



Comparative Performance Evaluation: Quality.

DOI:

[10.1016/B978-0-12-375678-7.01313-4](https://doi.org/10.1016/B978-0-12-375678-7.01313-4)

[Link to publication record in Manchester Research Explorer](#)

Citation for published version (APA):

Fichera, E., Nikolova, S., & Sutton, M. (2014). Comparative Performance Evaluation: Quality. In A. J. Culyer (Ed.), *Encyclopedia of Health Economics* (Vol. 1, pp. 111-116). Elsevier BV. <https://doi.org/10.1016/B978-0-12-375678-7.01313-4>

Published in:

Encyclopedia of Health Economics

Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

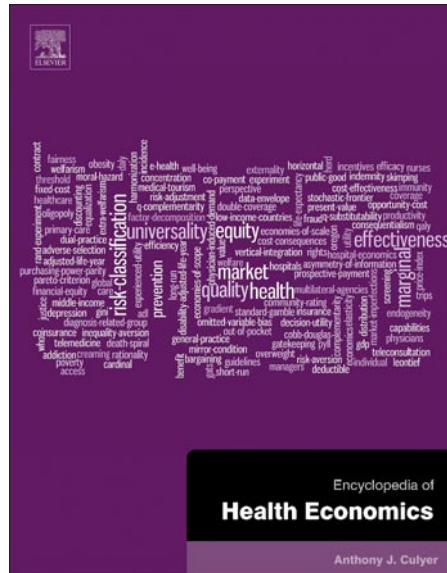
Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact openresearch@manchester.ac.uk providing relevant details, so we can investigate your claim.



**Provided for non-commercial research and educational use only.
Not for reproduction, distribution or commercial use.**

This article was originally published in *Encyclopedia of Health Economics* published by Elsevier, and the copy attached is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who you know, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<http://www.elsevier.com/locate/permissionusematerial>

Fichera E., Nikolova S., and Sutton M. Comparative Performance Evaluation: Quality. In: Anthony J. Culyer (ed.), *Encyclopedia of Health Economics*, Vol 1. San Diego: Elsevier; 2014. pp. 111-116.

© 2014 Elsevier Inc. All rights reserved.

Comparative Performance Evaluation: Quality

E Fichera, S Nikolova, and M Sutton, University of Manchester, Manchester, UK

© 2014 Elsevier Inc. All rights reserved.

Glossary

Agency relationship The relationship between an agent and a principal. Classically in health care, the role of a physician or other health professional in determining the patient's (or other client's) best interest and acting in a fashion consistent with it. The patient or client is the principal and the professional is the agent. More generally, the agent is anyone acting on behalf of a principal, usually because of asymmetry of information. In health care, other examples include health managers acting as agents for their principals such as owners of firms or ministers, regulators as agents for politically accountable ministers, ministers as agents for the electorate. In health care, the situation can become even more complicated by virtue of the facts, first, that the professional thereby has an important role in determining the demand for a service as well as its supply and, second, that doctors are expected (in many systems) to act not only for the 'patient' but also for 'society' in the form, say, of other patients or of an organization with wider societal

responsibilities (like a managed health care organization), or taxpayers, or all potential patients. There can be much ambiguity, as in seeking to understand the agency relationships in overseas aid giving and management, and as in establishing the extent to which formal contracts can enhance efficiency.

Incentive contracts The contracts between insurers or other third party payers and the providers of health care that embody incentives and penalties (both usually financial) for failing to meet particular conditions.

Yardstick competition An industrial regulatory procedure under which the regulated price is set at the average of the estimated marginal costs of the firms in the industry.

Zeckhauser's dilemma A problem with incentive contracts when those who are incentivized to behave in particular ways cannot fully control the consequences of their actions. They then require compensation in some form to offset this increase in the risk they face of failure, which raises the cost of the contract relative to the benefits anticipated by the principal.

Introduction

Health care purchasers and regulators often make comparisons between providers on indicators of quality. In this article the rationale for such comparisons is described, the options for this form of monitoring are considered and how this type of evaluation has evolved over time is outlined. Then, using a recent example of a quality program that links financial rewards to comparative performance in the UK, the key issues with this kind of performance evaluation are highlighted.

