employing around 1.2 million people and accounting for 5% of total employment (Elliott et al., 2002). Although health care is becoming increasingly technologically advanced, it remains a labour intensive industry with almost 70% of total expenditure on labour. The drive to improve the performance of public services is a key aspect of government policy. During 2004, the national employment contracts of hospital consultants, GPs, nurses and all other non-medical staff will change. Workforce issues have become a major area of policy development. This is set against the context of the Wanless report that justified unprecedented increases in NHS expenditure and therefore increases in the demand for labour and other factor inputs (Wanless, 2002).

The medical profession plays a key role in the delivery of NHS services. There were 84,000 doctors employed in the NHS in 2000, about 7% of the NHS workforce. Although a relatively small proportion, doctors are responsible for all decisions regarding patient care, and so have a significant influence on costs and quality. The role of incentives is key in improving performance and productivity.

Although there are employment opportunities for doctors outside of the NHS, this is an option for few except the most senior doctors. Labour markets internal to the NHS therefore play a major role. The internal labour market and the structure of careers, promotions, administrative rules and internal remuneration within the NHS therefore become key motivating factors in a public service that is under great pressure to enhance performance and productivity. The operation of the internal NHS labour market is particularly important as recent policy changes include a new consultants contract, ambitious targets for the recruitment of new consultants, and changes to the medical career structure and medical training. Within a fixed system of salaried payment and in a dynamic context, each of these can potentially have major effects on behaviour, effort and productivity.

Internal labour markets and incentives in health care

The analysis of internal labour markets in firms has developed since the 1990s, and has grown from the literature on agency and incentives into a broader literature on

personnel economics (Prendergast, 1999; Lazear, 2001). This literature recognises that performance pay schemes are rarely observed in practice because of costly monitoring and complex jobs (Burgess and Metcalfe, 1999). Research on internal labour markets has therefore focussed on the incentives for effort that are implicit in the hierarchical structure of promotions and careers within specific firms, and the use of promotion tournaments as the most efficient type of incentive contract (Lazear and Rosen, 1981).

The idea of internal labour markets suggests that workers already within the firm are protected to an extent from external labour market pressures due to limited ports of entry and exit, and the existence of administrative rules, customs and career paths specific to the firm (Doeringer and Piore, 1971). Key empirical studies by Baker et al (1994) and Treble et al (2001) have provided evidence as to the extent to which internal labour markets exist. It is therefore important to examine how general these results are in a large public sector organisation such as the NHS, where stronger evidence of an internal labour market might be expected.

In health care, previous research on incentives has focussed on a traditional principal-agent approach and the effect of explicit performance incentives, and the effects of changing remuneration systems (e.g. fee-for-service and salaried payment) on behaviour and performance (Gosden et al., 2001). This empirical research is more common in the United States and in primary care in the NHS (Robinson, 2001; Gravelle et al., 2001).

However, a key issue for hospital doctors in the UK is the provision of incentives for effort where they are paid by salary under a centrally determined remuneration system. There is a well-defined career structure for doctors and wages are determined centrally by administrative rules and the nature of jobs, rather than by worker's human capital characteristics, although ultimately this is an empirical question.

The observation that hospital doctors work very long hours and invest heavily in human capital under fixed salaries may be partly due to the patient's utility being an argument in the doctor's utility function, which has been cited as a reason why strong performance incentives are not observed in health care (Mooney and Ryan, 1993).

However, doctors' own motivations, particularly for future income and reputation through career advancement and promotion to hospital consultant, may also be an important factor in observing high levels of hours worked. Promotion to a hospital consultant confers prestige, responsibility and higher financial rewards. The higher financial rewards include bonuses, through distinction awards, and access to private practice. Hospital consultants can be awarded distinction awards, the highest of which can almost double their salary. Although it is not the case that performance will double, arguably the scheme provides incentives for newly appointed consultants and junior doctors to supply effort.

