Performance differences between for-profit and not-for-profit health providers: Relevance for the NHS

Bernarda Zamora[1], Will Bartlett[1], Virginie Perotin[2], Pauline Allen[3]

PRELIMINARY DRAFT – 9 December 2008

1. Introduction

At the present stage of the NHS reform, the main focus is on quality (*High Quality for All: NHS Next Stage Review Final Report*). Among the chosen outcome indicators of quality in acute care, the 2008 report mentions waiting times, better access to diagnostics in convenient locations, rates of premature mortality, rates of adverse events such as infections, patient experiences and degree of satisfaction, and degree of innovation. The main instruments which are expected to operate as a booster of quality are, firstly, a greater patient choice among different providers based upon better access to information on quality outcomes, secondly, a system of payments to hospitals linked to volume and quality, and, thirdly, a system of professionally led independent regulation.

Our aim in this paper is to analyze the expected impact of the introduction of competition among government and independent providers (for-profit and not-for-profit) on health outcomes (quantity and quality) under the new system of payments: Payment by Results (PbR). To this end, we rely first on predictions of health economics, based both on existing models and on our particularization of some of

[1] School for Policy Studies, University of Bristol

[2] Leeds University Business School

[3] London School of Hygiene and Tropical Medicine

Correspondence address: Bernarda Zamora, School for Policy Studies, University of Bristol, 8 Priory Road, Bristol BS8 1TZ
Email: Bernarda.Zamora@bristol.ac.uk
Tel: (0)117 9546604

them to monopolistic competition under PbR. Secondly, we analyze some evidence on past experiences, mainly the internal market during the 1990s, that share some characteristics with the ongoing reform. Lastly, we summarize the state of the ongoing reform regarding the level of competition, the implementation of PbR, and the entry of the independent sector.

In regard to the theoretical model, it is broadly accepted that health markets can be considered as monopolistically competitive, including an element of monopoly (downward-sloping demand depending on prices or responsive to any other signal reflecting quality) and an element of competition (large number of competitors – each firm ignoring strategic interactions) among which the patient has some possibility of choice. McGuire (2000) mentions numerous studies that model health markets as being monopolistically competitive. In the UK, there is evidence of a demand responsive to quality and costs as measured by time in the waiting list and distance to hospital (Gravelle et al., 2003, Windmeijer et al., 2005). Also, the ongoing reform of the NHS promotes patient's choice and entry of private sector providers as a main element of the reform to enhance competition. The effect of competition on quality and quantity with regulated prices has been analyzed under different models (see Gaynor, 2006, for a thoughtful summary).

We start from McGuire (2000) model. We particularize it to the case of the PbR reimbursement system. This payment system is a supply-side costs sharing mechanism with fully supply-side cost sharing. The provider bears all the costs once the patient has been admitted. As in McGuire (2000), our baseline model assumes profit-maximizing provider behaviour. We derive from it the optimal PbR tariff necessary to achieve a socially optimal quantity of treatment for each patient. We generalize the baseline model to consider any level of provider altruism so as to model the not-for-profit provider behaviour. This model generalizes Ellis and McGuire (1986) model to the introduction of quality and competition. The resulting levels of quantity and quality allow us to analyze the effect of altruism and competition on them as well as to compare them with the socially optimal levels. Later, we model not-for-profit behaviour by means of a binding ethics constraint on the baseline model. Finally, we briefly review some existing models with alternative hypothesis of not-for-profit behaviour.

The experience of competition in the NHS is limited and there is not much empirical evidence on its effects outside the period of the internal market in the 1990s. We summarize some results that discuss the level of competition and its effects during the internal market. Furthermore, we discuss the potential effects of the current reform on competition.

The rest of the paper is organized as follows. Section 2 presents the theoretical framework. Section 3 presents empirical evidence on competition in the NHS and discuss the implementation of PbR and entry of independent providers. Finally, we present some concluding remarks.

2. Theoretical Framework

2.1. Baseline Model: Profit Maximization

Following McGuire (2000), we present a model of monopolistic competition where the provider maximizes profits by setting quantity and quality. These arguments of the profit function must be understood as inputs of the health production function, in which quantity is contractible but quality is not. For example, the number of visits may be contractible but the time of the consultation may be not. We particularize McGuire's model to the case in which the insurer or government reimburses the provider according to a fully-capitated system, also known as pure prospective payment which corresponds with the PbR system in England. Moreover, health services are free at the point of use that is, the patient's co-payment is zero.

Within this framework, the provider's objective is to maximize profits, π, that depend on the number of patients treated (market share), n, on the quantity of treatment given to each patient, x, on the prospective payment or fixed reimbursement per patient, R, and on the costs of treatment, $c(e)$, that are related to the level of quality, e, (effort). Costs are increasing and convex ($c_e > 0, c_{ee} > 0$). The provider is an "agent" that takes into account the expected patient's reaction to his supply of quality and quantity.

Following McGuire and others, we consider that the health production function or benefit function, $B(x,e)$, represents the patient's demand. Demand for x and e is independent of patients' income[4]. We assume decreasing marginal returns such that $B_x > 0, B_e > 0, B_{xx} < 0, B_{ee} < 0.$ Since the price of health is zero for patients, they demand x and e up to the point in which $B_x = 0$ and $B_e = 0$. This implies excessive demand for quantity and quality (moral hazard). We can see below that the mechanism of supply-side costs sharing in the prospective reimbursement system restraints moral hazard since the patient can be enforced to accept less quantity than desired.