Principal-Agent Problems

The health care sector is characterized by a series of agency relationships. Patients delegate decision-making to doctors and payers give responsibility for supplying health care to providers. This delegation of decision-making or provision would be unproblematic if there was symmetric information and identical objectives were shared between the parties. In reality, two general problems are suggested by the principal-agent analysis. First, the task itself (i.e., delivering health care) is only partially observable or verifiable. This is called the moral hazard or hidden action problem. Second, the agent's capabilities are unknown to the principal but are known to the agent before the parties enter into the contract. This may adversely affect the principal's payoff and is called the adverse selection or hidden information problem.

The solution adopted in practice is to use a set of performance indicators to measure the output of the agent.

However, this is only a partial solution because the information problems persist when the correlation between such indicators and the agent's effort is noisy and determined by a random component that often varies across agents. The extent to which the agent is in control of such variation is also unknown to the principal. The principal must therefore design a contract or system of incentives that elicits a second-best outcome from the agent.

Problems with Incentive Contracts in Healthcare

It is often claimed that the design of incentive contracts is more difficult in the health care sector than in other sectors. This is particularly the case when the principal's problem of ensuring that the agent delivers a high quality service is considered. There are five problems that are germane:

1. One of the best known concerns about incentive contracts is the trade-off between incentives and risk (so-called Zeckhauser's dilemma). Theoretically, incentive contracts impose a risk on agents and risk-averse agents will require a higher mean level of compensation. This premium will increase with the riskiness of the environment. Although empirical research has not found convincing evidence that higher incentives are given in riskier environments, health care providers provide an uncertain output (the well-being of the patient) which is only partially dependent on their actions.
2. When multiple actions are substitutes, incentive schemes may cause diversion of effort. For instance, an incentive

- scheme focused on observable indicators will induce health care providers to game the system and reduce their effort on unobservable dimensions.
3. Because patients differ in their expected health outcomes and the agent has more information on the expected health outcomes than the principal, the agent may engage in 'cherry-picking' of patients, providing care only to patients at low risk of adverse outcomes.
 4. As health care provision often requires input from more than one agent, the aggregation of agents into groups (for example, hospital teams) for incentive contracts creates externalities. These externalities may be positive (through monitoring of effort by close peers) or negative (caused by free-riding on others' efforts).
 5. Health care providers have a social role and are trained to adopt professional ethics. Their utility functions are typically assumed to contain an altruism component that values the benefits to their patients. Incentives have the danger of crowding-out this intrinsic motivation.

Information on Quality

Quality assessment in health care is bedeviled with measurement problems. The measurement of output, or more strictly the agent's effort in producing output, is particularly difficult. Quality can be measured in terms of the quality of inputs, processes, or outcomes. Input quality measurement, for example, would involve assessing the capabilities and training of the labor force, the standard of the capital facilities and equipment, and the input mix. Such an approach is often taken by health care regulators seeking to maintain a register of qualified providers. Process quality measurement, however, would involve assessing whether agents are performing actions that are most likely to generate good quality outputs. In health care, this might involve assessing whether providers are adhering to best-practice guidelines and offering patients effective treatment regimes. Finally, quality output measurement would focus on the benefits that have been achieved for patients, regardless of how they have been achieved. Such benefits should include gains in survival and quality of life and increasingly capture patients' experience of using health care services.

The difficulty for the principal is to know which type of quality measurement offers the most accurate information on the agents' efforts. Quality inputs are a necessary but not sufficient condition for quality outputs. When assessing the quality of processes, principals are frequently forced to rely on agents' reports of their processes. These may be deliberately misreported, or may be applied to the least-costly patients who may be less likely to gain substantial benefits. The main problem with direct measurement of the quality of the agent's outputs is that these are noisy signals of their effort because patient outcomes reflect historical events, the patient's own actions, and the actions of other agencies. These are largely unobservable and contain a substantial random element.

For these reasons, principals often adopt a portfolio of quality indicators across each of these levels. This reduces, but

does not eliminate, the problems with each of the individual indicators. However, it generates new problems of how the agent's performance on each indicator should be aggregated to form an overall signal of their effort.

Comparative Performance Evaluation

Broadly speaking, incentive contracts can be classified into two types of performance measurements: (1) absolute and (2) comparative performance. Under absolute performance, the agent is set standards on performance measures that must be achieved, for example, 80% compliance with a care guideline. Under a comparative performance scheme, the agent's performance is benchmarked against a relative standard.