The aim of this paper is to explore the relationship between promotions and the supply of effort. We use a unique panel dataset on individual hospital doctors in the NHS in Scotland to examine whether effort in previous periods influences promotion to consultant in the current period.

2. Theoretical framework

Tournament theory has been used to examine the incentives in promotion tournaments. The main prediction is that the larger the difference between the reward of the winner and the reward of the loser (the wage spread), then the higher the level of effort. Promotion is a function of the relative ranking of individuals' output, dependent on effort and a random term reflecting luck. Attitudes to risk and the cost of effort are also relevant (Lazear and Rosen, 1981; Audas et al., 2002). This is useful, in that changing the size of the prize at a given level in the promotion hierarchy (e.g. a new consultant contract) will change incentives to supply effort lower down in the promotion hierarchy. Tournament theory therefore focuses on the relative pecuniary rewards between different levels in the promotion hierarchy.

Since pay is fixed for hospital consultants, there will be little variation in rewards across different promotion opportunities in any given year. A more general prediction of tournament theory is that promotion depends on relative effort and output. This assumes that promotion is not automatic, but depends on individuals' human capital and on their supply of effort. Automatic promotion may occur when employers base promotion on seniority and 'time served' in a post, or other factors that are easily observable. Bureaucratic promotion rules may be a rational response to the high costs

faced by employers in measuring and observing effort or output in complex jobs. The main hypothesis we test is therefore whether effort influences promotion or whether promotion is automatic and pre-determined.

2. The Data

Data were obtained from ISD Scotland. The Medical and Dental Census (MEDMAN) includes individual level data for all doctors and dentists working in hospital and community health services. The accuracy of the data relies on NHS Trusts who supply the data to ISD on an ongoing basis. For example, every time a doctor leaves or joins the NHS, a return is sent to ISD and the database updated. At the end of September every year, the annual dataset is produced by ISD and used to report on the size and nature of the workforce. Each annual dataset between 1990 and 2000 was merged using a unique individual identifier to create a longitudinal dataset. Although the data extends before this period, definitions and the structure of the data have not been standardised with later years. This provides data on between 8000 and 9000 hospital doctors per year. A descriptive analysis of this data can be found in Elliott et al. (2002).

The data contain information on age, ethnic origin, country of birth, country of qualification, year of qualification, grade, specialty, type of contract, whole time equivalent (WTE), date started in post, date appointed to grade, provider (hospital) and health board.

The analysis concentrates on the promotion to consultant in any given year, t . The year 1990 had to be excluded because of incompatibilities in the coding of specialties. All consultants at the start of the observation period were excluded from the sample as they had already been promoted. The sample was then restricted to all those who were in the following grades at some point in time during the observation period: registrar, senior registrar, specialist registrar, associate specialist, and staff grade. This represents the pool of those doctors who could potentially be promoted to consultant during the period. Individuals who were observed for only one period were excluded. A promotion variable was generated which took the value one for those who were not consultants in the previous period and consultants in the current period (promotion) and zero for those who were not consultants in both the previous and current period

(no promotion). The promotion variable was used as the dependent variable in all estimations.

Experience was calculated for each period as the number of years after graduation.¹ Eight experience dummies were generated for experience from 9 to 16 years, in order to investigate in a semi-parametric context the influence of experience on the probability of promotion. Four country of qualification dummies were generated; England, Scotland, rest of the UK, India, with the rest of the world as the reference category. A year variable was included to control for any time trend in promotion probabilities. An ordinal variable representing the size of the origin employer is included to test for the possibility that the size of employer may influence future promotion probabilities (small, medium large). This is based on the total number of NHS staff employed by the Health Board.² The skill mix in the individuals' Health Board of promotion was also included. The ratio of several staff categories to the number of Doctors employed in each health board was included for PAMs, Nurses and Managers. A gender variable, which takes the value of one for females, was also included.