To introduce competition, we consider that the provider may attract more patients by offering a higher level of health so that we can express the number of patients visiting a provider by the increasing function $n = n(B)$.

In these terms, the profit maximization problem for the provider is:

---

[4] The patient's utility function can be expressed as $U(y + B(x,e))$, where y is income.

(1) $\underset{x,e}{Max}\,\pi = n(B)[R - c(e)x]$

whose first order conditions are:

(2) $\pi_x = n' B_x [R - cx] - nc = 0$

(3) $\pi_e = n' B_e [R - cx] - nc_e x = 0$

From the relationship between both conditions we obtain:

(4) $e = \frac{c}{c_e} \frac{\varepsilon_{B,e}}{\varepsilon_{B,x}}$

where $\varepsilon_{B,e} = B_e \frac{e}{B}$ and $\varepsilon_{B,x} = B_x \frac{x}{B}$ are the elasticities of the health production function with respect to quality and quantity, respectively.

This has a first immediate implication on the allocation of quality. If the health production follows a constant elasticity of substitution function (CES), the prospective payment induces an allocation of quality depending only on technology and independent of the reimbursement and of the allocation of quantity[5].

In general, with variable elasticity, condition (4) is similar to the Dorfman-Steiner (1954) condition but, in our case, prices do not enter[6]. The level of quality goes up if the quality elasticity of demand increases or the quantity elasticity of demand decreases.

To analyze the allocation of quantity and quality as well as the effects of competition, we can reorder the first order conditions as follows:

(5) $\pi_x : B_x = \frac{c}{\frac{\varepsilon_{B,n}}{B}(R - cx)}$

(6) $\pi_e : \frac{B_e}{c_e} = \frac{x}{\frac{\varepsilon_{B,n}}{B}(R - cx)}$

---

[5] If the system introduces patient co-payment, a CES production function does not imply constant quality. In this case the allocation of quality depends on the allocation of quantity.

[6] Gaynor (2006) finds the Dorfman-Steiner condition for a monopolist that sets prices and quality.

where $\varepsilon_{n,B} = n' \frac{B}{n}$.

Before analyzing the provider optimal decision on quantity and quality, we review the meaning of competition in this framework. Under perfect competition, the provider can steal the whole market by increasing his health production beyond their competitors, that is, $n' = \infty$. Therefore, in perfect competition, all the providers offer the same level of health.

In contrast, if the provider has some market power because providers are not perfect substitutes, their competitors cannot steal his market share by offering more health. In the extreme case of a monopolist, his market share - one - will be constant for any supply of health $(n' = 0)$ and the relationship between number of patients and level of health breaks. Under monopolistic competition, $n' > 0$ (finite), and increasing with the level of competition.

From condition (5), it is clear that the provider allocates quantity below patient's demand, such that $B_x > 0$, whenever there is not perfect competition $(\varepsilon_{n,B} < \infty)$ and the reimbursement is above average costs. In this sense, the provider has incentives to restraint moral hazard. Moreover, competition increases the optimal level of quantity.

Condition (6) determines the level of quality for the chosen value of quantity. Since $B_{ee} < 0$ and $c_{ee} > 0$, $\frac{B_e}{c_e}$ is a decreasing function in . Therefore, an increase in the average profitability leads to an increase in quality. Since quality is costly, an increase in profitability can only be achieved through an increase in the reimbursement fee R. Competition decreases the level of profitability because increases quantity of treatment. Therefore, quality has to increase to maintain the level of profitability after an increase in competition.

2.2. Optimal reimbursement

The socially optimal level is that in which marginal benefits for the patient equal marginal costs for the provider[7]. Let's x^* be the socially optimal level in which $B_x = c$. Based upon the above baseline model, we can determine the optimal reimbursement according to equation (6) as:

[7] McGuire (2000)

(8) $R^* = \left(1 + 1/A\varepsilon_{n,B}\right)cx^*$

where $A = \dfrac{B_x}{B/x} = \dfrac{c}{B/x}$.

The optimal reimbursement will be higher the higher the marginal costs of treatment[8] and the lower the patient's possibility of choice among providers. Therefore, more competition allows achieving the optimal quantity with a lower reimbursement.

We envisage three problems regarding the implementation of a socially optimal level of treatment. First, the heterogeneity of the provider's costs. Second, the estimation of the elasticity of the health production function with respect to quantity of treatment. And third, the estimation of the market elasticity with respect to a level of health.

The heterogeneity of costs implies that the optimal reimbursement should be higher for providers with higher marginal costs. If the regulator considers the average of marginal costs across providers to estimate the optimal level of reimbursement and reimburses all providers with the same capitation fee, only the average provider will offer the optimal quantity; those providers with higher (lower) marginal costs will underprovide (overprovide).

Regarding the estimation of the elasticity of the health production function with respect to quantity at the point in which $B_x = c$, efficiency models and stochastic frontier analysis define health outcomes and health activity to estimate production functions and/or costs functions. The importance of the parameter A is relative; it is used as a measure to compare efficiency among providers, with a higher level of A for more efficient providers.