The relative standard in comparative performance evaluation can be set on two dimensions: time and reference group. The time dimension of comparative performance can be implemented in a static or in a dynamic setting (i.e., current or historical performance). The reference group dimension of comparative performance can be implemented across groups of agents within or between health care organizations. Although dynamic comparative performance may or may not be implemented across reference groups, static comparative performance is always relative to a reference group.

To set an absolute performance standard, the principal needs to have good information on the effort that the agent will need to make to reach that standard. Setting a relative standard based on the agent's own historical performance ensures that the agent improves quality (and thereby increases effort) period-on-period but can fall foul of secular trends and does not seek to induce equal effort across agents. Use of a static reference group benchmark isolates performance measurement from (common) secular trends, but relies on choice of an appropriate reference group and places the agent at higher risk.

If the reference group approach is selected, comparative evaluation can involve two broad types of comparisons against the other agents. It can involve comparison to the average (which is called benchmarking) or it can involve the construction of league tables (known as a rank-order tournament in the sport sector).

Benchmarking versus Rank-Order Tournaments

The primary purpose of relative performance evaluation is to mitigate the principal's imperfect information. However, comparative performance evaluation has a 'yardstick competition' effect as well as an information effect. Because rank-order tournaments will increase competition more than benchmarking, the latter is a lower-powered incentive whereas the former provides sharper incentives. Previous research has shown that wider variation in levels of performance will be induced by rank-order tournaments. The risk of such tournament-based incentives is that contestants who think they have little chance to earn a prize are not motivated by the scheme and wider variations in performance are created.

Comparative performance evaluation is optimal only when all agents face common challenges. When this is the case, the performance of one agent allows the principal to infer information about another agent's performance. However, if worse health conditions adversely affect performance and these are concentrated in specific areas, then these factors should be filtered out by comparing providers within the same area. However, an agent's rank-order within an area contains less information on the performance of an individual agent and will not generally represent an efficient use of information. Instead, aggregate measures like averages of similar organizations are more efficient because they provide sufficient information about common challenges.

Benchmarking is able to reduce the 'feedback' and the 'ratchet' effects of the reward mechanism. Feedback occurs whenever one agent's action affects the incentive scheme and thus changes the agent's own reward as well as the reward for other agents. As the number of agents affecting the overall standard is higher under benchmarking, the feedback effect will be lower than in the case of rank-order tournaments. The ratchet effect is in essence the dynamic counterpart of the feedback effect. Good agents may be better off by hiding or misreporting their 'true' performance for fear that the principal may raise the current target on the basis of past performance. Unless collusion between agents occurs, this gaming is mitigated by benchmarking.

More fundamentally, any judgment on which type of relative performance evaluation is most effective depends on the goals the principal is trying to achieve. The principal may be primarily concerned with maximizing efficiency or with minimizing inequity. If the principal is mainly concerned with increasing the efficiency of health care provision then they will seek to use comparative performance evaluation to increase the average level of performance and will likely adopt a rank-order tournament. Alternatively, the principal may be motivated by the distribution of agents' performance levels as they care most about equity of health care provision. In this case, they will seek to use comparative performance evaluation to close the gap between outstanding and poorly performing health care providers. In this case, the principal may be reluctant to use rank-order tournaments as this may increase the gap in performance between agents at the top and the bottom of the league.

The Development of Comparative Performance Evaluation

Comparative performance evaluation began as an informal exercise in the private sector and became more structured in the late 1970s in response to Japanese competition in the copier market. It typically took the form of rank-order tournaments as the extent of market competition was high.

More recently, benchmarking has been used in the public sector. For example, from April 1996 the Cabinet Office and HM Revenue and Customs in the UK have run a project, the Public Sector Benchmarking Service, to promote benchmarking and the exchange of good practice in the public sector.

Box 1 shows some key definitions of benchmarking. It highlights the competitive definition of benchmarking by the private company Xerox and the less competitive definition of benchmarking, focused on learning from comparisons, by the public sector.

These developments have been mirrored in the health care sector. Initially, governments in their roles as payers and regulators, made use of the availability of electronic information to give feedback to providers on their relative performance. These initiatives were frequently undertaken under the auspices of professional bodies and the focus was deliberately on information-sharing and supporting intrinsic motivation. Providers were often given data on their own performance and the performance of the average provider or their rank in the distribution of performance over anonymized providers.