An important part of this analysis is the investigation of whether effort influences promotion or not. The main information in the data relating to the level of individual effort is the individual whole time equivalents (WTE) supplied in a given year.³ Two measures of effort are used. The first is the absolute level of WTE in the year before promotion. This assumes that employers observe labour supply and promote those with the highest WTE. The second measure of effort is whether an individual increased their WTE between the two periods prior to promotion (between $t-1$ and $t-2$) This assumes that individuals increase their own labour supply before promotion in order to signal their willingness to put in higher effort. The use of information from past periods avoids the problem of endogeneity in the promotion equation.

¹ This variable may be biased for sub-groups which may have taken longer or shorter career breaks, but the data does not have any relevant information.

² There was no consistent hospital identifiers over time due to the gradual creation of NHS Trusts during the early 1990s, and the mergers of NHS trusts in the late 1990s.

³ This is based on contracted hours worked, not actual hours, and so may underestimate the actual effort supplied. We assume that those who work more than their contracted hours is distributed randomly across the sample. At best, this variable will distinguish those who work part time and full time.

3. Estimation and results

Descriptive statistics of the sample are shown in Table 1. The average promotion incidence across the whole period was 8%. This does not include new entrants to NHS Scotland who are a consultant, but were not in the previous year. Three logit models were estimated and compared, a Pooled, a Random Effects (RE) and a Fixed Effects (FE). The first grouping variable was individuals, thus accounting for unobserved heterogeneity of individuals over time. However, estimation with the individual as the group showed that there was no panel variation in the data, i.e. little variation in individual behaviour over time. This reflects the fact that promotion to consultant can only occur once per individual and in one direction. This would favour the use of a pooled logit model.

As an alternative, we examine the FE, RE and pooled models with specialty as the grouping variable. This accounts for the unobserved heterogeneity at specialty level and that promotions are related to the demand for consultants and are therefore correlated within specialties.

The differences and similarities between the three models can be informative. The pooled model assumes that there is no panel-level variation in the data and can produce biased results because of this. The two panel models (FE and RE) explicitly account for any panel variation and both control for unobserved heterogeneity where this may be present. The choice of fixed or random effects can often be very difficult, especially when they produce different coefficients, as the assumptions on which they are based and their statistical properties are different. Table 2 shows the results from the fixed and the random effects estimations to be sufficiently close for this choice to not be a worry in the present analysis. This is a good indication for the soundness of the estimation as it indirectly supports the specification of the model and the RE assumption of independence between the random effect and the independent variables.

Noting that both RE and FE models control for unobserved heterogeneity, the main difference that remains between the two models in this situation is the interpretation of the inference that can be made. The FE assumption is tantamount to assuming that the sample estimated is the whole population that we are interested in. Hence, any

inference applies to the sample itself and should not be extended to a wider population. The upside is that the estimated specialty effects have a precise interpretation, namely, they can be directly attributed to the group they belong to. By contrast, the RE assumption is tantamount to assuming that the specialty effects are a random draw from a wider population. Hence, any inference can be extended to a wider population of which the estimated sample is representative. The downside is that the specialty effects have a far less clear interpretation. Which of the two assumptions is made will often depend on the problem at hand. In the case of this paper where the panel-level is the specialty of doctors in the NHS, the natural route is to assume the FE model. The use of RE and the Pooled estimates then is restricted to the judgement of the statistical properties of the model in general. The great similarity of the FE and RE estimates suggests a sound specification. Contrasting the results of the FE and RE with the pooled model suggests that controlling for unobserved heterogeneity has had a marked effect on the coefficients of the effort variables, although they remain insignificant. The coefficient for gender is noticeably larger in the panel models, as are the coefficients for experience from 13 years onwards.

The specialty specific fixed effects will control for unobserved heterogeneity due to specialty-specific differences in the characteristics of doctors, such as specialty-level differences in the average ability and skill of doctors. They will also reflect demand side factors in terms of consultant turnover and vacancy rates by specialty. The data at hand cannot distinguish whether these difference arise due to limited supply or to changes in the demand for individual specialties.