The estimation of $\varepsilon_{n,B}$ requires to estimate a demand function for health in terms of a measure of health or quality. For the UK Gravelle et al., 2003 estimate a demand function for health in which aggregate demand is the number of patients who had been waiting three months or less at the end of the quarter, divided by the population in the region. Their measure of (inverse) health or quality is the waiting time perceived by patients when consider joining the waiting list. In these terms, they find that $\varepsilon_{n,-B} = -0.3$ (or -0.2 in an alternative model).

---

[8] Since $\dfrac{\partial R}{\partial c} = x^*$.

The practical implication on the level of optimal reimbursement is not to determine the absolute value of the margin over variable costs $\left(1/A\varepsilon_{n,B}\right)$ but only its variability across regions and time to adapt the optimal reimbursement accordingly.

2.3. Not-for-profit behaviour

In the baseline model, we assumed profit-maximization as the purpose of the provider organization. The literature considers alternative purposes for government and not-for-profit providers. The ambiguous objective of the not-for-profit provider has been formalized under different models. McGuire (2000) summarizes the not-for-profit models in two kinds of settings: (i) provider's intrinsic motivation that considers the patient's best interest, and (i) medical ethics as a constraint of choices. Sloan (2000) summarizes models of not-for-profit hospital behaviour as derived from the pioneering model of Newhouse (1970), assuming the maximization of quality and quantity subject to zero profit. Lastly, both Sloan (2000) and McGuire (2000) summarize models of "physician's cooperative" and "target income". In both cases, the decision-maker is the physician whose objective is to maximize or achieve a level of income.

2.3.1 Intrinsic motivation

The first way of considering not-for-profit behaviour is through the provider intrinsic motivation model which takes into account the patients' welfare into the provider's utility function. Within the framework of our baseline model, our model of intrinsic motivation considers the provider as maximizing a utility function that depends on profits and on patient benefits: $U\left(\pi\left(x,e\right),B\left(x,e\right)\right)$. Ellis and McGuire (1986) also model the not-for-profit behaviour assuming this utility function. However, in their model, quantity is the unique argument of profits and patient's benefits, and they assume a monopolist provider. They show that hospital incentives are distorted under a pure prospective reimbursement system when the decision maker weights hospital profits more than patient's benefits, that is, the decision maker is not a perfect agent. In this case, the provider undersupplies services (skimping), encourages low-value admissions (creaming) and discourages high-cost cases (dumping). The introduction of competition diminishes the distortion created by the prospective payment system (Pope, 1989).

We generalize Ellis and McGuire (1986) model by introducing quality and competition in the same terms than in the baseline model. Now, physicians, as managers of the not-for-profit hospital, decide on quality and quantity to solve the following problem:

(8) $\underset{x,e}{Max}\ U(\pi(x,e), B(x,e))$

where $\pi(x,e)$ is defined in program (1). The first order conditions give:

(9) $U_x = \frac{\partial U}{\partial \pi}\left[n' B_x (R - cx) - nc\right] + \frac{\partial U}{\partial B} B_x = 0$

(10) $U_e = \frac{\partial U}{\partial \pi}\left[n' B_e (R - cx) - nc_e x\right] + \frac{\partial U}{\partial B} B_e = 0$

The first characteristic of this solution is that the level of quality is not determined by the technology of health production as it may happen under the profit maximizing condition (4).

From condition (9) we can derive the same condition than in Ellis and McGuire (1986):

(11) $-\pi_x = \frac{(\partial U / \partial B)}{(\partial U / \partial \pi)} B_x = MRS_{\pi,B} B_x = \alpha B_x$

and a similar condition for quality is derived from (10):

(12) $-\pi_e = \frac{(\partial U / \partial B)}{(\partial U / \partial \pi)} B_e = MRS_{\pi,B} B_e = \alpha B_e$

where $\alpha = MRS_{\pi,B}$ is the weight that the physician gives to patients' benefit relative to hospital's profit. If $\alpha = 1$, the physician is a perfect agent with the same objectives than the social planner. If $0 < \alpha < 1$, the physician values hospital's profits more than patient's benefits. This is the most realistic assumption according to Ellis and McGuire (1986). At one extreme, $\alpha = 0$ for the pure for-profit. At the other extreme, $\alpha = \infty$ for the fully-altruistic provider.

To compare the solutions of the above problem with the socially optimal quantity and quality, let's solve the social planner problem: the problem of a perfect agent monopolist.

The socially optimal allocation of quantity is obtained according to the equality:

(13) $-\pi_x = B_x$.

Since, $\pi = R - c(e)x$ for the monopolist, we confirm that the perfect agent allocates the optimal quantity x^* such that

(14) $B_x = c$

The socially optimal quality is obtained under the first order condition:

(15) $-\pi_e = B_e$

Therefore, the optimality condition gives a relationship between quality and quantity such as:

(16) $x^* = \frac{B_e}{c_e}$

Our model is not based on a social planner model since we introduce competition. Therefore, we can compare allocations of quality and quantity in our model with the socially optimal levels and analyze the effects of competition and of the level of provider's altruism.