Later, these data were deanonymized and sometimes publicly reported. This was viewed as a natural progression. Once providers were content that the information on their performance was accurately recorded and consistently collected across providers, the public could be reassured that quality in the public health care sector was consistently high.

However, when quality first became linked to penalties and rewards, it was typical to use absolute performance standards. The introduction of waiting time targets in the UK National Health Service (NHS), associated with stringent monitoring and strong personal penalties, for example, was enforced using absolute maximum standards. These were frequently criticized for distorting priorities and inducing gaming, though the empirical evidence on patient reprioritization is scant and previous research finds no support for gaming. Similarly, the introduction of highly powered financial incentives for UK general practices in the form of the Quality and Outcomes Framework were based on absolute standards. The lack of data on baseline performance meant that these standards were set too low and that only modest gains in quality were delivered, some of which have been shown to be due to gaming of the self-reported performance information.

The second generation of financial incentives for improving quality in the UK NHS make greater use of comparative performance evaluation. There are a number of national schemes that emphasize local flexibility and payment for quality improvement rather than achievement of absolute standards. The forerunner to these was introduced in one region in England

Box 1 Definitions of benchmarking

The Public Sector Benchmarking Service defines benchmarking as: 'Improving ourselves by learning from others.'

The Cabinet Office calls benchmarking: 'The process of comparing practices and performance levels between organizations (or divisions) to gain new insights and to identify opportunities for making continuous improvements.'

The European Benchmarking Code of Conduct states that: 'Benchmarking is simply about making comparisons with other organizations and then learning the lessons that those comparisons throw up.'

Xerox, a pioneer of private sector benchmarking in the copier market says that it is: 'The continuous process of measuring products, services and practices against the toughest competitors or those companies recognized as industry leaders.'

and provides a good example of the limitations of using financial incentives linked to comparative performance evaluation. This scheme is described in the next section.

The Advancing Quality Program

The Advancing Quality (AQ) program was launched in October 2008 for 24 acute hospital trusts in the North West of England. Trust performance is summarized by an aggregate measure of quality – the composite quality score – within each of five clinical domains. The five incentivized clinical conditions are acute myocardial infarction, coronary artery bypass graft surgery, hip and knee replacements, heart failure, and pneumonia. The composite quality scores are derived by equally weighing achievement on a range of quality metrics which include process and outcome measures. **Table 1** lists the quality metrics used in AQ.

Table 1 Quality measures used in the advancing quality program

<i>Patients with acute myocardial infarction</i>
Aspirin at arrival
Aspirin prescribed at discharge
ACEI ^a or ARB ^b for LVSD ^c
Adult smoking cessation advice/counseling
Beta blocker prescribed at discharge
Beta blocker at arrival
Fibrinolytic therapy received within 30 min of hospital arrival
Primary PCI ^d received within 90 min of hospital arrival
Standardized survival index
<i>Patients with heart failure</i>
Evaluation of left ventricular function
ACEI or ARB for LVSD
Discharge instructions
Adult smoking cessation advice/counseling
<i>Patients receiving coronary artery bypass grafting</i>
Aspirin prescribed discharge
Prophylactic antibiotic received within 1 h before surgical incision
Prophylactic antibiotic selection for surgical patients
Prophylactic antibiotics discontinued within 48 h after surgery end time
<i>Patients receiving hip and knee replacements</i>
Prophylactic antibiotic received within 1 h before surgical incision
Prophylactic antibiotic selection for surgical patients
Prophylactic antibiotics discontinued within 48 h after surgery end time
Recommended venous thromboembolism prophylaxis ordered
Received appropriate venous thromboembolism prophylaxis within 24 h of surgery
Readmission avoidance rate – 28 days post discharge
<i>Patients with pneumonia</i>
Oxygenation assessment
Initial antibiotic selection for immunocompetent patients
Blood culture performed in A&E before initial antibiotics received in hospital
Initial antibiotic received within 6 h of hospital arrival
Adult smoking cessation advice/counseling

^aAngiotension converting enzyme inhibitor.

^bAngitensin receptor blocker.

^cLeft ventricular systolic dysfunction.

^dPercutaneous coronary intervention.