The FE model in Table 2 is used to explain the main results. Neither of the effort variables are statistically significant. Females appear to have a lower probability of promotion to consultant. The country of qualification suggests quite clearly that Indian degrees carry the lowest probability of promotion and Scottish ones the highest. This may reflect a preference of employers to promote ‘insiders’, those who graduated in Scotland and who therefore are known to potential employers. This may reflect the nature of the sample, as it was not possible to include promoted consultants who were previously working outside of Scotland.

The Year variable has been used to show any time trend in the promotions to consultant level in NHS Scotland in the 1990s. It suggests that the promotion probability increased during the 1990s. It should be noted that, in the multivariate context of the analysis, this trend is *over and above* any changes attributable to any of the other explanatory variables in the model as well as to any unobserved (but accounted for) differences between specialties.

Promotions appear to be positively influenced by the size of the origin health board. This result probably reflects the fact that health boards with higher numbers of staff offer better training and career development opportunities.

Experience is measured as the years after qualification. The impact of experience resembles that of a quadratic form, that is experience increases the probability of promotion up until 12 years experience, and its effects after that remain high slightly erratic. This suggests that the timing of promotion varies considerably, but the eventual probability of promotion does not. There are few cases where promotion happens in less than 11 years of medical experience, but the bulk of the promotions happen between 12 and 16 years. It is worth noting that this effect appears after any possible specialty differences have been accounted for. The fact that observed experience is the same for males and females and that the data does not allow the identification of individual career breaks which may be gender related, leaves the origin of the gender differential in promotions unaccounted for. Interaction terms between Experience and Female were tried but were found to have no statistical significance.

Both the number of managers per doctor and the number of nurses per doctor were associated with a higher chance of being promoted. This reflects the effect of skill mix in each health board. Boards where doctors have more support offer better promotion chances. The number of managers per doctor has the largest effect on promotion.

The total number of staff in the health board of employment in the year before promotion also has a large and significant association with promotion probability.

Larger health boards may offer more and varied training opportunities compared to smaller boards.

Finally, both of the main effort variables are not significantly associated with promotion. Their signs are in the predicted direction and the absolute effort variable (WTE_{t-1}) is close to significance at the 10% level.

4. Discussion

Our main finding is that effort in previous periods does not influence promotion chances in the current period. There are a number of potential explanations and caveats to this result. The first explanation is that the medical career structure is pre-determined and highly selective at the point of selection into, and during, medical school and selection into specialty. Those who wish to supply lower WTEs (who have strong preferences to work part time) will select themselves into General Practice.⁴ This means that any resulting differences in effort (WTEs) for those in the promotion pool do not matter (or are not large enough) to employers. This fits with the characterisation of the medical career structure where training periods are relatively fixed and where promotion is determined by network effects (higher chances of promotion for those who graduated in Scotland), experience and gender. Individual heterogeneity plays an insignificant role since there is little variation in effort and ability amongst doctors in the promotion pool.

Another explanation concerns the role of incentives in promotion hierarchies, in that incentives are weaker where the difference in rewards from the bottom of the hierarchy to the top is small. In this case, incentives might be dampened if there is evidence of within-job pay compression, which is common in the public sector. This leads to weaker incentives in promotions that may be associated with employer preferences for a more co-operative team-based work environment, such as the NHS.

A further explanation is concerned with the role of WTEs as a proxy for effort. The ISD data use contracted hours or sessions to construct the WTE variable (actual

⁴ Part time working in General Practice has increased substantially during the 1990's, characterised by new entrants working part time rather than GPs stepping down from full time to part time. However, in hospital medicine part time working has remained stable (Elliott et al., 2002).

contracted hours/whole time weekly hours) and so will underestimate total hours worked and WTEs. However, if the difference between contracted hours and actual hours is distributed randomly across doctors then this is less of an issue. The specialty specific fixed effect will also control for systematic differences in average hours due to specialty. In complex jobs, the difficulty of measuring effort or output is considerable. In the general literature on promotions, studies have measured effort using sickness absence (Audas et al., 2004). However, this also reflects a low health status. WTEs are a more direct measure of effort supplied that is more akin to the use of hours worked in standard labour supply models.