From condition (9), the allocation of quantity in the case of monopolistic competition gives:

(17) $B_x = \frac{c}{\frac{\varepsilon_{n,B}}{B}(R - cx) + \frac{\alpha}{n}}$

Therefore, we ascertain that the perfect agent allocates the optimal quantity if the provider is a monopoly ($\varepsilon_{n,B} = 0, n = 1, \alpha = 1$). In other case, the perfect agent allocates quantity such that $B_x < c$, that is, above the socially optimal level.

If the agent is less altruistic than the perfect agent $(0 < \alpha < 1)$, there will be underprovision in the case of monopoly or zero average profitability, but the agent can achieve the optimal quantity in cases of monopolistic competition and positive average profitability.

Secondly, to analyze the allocation on quality, condition (10) gives:

(18) $\frac{B_e}{c_e} = \frac{x}{\frac{\varepsilon_{n,B}}{B}(R - cx) + \frac{\alpha}{n}}$

As we saw before, the perfect agent monopolist ($\alpha = 1, n = 1,\ \varepsilon_{n,B} = 0$) allocates the optimal level of quality. In the case of monopolistic competition ($n < 1,\ \varepsilon_{n,B} > 0$) we can see that the perfect agent allocates quality so that

(19) $\frac{B_e}{c_e} < x$

Therefore, the perfect agent overprovides under condition (19). Condition (18) shows that, in monopolistic competition, there is scope to achieve the optimal quality under the more realistic assumption of a less altruistic agent than the perfect agent. The pure for-profit can also achieve the optimal quality in a

competitive market. The higher the degree of competition the higher the probability the agent overprovides quality. As in Pope (1989), this result indicates that the introduction of competition diminishes the distortion (to underprovision) created by a non-altruistic behaviour. Analogously, the higher the level of altruism the higher the probability of overprovision of quality.

2.3.2. Ethics constraints

Ethics has been argued to dictate the type and quantity of treatment that a patient should receive. McGuire (2000) states that, in this case, economic incentives on either the demand or the supply side would not determine health care choices and they do. However, if we expect that ethics govern the provider behaviour besides profit, it should play a stronger role in the not-for-profit than in the for-profit provider.

Ma and McGuire (1997) study a model of interaction among the insurer, the provider and the patient to derive insurance contracts to consumers and types of reimbursement to providers such that these contracts are optimal in the trade-off between risk-sharing (welfare enhancing due to the insurance) and moral hazard (welfare hampering due to the "hidden action" of providers and patients). The provider is modelled as a profit-maximizing agent with an ethics constraint. The behaviour of a not-for-profit is modelled subject to a tighter ethics constraint.

McGuire (2000) and our baseline model are related to Ma and McGuire's (1997), where competition is introduced. We introduce ethics constraints in our baseline model so that the provider's problem is:

$$(20)\quad \left.\begin{array}{l} \underset{x,e}{Max}\ \pi = n(B)\left[R - c(e)x\right] \\ subject\ to:\ \ B(x,e) \geq \overline{B} \end{array}\right\}$$

If the ethics constraint $\left(B(x,e) \geq \overline{B}\right)$ is not binding, the provider's behavior is identical to that in the baseline model.

The value $\overline{B}$ is assumed to be higher for the not-for-profit provider so that the ethics constraint either is binding for the not-for-profit and non-binding for the for-profit or is binding for the two types of providers but at different levels.

If the ethics constraint is binding, the provider solves the Lagrange problem:

$$(21)\quad \underset{x,e,\lambda}{Max}\ L = n(B)\left[R - c(e)x\right] - \lambda\left(\overline{B} - B(x,e)\right)$$

Quality and quantity are the solutions of the two following conditions:

(22) $\frac{B_e}{c_e} = x\frac{B_x}{c}$

(23) $B(x,e) = \overline{B}$

As before, a socially optimal quantity, $B_x = c$, is associated with the socially optimal quality. Moreover, overprovision of quantity $(B_x < c)$ implies overprovision of quality $\left(\frac{B_e}{c_e} < x\right)$.

In this case, competition does not affect the solution.

2.3.3. Alternative models of not-for-profit behaviour

Physician cooperative and target income

These theories were developed to acknowledge the power of physicians to induce demand and to maintain their income levels. The target income theory was used in the 80's to explain some anomalies such as the increase in price in medical services at times of increase in competition among physicians. Later, with the regulation of fees by the insurance companies, the negative relationship between fees and quantity was also explained by target income theories. McGuire (2000) discusses how it is not necessary to view "target income" as an alternative to profit maximization. In the context of profit maximization, income effects and participation constrains can explain the negative relationship between fees and quantity.

On the side of physicians' cooperative, the maximization of income per medical staff member can explain how an increase in demand can lead to an increase in price (if prices are not regulated) and a reduction in output and medical staff size. Physicians have more control in not-for-profit hospitals because of the lack of a well-defined residual claimant (shareholders act as residual claimant in for-profits). In this case, physicians can act as a cartel as residual claimants. In concrete, the model by Pauly and Redisch (1973) stress that the only difference between the short-term allocation (fixed medical staff) of a not-for-profit physician cooperative and an orthodox for-profit is that, in the first, economic profits are captured by physician inputs and, in the latter, by nondebt capital (shareholders).