The AQ scheme is similar to the Hospital Quality Incentive Demonstration (HQID) in the US. Both schemes started as pure rank-order tournament systems. At the end of the first year, hospitals in the top quartile received a bonus payment equal to 4% of the revenue they received under the national tariff for the associated activity. For trusts in the second quartile, the bonus was 2% of the revenue. For the next two quarters, the reward system changed to the same structure that was adopted by HQID after 4 years; bonuses were earned by all hospitals performing above the median score from the previous year and hospitals could earn additional bonuses for improving their performance or achieving top or second quartile performance. There was no threat of penalties for the poorest performers at any stage.

Evidence from HQID and AQ initiatives suggests that providers quickly converge to similar values on the process metrics and differences in performance must be measured at a very high level of precision to discriminate among providers. In addition, on some of the process measures most providers scored (close to) maximum scores. Because of the small variability in the measures and these ceiling effects, the schemes end up rewarding trusts based on small differences in performance.

Under the HQID and AQ scoring mechanisms, all of the targeted indicators are given equal weight regardless of their underlying difficulty. Thus, the quality score methodology involves a risk that providers will divert effort away from more difficult tasks toward easier tasks. However, despite the clear incentive to do so, research from the US suggests no consistent evidence that providers engaged in such behavior.

From the perspective of public health and policy making, the more important question, however, is whether health outcomes have changed as a result of the introduction of HQID and AQ initiatives. Here, the US and UK experiences are contradictory. A comprehensive US study found no evidence that HQID had affected patient mortality or costs. The first evidence from the UK shows that the introduction of AQ initiative was associated with a clinically significant reduction in mortality.

In both countries, studies have found weak links between process measures and patient mortality and ruled out causal effects on the health outcome. This appears to show that improved performance on the process measures alone could not explain the association with reduced mortality in the North West.

The critical questions now are how and why AQ scheme was associated with robustly estimated mortality reductions when similar studies have found little evidence of an effect of process metrics on patient outcome.

The qualitative evaluation of the AQ scheme found that participating hospitals adopted a range of quality improvement strategies in response to the program. These included employing specialist nurses and developing new and/or improved data collection systems linked to regular feedback of performance to participating clinical teams.

Compared to HQID, the larger size and greater probability of earning bonuses in AQ may explain why hospitals made such substantial investments. The largest bonuses were 4% in AQ compared to 2% in HQID and the proportion of hospitals

earning the highest bonuses was 25% in AQ compared to 10% in HQID.

In addition, the participation process may be important. To participate in HQID, hospitals had to (1) be subscribers to Premier's quality-benchmarking database, (2) agree to participate, and (3) not withdraw from the scheme within 30 days of the results being announced. The 255 hospitals that participated represented just 5% of the total 4691 acute care hospitals across the US. In contrast, the English scheme was a regional initiative with participation of all NHS hospitals in the region. This eliminated the possibility of participation by a self-selected group that might already consist of high performers or be more motivated to improve. Further experiments would be required to identify whether pay for performance schemes are more effective when participation is mandatory or targeted at poor performers.

Despite the 'tournament' style of the program, staff from all AQ participating hospitals met face-to-face at regular intervals to share problems and learning, particularly in relation to pneumonia and heart failure, where compliance with clinical pathways presented particular challenges and where the largest mortality rate reduction can be found. Similar shared learning events were run as 'webinars' for HQID. The face to face communication, regional focus, and smaller size of the scheme in England may have made interaction at these events more productive.

The fact that a scheme that appeared similar to a US initiative was associated with different results in England reinforces the message from the rest of the literature that details of the implementation of incentive schemes and the context in which they are introduced have an important bearing on their effects.

Concluding Remarks

To summarize, the asymmetry of information between the principal and the agent is particularly acute in the case of information on quality. Principals design incentive contracts under these circumstances to induce agents to increase their effort. One way in which principals can retrieve information on the efforts being made by agents is through comparisons of performance across time and/or across agents. Such comparative performance evaluation can involve comparison to own historical achievements or a reference group's achievements. The principal can benchmark agents to the average or create a rank-order tournament.

Although both types of comparative performance evaluation can improve efficiency by reducing the principal's information problem, rank-order tournaments are more likely to increase the gap between performance at the top and the bottom of the league. Benchmarking minimizes feedback and ratchet effects, but it can also weaken competition between agents. Ultimately, the choice between benchmarking and rank-order tournaments depends on the objectives of the principal.