A further possible explanation is that it is not WTEs that matter, but human capital acquisition. The propensity of doctors to collect degrees, spend time working as clinical research fellows, and producing published output maybe more important than WTEs. Those doctors who spend more time on research maybe the driving factor influencing promotions. At the moment differences in research activity by specialty are captured in the fixed effect, although individual-level differences within specialties are still likely to exist.

Another key finding was that females are less likely to be promoted than male doctors. The source of this differential is not due to differences in experience caused by career breaks, as an interaction term between gender and experience was not significant. This potential source of gender discrimination is important for future workforce planning as over 50% of medical graduates are now women.

The main policy implications arise when considering changes to the medical career structure and to the contracts of doctors. Most of these changes, including changes to specialist medical training, the junior doctors contract introduced in 2000, and the new consultant contract, are all likely to influence the rate of promotions as relative rewards change. Recent changes to specialist medical training, including the Calman reforms, are making progression to consultant even more prescriptive. This will further limit the scope for individual heterogeneity having an impact on promotions, and introduces further selection into the training and promotion process. This maybe efficient if it is more costly to observe actual effort and productivity. The Department of Health has signalled an intention to reduce the length of specialist training (perhaps

influenced by meeting government targets for numbers of consultants). However, with a fixed promotion pool, the main effect of this will be to increase labour costs without an increase in the quality of consultants. The average quality of hospital consultants may even fall.

References

Audas R., Barmby T., Treble J. Luck, effort and reward in an organisational hierarchy. *Journal of Labour Economics* (forthcoming).

Baker G., Gibbs M., Holmstrom B. The internal economics of the firm: evidence from personnel data. *Quarterly Journal of Economics* 1994;59: 881-919.

Booth A., Francesconi M., Frank J. A sticky floors model of promotion, pay and gender. *European Economic Review* (forthcoming)

Burgess S, Metcalfe P (1999) *The use of incentive schemes in the public and private sectors: evidence from British establishments*. Centre for Market and Public Organisation, Working Paper 00/15, University of Bristol.

Doeringer P.B., Piore M.J. Internal labour markets and manpower analysis. Lexington: Heath Lexington Books, 1971.

Elliott R., Mavromaras K., Scott A., Bell D., Gerova V., van der Pol M. NHS Labour markets in Scotland. Final report to Scottish Executive Health Department. Health Economics Research Unit, University of Aberdeen, 2002.

Jovanovic B. Job matching and the theory of labour turnover. *Journal of Political Economy* 1979;57:972-990.

Gosden T, Forland F, Kristiansen IS, Sutton M, Leese B, Guiffrida A, Sergison M, Pedersen L (2001) Impact of payment method on the behaviour of primary care physicians: a systematic review. *Journal of Health Services Research and Policy*, **6**, 44-55.

Gravelle H., Guiffrida A. Inducing or restraining demand: the market for night visits in primary care. *Journal of Health Economics*, 2001;20: 755-779

Lazear E.P., Rosen S. Rank-order tournaments as optimum labour contracts. *Journal of Political Economy* 1981;89:841-864.

Lazear E.P. The future of personnel economics. *The Economic Journal* 2000;110:F611-F639.

Mooney G, Ryan M Agency in health care: getting beyond first principles. *Journal of Health Economics*, 1993; **12**, 125-238.

Prendergast C (1999) The provision of incentives in firms. *Journal of Economic Literature*, **37**, 7-63.

Robinson JC *Theory and practice in the design of physician payment incentives*. *Milbank Quarterly*, 2001;79; 149-177.

Rosen S. Prizes and incentives in elimination tournaments. *American Economic Review* 1986;76: 701-715.

STATA 8 manual Cross-Sectional Time-Series, 2003, Stata Press, Texas

Treble J., van Gameren E., Bridges S., Barmby T. The internal economics of the firm: further evidence from personnel data. *Labour Economics* 2001;8:531-552.