Capital Price Theories

An important difference between for-profit and not-for-profit hospitals is the type of capital they use. For-profit capital is mainly equity while not-for-profit capital is obtained from tax-deductible donations and debt, mainly tax-exempt debt. These differences may drive different investments in the initial fixed capital or high technological equipment. For example, Wedig et al. (1996) argue that differences in investment

decisions between for-profits, not-for-profits and government hospitals are explained because not-for-profits and government hospitals are exempted from corporate tax, have a project financing constraint (this rule requires the organization to undertake a set of approved capital projects whose dollar value equals or exceeds the flow of tax-exempt financing), and are constrained against share issuance. Lakdawalla and Philipson (1998) argue that not-for-profit penetration is expected to be higher when opportunity costs of the capital invested is lower. The differences in opportunity costs of capital are mainly explained by location and, in fact, in the U.S., for-profit hospital market share is higher in states in which capital cost recovery is better (Sloan, 2000). Chakravarty et al. (2006) argue that, since not-for-profits need preferential financing conditions and sometimes they cannot obtain them, the easiness to obtain capital would act in favour of the responsiveness of for-profits to exit and entry when profitability conditions change.

Market Theories

Lastly, market theories explain behavioural differences based on competition among for-profits and not-for-profits. Some empirical studies find evidence that for-profits influence not-for-profits (e.g. Dugan, 2000; Silverman and Skinner, 2000; Cutler and Horwitz, 2000). The study done by Cutler and Horwitz (2000) with U.S. hospitals shows that for-profits are more financially successful than not-for-profits because of their greater skills to capture public sector reimbursements. This success may explain why not-for-profits followed for-profits' behaviour, in particular, for exploiting Medicare loopholes so as to increase reimbursements from Medicare. As competition increases, there is less scope to inefficiencies linked to not-for-profit behaviour. However, this does not imply a direct conversion to for-profit ownership. In many case, mergers and vertical integration of insurance companies, hospital and purchasers combine different types of ownership which may produce a variety of outcomes.

3. The NHS: Empirical evidence

3.1. Competition in the NHS

Competition in the NHS was introduced by the Conservative government during the 'internal market' reforms under the *NHS and Community Care Act* of 1990. The internal market empowered health authorities, NHS trusts and GP fundholders to purchase clinical services from the independent sector. From 1997 to the beginning of 2000 ('old Labour' phase), the Labour Government introduced a number of policies that restricted competition, among them, the abolition of GP fundholding in 1999. The 'new Labour' phase starting in 2000 with the *The NHS Plan* and the *'Concordat'* defined a national framework for partnership or 'concordat' between the independent sector and the NHS. *The Choice at Referral* –

Guidance Framework for 2006/07 published in April 2006 states that 'the IS [independent sector] providers who wish to offer services on the national menu that are not covered by existing ISTC contracts, must be approved to do so by the *Extended Choice* procurement due for completion in summer 2006. PCTs will be strongly encouraged to add local approved IS providers to their local menus in a fair, transparent and non-discriminatory manner.'

In 2008, patients will have the right to choose from any provider approved by the Healthcare Commission which provides services within the national maximum price (PbR tariff). Most of the major independent hospitals groups now have hospitals registered to accept referrals form the NHS. According to Laing's Healthcare Market Review 2007-08, there have been no reports of significant or unexpected growth of referrals under the *Extended Choice Network* since the start of 2007.

An increase in competition is also expected in management and commissioning services, since some of these services will be outsourced to the private sector (under the *Framework for Procuring External Support for Commissioners*). This competition is however limited. Laing's Healthcare Market Review 2007-2008 predicts that the only management areas open to competition with the private sector are data collection and contract management, but not commissioning services on behalf of PCTs.

On the other hand, Laing's Healthcare Market Review 2007-2008 points out a threat to competition of the *2000 NHS Plan*: competition with privately funded health may decrease. Competition is aimed for NHS patients and, by delivering a more accessible, timely and efficient elective surgery which is free at the point of use, privately funded health may lose market share which represents a 1.4 percent of GDP in 2007-08. If the NHS ends covering the whole health market, the consequences of competition on social welfare will be different. In Gaynor's (2006) words, if competition on quality leads only to demand stealing consumers are not better off as a result, but if it leads to any increase in market demand then consumers are unequivocally better off.

On the demand side, there is not a direct link between patient choice and competition since buyers are PCTs and GPs who place contracts on behalf of patients. During the internal market, buyers were GP fundholders and District Health Authorities. Currently, with the introduction of practice based commissioning under the 2006/07 *Directed Enhanced Service* and locally agreed incentive schemes, GPs have incentives to make savings from their budgets. Therefore, GPs, acting as agents for patients, are interested in price as well as quality since they must serve patient's wishes. The GPs' interest in price may generate a conflict between purchasers' and patients' choice. According to Propper et al. (2006), "greater patient choice may mean less patient choice as payers seek to develop relationships with a limited set of providers". Moreover, patients may accept passively GPs referrals due to lack of information, difficulty to evaluate such information, and the necessity of trust in doctors and practitioners.