In practice, comparative performance evaluation for improving quality was used quite widely and with little controversy when it appealed only to intrinsic motivation. Linkage

of comparative performance evaluation to financial rewards, however, has led to a sharper focus on its limitations. In this regard, the experiences with the HQID and AQ initiatives display many of the conundrums of using comparative performance evaluation. There is a great deal of uncertainty over, and little empirical evidence to support, the choice of comparator. The frequently adopted strategy of using a portfolio of indicators leads to problems of appropriately weighing the calculation of overall performance to avoid re-prioritization of effort. Finally, incentivization of improvements in the quality of processes reported by agents does not in itself lead to outcome improvements.

Overall, the evidence base on the effects of comparative performance evaluation is weak. Although there has been a great deal of (well-intentioned) experimentation, these initiatives have been adapted too frequently and have not been rigorously evaluated. Ultimately, the main challenges for principals considering the use of comparative performance evaluation are how to measure hospital quality, how to identify similar agents to make accurate comparisons, whether to appeal to extrinsic or intrinsic motivation, and how to devise and implement the pay-for-performance initiative given the context in which it is introduced.

See also: Competition on the Hospital Sector. Heterogeneity of Hospitals. Markets in Health Care. Pay-for-Performance Incentives in Low- and Middle-Income Country Health Programs

Further Reading

- Benabou, R. and Tirole, J. (2006). Incentives and prosocial behaviour. *American Economic Review* **96**(5), 1652–1678.
- Burgess, S. and Metcalfe, P. (1999). Incentives in organisations: A selective overview of the literature with application to the public sector. *CMPO Working Paper Series No.00/16*.
- Burgess, S. and Ratto, M. (2003). The role of incentives in the public sector: Issues and evidence. *Oxford Review of Economic Policy* **19**(2), 285–300.
- Chalkley, M. (2006). Contracts, information and incentives in health care. In Jones A. M. (ed.) *The Elgar companion to health economics*, pp. 242–249. Cheltenham: Edward Elgar Publishing Ltd.
- Dixit, A. (2002). Incentives and organizations in the public sector: An interpretative review. *Journal of Human Resources* **37**(4), 696–727.
- Frey, B. S. (1997). A constitution for knaves crowds out civic virtues. *Economic Journal* **107**, 1043–1053.
- Glaeser, E. L. and Shleifer, A. (2001). Not-for-profit entrepreneurs. *Journal of Public Economics* **81**(1), 99–115.
- Grout, P. A., Jenkins, A. and Propper, C. (2000). *Benchmarking and incentives in the NHS*. London: Office of Health Economics, BSC Print Ltd.
- Holmström, B. and Milgrom, P. (1991). Multitask principal-agent analysis: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization* **7**, 24–52. (Special Issue).
- Lazear, E. P. and Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy* **89**(5), 841–864.
- Lindenaauer, P. K., Remus, D., Roman, S., et al. (2007). Public reporting and pay for performance in hospital quality improvement. *New England Journal of Medicine* **356**(5), 486–496.
- Nicholas, L., Dimmick, J. and Iwashyna, T. (2010). Do hospitals alter patient care effort allocations under pay-for-performance. *Health Services Research* **45**(5 Pt 2), 1559–1569.
- Prendergast, C. (2002). The tenuous trade-off between risk and incentives. *Journal of Political Economy* **110**(5), 1071–1102.

- Propper, C. (1995). Agency and incentives in the NHS internal market. *Social Science and Medicine* **40**, 1683–1690.
- Ryan, A. M. (2009). Effects of the premier hospital quality incentive demonstration on medicare patient mortality and cost. *Health Services Research* **44**(3), 821–842.
- Ryan, A. M., Tomkins, C., Burgess, J. and Wallack, S. (2009). The relationship between performance on Medicare's process quality measures and mortality: Evidence of correlation, not causation. *Inquiry* **46**(3), 274–290.
- Shleifer, A. (1985). A theory of yardstick competition. *RAND Journal of Economics* **16**(3), 319–327.
- Siciliani, L. (2009). Paying for performance and motivation crowding out. *Economics Letters* **103**(2), 68–71.