Wanless D. (2002) *Securing our future health: taking a long-term view*. Final Report. London: HM Treasury.

Wooldridge J M, 2002, Econometric Analysis of Cross Section and Panel Data, MIT Press, London.

Table 1. Descriptive statistics (n=9589)

Variable	Mean	Standard deviation	Min	Max
Female	0.421	0.494	0	1
<i>Country of qualification:</i>				
Qual-England	0.134	0.341	0	1
Qual-Scotland	0.648	0.478	0	1
Qual-India	0.084	0.278	0	1
Qual-RestUK	0.023	0.149	0	1
Year	7.911	2.180	4	11
<i>Experience:</i>				
9 years	0.101	0.301	0	1
10 years	0.098	0.297	0	1
11 years	0.085	0.279	0	1
12 years	0.060	0.236	0	1
13 years	0.046	0.209	0	1
14 years	0.031	0.174	0	1
15 years	0.026	0.158	0	1
16 years	0.021	0.142	0	1
<i>Staff ratios in destination health board:</i>				
PAM/Doctor ratio	0.895	0.194	0.69	2.07
Manager/Doctor ratio	2.352	0.471	1.8	5.86
Nurse/Doctor ratio	7.314	1.860	5.32	17.64
WTE _{t-1}	0.930	0.170	0.05	1.454
Change in WTE (WTE _{t-1} – WTE _{t-2})	-0.010	0.103	-0.9	0.875
Health Board size	2.727	0.498	1	3
Promotion	0.080	0.272	0	1

Table 2. Regression results

	Fixed effects logit		Random effects logit		Pooled logit	
	β	t	β	t	β	t
Female	-0.667	-6.68	-0.622	-6.39	-0.545	-4.93
<i>Country of Qualification:</i>						
England	0.524	2.90	0.532	2.98	0.531	2.42
Scotland	0.768	4.82	0.750	4.77	0.796	3.95
India	-0.480	-1.70	-0.513	-1.82	-0.572	-1.66
Rest of UK	0.682	2.39	0.658	2.32	0.642	2.13
Year	0.178	5.53	0.175	5.49	0.164	6.45
<i>Experience</i>						
9 years	1.268	6.57	1.269	6.58	1.273	4.65
10 years	2.053	12.28	2.067	12.40	2.075	7.32
11 years	2.801	17.70	2.800	17.75	2.785	9.81
12 years	3.267	19.92	3.279	20.07	3.217	15.21
13 years	3.208	18.15	3.170	18.06	3.087	12.82
14 years	3.323	16.74	3.288	16.64	3.189	12.75
15 years	3.126	14.39	3.113	14.45	3.011	14.10
16 years	3.210	13.86	3.202	13.89	3.118	13.19
<i>Staff ratios in destination health board:</i>						
PAM/Doctor ratio	-0.252	-0.46	-0.234	-0.44	-0.200	-0.45
Manager/Doctor ratio	1.506	4.38	1.508	4.42	1.535	3.65
Nurse/Doctor ratio	0.133	1.58	0.131	1.58	0.118	1.54
Health Board size	2.360	13.76	2.350	13.85	2.307	8.23
WTE _{t-1}	0.036	0.13	0.065	0.24	0.186	0.54
(WTE _{t-1} – WTE _{t-2})	0.190	0.45	0.201	0.48	0.124	0.43
Constant			-17.311	-21.11	-17.091	-12.81
Restricted LogL	-2490		-2759		-2759	
Unrestricted LogL	-1849		-1999		-2021	
Model chi2 (df)	1480 (20)		878 (20)		1346 (20)	
n	9779		9589		9589	
Number of specialties	55		62		-	
Average specialty size	178		159		-	
Rho (p-value)	-		0.045 (p<0.0001)		-	
Pseudo R2	0.59		0.32		0.28	

Note: Random and Fixed effects LOGIT: group variable is specialty. Source of data ISD Scotland. Estimation carried out using STATA xtlogit and logit commands.