Did competition really increase during the internal market? Propper et al. (2003) summarize contrasting views on the effect of the internal market policies on competition. Firstly, competition was limited among NHS Trusts, since buyers did not refer to independent sector providers. Secondly, competition was primarily within geographical areas. Third, competition took place through annual negotiations over contracts for block of services with regulated prices that were supposed to equal average costs. Among negative views, Allen (2002) points out that the internal market was not conductive to increase choice for patients. Le Grand (1999) also states that patient choice did not increase, GP fundholding did not have any impact, and control was centrally retained. Duseiko et al. (2007) state that the internal market did little to change historical patterns of commissioning by Health Authorities and non-fundholding practices. Among the positive views, Propper et al. (2003) conclude that, although the number of competitors in any one local market did not increase during the internal market period, the policy reform introduced competition. This was due to the introduction of fundholding practices that had incentives to 'shop around' in search of lower prices and better quality (Propper, 1996; Glennerster et al., 1994; Strong and Lloyd, 1997), and hospitals were more responsive to fundholdres than to Health Authorities (Mays et al., 2000). Yet Ranade (1994) presents evidence on changes of providers in the District Health Authorities' contracts. Duseiko et al. (2007), using concentration indexes, find that fundholding practices used more hospitals than non-fundholders.

Based upon the results of GP fundholders during the internal market, Duseiko et al. (2007) predict that GP practice-based commissioning, with reintroduction of practice budgets, "will lead to a small but noticeable reduction in concentration[9]". They also find that larger practices with more patients per GP have less concentrated admissions. Therefore, they also predict that the decrease in the number of PCTs (from 303 in April 2002 to 152 at the end of 2008, and the Department of Health has indicated its intention to reduce the number of PCTs to less than 100) will reduce concentration. Interestingly, the study includes a measure of commissioning concentration on 15 Health Resource Groups (HGR) that reimbursed by PbR since 2004, including cataract surgery. They find a slight increase in concentration index for these HGR through the period 1997/90 to 2002/03. The concentration index for cataract surgery was higher but did not change over the period. The effect of PbR cannot be evaluated since these HGR concentration indexes are not measured after 2004. Nevertheless, the increase in the number of hospitals with the entry of independent sector hospitals is expected to reduce concentration.

[9] They define concentration as the degree in which commissioners of elective hospital concentrate their admissions among hospitals. Their empirical study covers the period from 1997/98 to 2002/03.

Assuming that competition increased during the internal market, there is no evidence of a beneficial effect on quality. Propper at al. (2003) find an adverse effect: 'death rates (from acute myocardial infarction) were higher in areas subject to competition than in areas not subject to competition during the first part of the 1990s but declined somewhat from 1996 onwards'. We cannot claim that this finding contradicts our results because contracts during the internal market were based on block contracting. To the best of our knowledge, there are no studies analyzing the effect of competition under reimbursement systems based on block contracting. Yet Chalkley and Malcomsom (1998) find that block contracting is not an optimal contract in a realistic scenario; it can only be an optimal contract if the provider is fully benevolent.

3.2 Reimbursement system: PbR

'Most independent sector hospitals operators view this [payment by results] as being the truly fundamental change introduced under the government's reform programme, and the major platform upon which sustainable NHS funded activity can be built in the long term, subject to continuing political support' (Laing's Healthcare Market Review 2007-2008).

The PbR tariff was launched in *Delivering the NHS Plan* (July 20002) with the announcement of the Health Resource Group (HRG) benchmark that would be used to establish a standard national tariff for the same treatment regardless the provider. With *Reforming NHS Financial Flows* (October 2002) the DH set the plan to substitute block contracting with Service level Agreements, initially on a cost and volume basis, and with the PbR tariff fully applicable from 2008. PbR is expected to be the main driver of competition for patients since any provider with approved standards can supply services if it does within the tariff.

The independent sector has several costs disadvantages *vis* a *vis* NHS providers. The most important is the consideration of return to capital and profit in the tariff. NHS reference costs allow for a 3.5% return on the value of all of trusts' relevant assets, which is well below of an expected 15% IIR (Internal Rate of Return) for many hospital operators.

To facilitate entry of new providers with high start-up costs and that may offer new services, PCTs are allowed to pay providers more than the national tariff and to offer guarantees such as long-term contracts, minimum income levels, or different types of ownership that reduce initial capital investment.

These differences in tariff applied to different providers according to Market Forces Factors (MFF) may be a threat to competition since they may serve the NHS commissioner as an argument to drop providers (in particular, independent providers) from their Choose and Book menu. This has been the case of Leeds

Primary Care Trust which advised GPs to refer only to NHS hospitals and ISTC and retire referrals to BUPA and Nuffield private hospitals for which the tariff was a 4% higher.

The adjustment rate MFF depends on costs of land, buildings and staff in each geographical area. We have seen that the MFF should equal $\left(1+1/A\varepsilon_{n,B}\right)$ to achieve an optimal provision under profit maximization. The elasticity parameter A depends on costs per average output and this is taken into account in the calculation of the MFF. However, the MFF is not adjusted to the level of competition in a geographical area.

In accordance to the problem of heterogeneity of costs in the determination of the optimal tariff, Street and Maynard (2007a) show that day case activity is less costly than inpatient activity for three HRGs. A tariff of the HRG that is a weighted average of day case costs and inpatient costs incentivises hospitals to undertake cheaper day case work. Malcomson (2007) points out that the tariff will result in both a rent for low-cost providers (day case) and the incentive to keep costs down for high-cost providers (inpatient). If tariffs are too low, the incentives to underprovide such as to provide the cheaper day-case treatment, even when inpatient treatment would be more appropriate, should be not ignored since high cost providers or treatments with higher costs for the same diagnosis are not fully reimbursed.

Regarding the administration of the system, Marini and Street (2007) estimate that PbR has increased transaction costs facing PCTs and hospitals by £130,000, on average, in the fiscal year 2004/2005 due mainly to recruitment of additional staff for clinical coding activities and financial services. These expenditures are expected to decrease once the initial investment in the system of coding and accounting is implemented.

Lastly, since PbR tariffs are based on both recurrent and capital costs, the system results in under-financing of new capital stock (Palmer, 2006). We have mentioned that this may represent a more serious problem in the case of for-profit providers with higher opportunity costs of capital, but we also mentioned the factors that, according to Wedig et al. (1996), may affect investment decisions apart from the tariff.

To deter adverse selection in the case mix and to control demand, Malcomson (2007) and Street and Maynard (2007a, 2007b) propose to implant two-part tariffs with supply-side cost sharing, a system that promises to be a good mechanism given the experience in other countries. Nevertheless, the amount of the marginal tariff necessary to achieve an optimal level diminishes with competition. We have seen that a fully-capitated system (with a zero marginal tariff) like PbR may induce an optimal provision given adequate levels of competition and/or provider's altruism.

3.3 NHS providers and the Independent Sector: For-profit and Not-for-Profit providers

In the U.S. the majority of hospitals are tax-exempt entities operated by charities or government agencies (5,100 hospitals out of a total of 6,500). In contrast, the majority of private hospitals in England are owned by for-profit companies. The private sector in the UK is mainly funded by private insurance (11 percent of the population in UK has private medical insurance), although public funding is important since the DH expects that the private sector carries out a 15 percent of the NHS procedures by 2008. Because the growing importance of public funding and volume of patients, many private companies have entered the health care market in England since the inception of *The NHS Plan* in 2000. The oldest hospital companies were not-for-profit companies: Nuffield Hospitals, formed in 1957, and BUPA (British United Provident Association) hospitals, formed in 1947 by the amalgamation of a number of "provident" not-for-profit associations. More recently, for-profit providers are the only entrants in the private healthcare sector. Since 2007, when BUPA hospitals were acquired by the private equity company Cinven Ltd. from the for-profit provider Spire Healthcare, the unique relevant not-for-profit provider in England is Nuffield Hospitals that owned 41 hospitals in England by mid-2007. It remains to see if private providers can keep their costs in line with national tariffs and remain in the market.

On the other hand, NHS providers must also keep their costs in line with the national tariff because they will have Foundation Trust (FT) status and their financial performance must comply with standards approved by Monitor. In 2004/5, NHS Trusts that obtained good financial and quality performance (three stars) had the choice of becoming FT. According to Marini et al. (2007), `while FT may have greater freedoms to borrow and invest, capital investment under the PbR regime may make it more likely that they incur deficits' because the payments system averages recurrent and capital costs. In their study on financial performance of FT created since April 2004, these authors find that FT, especially those that gained foundation status in 2005, retained larger surpluses before the implementation of the policy in 2004/5 than after achieving FT status. Also, they find that FT showed lower Reference Cost Index (RCI) that is a weighted average of the HRG unit costs relative to the national average. Based on the previous experience of NHS Trust status at the beginning of the 1990s, Bartlett and Le Grand (1994) show that Trust status does not cause a reduction of costs per se since NHS providers that became Trusts were a selected group of already more efficient providers.

If the financial performance of NHS Foundation Trusts is not satisfactory, should they close as private providers do? The Labour leadership would then have to deal with the politically controversial issue of whether private competition should be allowed to drive out public provision, as Laing's Healthcare Market Review 2007-2008 reckons.

4. Concluding remarks

The models we have constructed evaluate the effect of competition and provider's altruism on quantity and quality of health when the reimbursement system corresponds to PbR. In two models: the baseline and the intrinsic motivation models, we obtain that competition and altruism have a positive effect on quality and quantity. Moreover, competition allows achieving a socially optimal level including in the case of non-benevolent providers (those that care more for hospital's profits than for patients' benefits).

We are not aware of any evidence pointing out differences in levels of provision to NHS patients (quality and quantity) between providers which are assumed to have different levels of altruism: for-profits, not-for-profits and government hospitals. The only available evidence is on the effects of competition (among NHS providers uniquely) during the period of the internal market in the 1990s. Despite the existence of theoretical models showing that, under regulated prices, competition increases quality, mortality rates following heart attacks were higher in areas subject to more competition during the internal market. The existing pricing mechanism in the internal market was block contracting under which NHS providers competed over contracts for block of services through annual negotiations. These kinds of contracts cannot generally achieve a socially optimal provision. We do not know any model studying the effect of competition on health under this reimbursement system.

Among the targets of the current reform of the NHS, competition is a fundamental one. It is expected to be achieved by promoting patient choice among any providers that offer services within the PbR tariff. This will drive providers towards efficiency. Whether competition will be limited or not so as to threat the survival of less efficient providers, among them NHS providers, remains a political issue.

References

Allen, P. (2002) Plus ca change, plus c'est la meme chose: to the internal market and back in the British National Health Service. *Applied Health Economics and Health Policy* 1, 171-178.

Bartlett, W. and J. Le Grand (1994) Costs and Trusts. *Studies in Decentralization and Quasi-Markets*. University of Bristol.

Chakravarty, S., M. Gaynor, S. Klepper, and W. Vogt (2006) Does the profit motive make Jack nimble? Ownership form and the evolution of the US hospital industry. *Health Economics* 15, 345-361.

Chalkley, M. and Malcomson (1998) Contracting for Health Services when patient demand does not reflect quality. *Journal of Health Economics,* 17(1), 1-19.

Cutler, D and J. Horwitz (2000) Converting hospitals from not-for-profit to for-profit status: why and what effects?, in *The Changing Hospital Industry: Comparing Not-for-profit and For-profit Institutions* 45 (David M. Cutler ed.).

Dugan, M. (2000) Hospital ownership and public medical spending. *Quarterly Journal of Economics* 115, 1343-1374.

Dusheiko, M., M. Goddard, H. Gravelle and R. Jacobs (2007) Explaining trends in concentration of healthcare commissioning in the English NHS. *Health Economics*. DOI: 10.1002/hec.1301

Ellis, R. and T. McGuire (1986) Provider Behavior under Prospective Reimbursement, *Journal of Health Economics* 5, 129-151.

Gaynor, M. (2006) What do we know about competition and quality in health care markets? NBER working paper 12301.

Glennerster H, M Matsaganis M, and P Owens (1994) *Implementing GP Fundholding: Wild Card or Winning Hand?* Open University Press: Buckingham.

Gravelle, H., P. Smith and A. Xavier (2003) Performance signals in the public sector: the case of health care. *Oxford Economic Papers-New Series* 55, 81-103.

Laing's Healthcare Market Review 2007-2008.

Lakdawalla D and T. Philipson (1998). Nonprofit production and competition. NBER Working Paper 6377.

Le Grand, J. (1999) Competition, Cooperation, or control? Tales from the British National Health Service. *Health Affairs* 18, 27-37

Ma and McGuire (1997) Optimal Health Insurance and Provider Payment. *American Economic Review*, 87(4), 685-704.

Malcomson, J.M. (2007) Hospital cost differences and payment by results. *Health Economics, Policy and Law* 2, 429-233.

Mays N, J Mulligan J, and Goodwin (2000) The British quasi-market in health care: a balance sheet of the evidence. *Journal of Health Services Research and Policy* 5, 49–58.

McGuire, T. (2000) Physician Agency. Handbook of Health Economics Vol 1A (Culyer and Newhouse eds.)

Marini, G. and A. Street (2007) A transaction costs analysis of changing contractual relations in the English NHS. *Health Policy* 83, 17-26.

Marini, G., M. Miraldo, R. Jacobs and M. Goddard (2007) Giving greater financial independence to hospitals - does it make a difference? The case of English NHS Trusts. *Health Economics* . DOI: 10.1002/hec.1292

Newhouse, J. (1970) Toward a Theory of Nonprofit Institutions: An Economic Model of Hospital. *American Economic Review* 60 (1), 64-74.

Palmer, K. (2006) *NHS Reform: Getting Back on Track.* Kings Fund: London.

Pauly M, and M. Redisch (1973). The not-for-profit hospital as a physician's cooperative. *American Economic Review* 63, 87–99.

Pope, G. (1989) Hospital nonprice competition and Medicare reimbursement policy. *Journal of Health Economics* 8, 137-172

Propper, C. (1996) Market Structure and Prices: The Responses of Hospitals in the UK National Health Service to Competition, *Journal of Public Economics*, 61, 307-335.

Propper, C., S. Burgess and D. Gossage (2003) Competition and quality: Evidence from the NHS internal market 1991-1999. Unpublished paper, University of Bristol.

Propper, C., D. Wilson and S. Burgess (2006) Extending choice in English health care: The implications of the economic evidence. *Journal of Social Policy* 35, 537-557.

Ranade, W (1994) *A Future for the NHS?: Health Care in the 1990's* , London: Longman

Silverman, E. and J. Skinner (2000) Are for-profit hospitals really different? Medicare upcoding and market structure. NBER working-paper 8133.

Sloan (2000) Not-For-Profit ownership and hospital behavior. *Handbook of Health Economics* Vol 1B (Culyer and Newhouse eds.)

Street, A. and A. Maynard (2007a) Activity based financing in England: the need for continual refinement of payment by results. *Health Economics, Policy and Law* 2, 419-427.

Street, Andrew and A. Maynard (2007b) Payment by results: qualified ambition? *Health Economics, Policy and Law* 2, 445-448.

Strong C, Lloyd SH. (1997) An application of Keaveney's model of service switching behaviour to GP fundholders. *Public Money and Management*, 47–51.

Wedig, G., M. Hassan and M. Morrisey (1998) Tax-exempt debt and the capital structure on nonprofit organizations: An application to hospitals. *The Journal of Finance*, 51(4), 1247-1283.

Windmeijer, F., H Gravelle and P Hoonhout (2005) Waiting lists, waiting times and admissions: an empirical analysis at hospital and general practice level. *Health Economics* 14, 971-985.