

THE UNIVERSITY *of York*

**CENTRE FOR HEALTH ECONOMICS**

**Capitation: Access and Quality**

Hugh Gravelle

***CHE Technical Paper Series 2***



# **Capitation: Access and Quality**

**Hugh Gravelle**

**December 1996**

## **CENTRE FOR HEALTH ECONOMICS TECHNICAL PAPER SERIES**

The Centre for Health Economics has a well established Discussion Paper series which was originally conceived as a means of circulating ideas for discussion and debate amongst a wide readership that included health economists as well as those working within the NHS and pharmaceutical industry.

The introduction of a Technical Paper Series offers a further means by which the Centre's research can be disseminated. The Technical Paper Series publishes papers that are likely to be of specific interest to a relatively specialist audience, for example papers dealing with complex issues that assume high levels of prior knowledge, or those that make extensive use of sophisticated mathematical or statistical techniques.

The series provides a maximum of 20 pages for authors. The content and its format are entirely the responsibility of the author, and papers published in the Technical Paper series are not subject to peer-review or editorial control, unlike those produced in the Discussion Paper series. Offers of further papers, and requests for information should be directed to Frances Sharp in the Publications Office, Centre for Health Economics, University of York.

### **THE AUTHOR**

Hugh Gravelle is Professor of Economics in the Centre for Health Economics. He leads the York contribution to the National Primary Care Research and Development Centre - a collaboration of the Universities of Manchester, Salford and York.

Price £4.00

© Hugh Gravelle

# Capitation: access and quality

Hugh Gravelle\*

## Abstract

This paper examines the implications of public and private funded capitation systems for the quality of care and the number of practices. Under private funding there are three types of market equilibria. Despite the fact that practices face downward sloping demand curves quality is efficiently supplied in all types of equilibria. However, there may be too few practices, with some patients not joining any list, or too many. Entry control by a self-regulating profession has ambiguous welfare implications but collusion over capitation fee and quality is welfare reducing. The optimal capitation fee in a public tax financed capitation system is derived. Quality and practice numbers are increasing in the capitation fee but welfare maximisation requires control of numbers as well as the fee. A median voter model of the capitation fee is outlined and shown to lead to too low a fee and quality. When voters bargain with a professional union the fee and quality are too high.

*Keywords:* Capitation. Quality. General practice. Product differentiation.

JEL No.: I1, L13

---

\*National Primary Care Research and Development Centre, Centre for Health Economics, University of York, Heslington, YO1 5DD; email: hg8@york.ac.uk. Support from the Department of Health to the NPCRDC is acknowledged. The views expressed are those of the author and not necessarily those of the Department of Health. I am grateful to participants in the Health Economics Study Group, Brunel University, June 1996, for helpful comments.

# 1 Introduction

The paper examines some of the properties of alternative forms of pure capitation systems where patients register with a doctor or practice and the doctor is paid a fee for each patient on her list. In exchange for the fee the doctor accepts the obligation to provide services to the patient at some point in the future.<sup>1</sup> Under a private capitation system the fee is paid by the patient. One example is the American HMO.

Under a public system, like the British NHS, the capitation fee is paid by the state rather than by the patient and is funded from taxation. Recent changes in the NHS have been placed a greater emphasis on the capitation system. Capitation payments to GPs for patients on their lists were increased under the 1990 contract, increasing the proportion of income received from this source. Amendments to regulations have made it easier for patients to move from one practice to another. The aim of the changes was to increase the incentives for GPs to attract and retain patients by maintaining and improving the quality of the services they provide.

Some changes in the characteristics of services delivered by GPs will be regarded as improvements by all patients: shorter waiting times for appointments, longer surgery opening hours, opening a branch practice... This is vertical product differentiation. Other changes may make some patients better off and others worse off. This is horizontal product differentiation. Examples include replacing an afternoon surgery with an evening surgery or moving the practice premises from one location to another. A patient's choice of practice depends on both aspects of the service. GPs compete for patients by both vertical and horizontal service differentiation.

To date there has been no analysis of the implications of private and public capitation systems for quality and the number of practices. This paper fills this gap in the literature by extending the Salop (1979) model of product differentiation to capture some important features of the market for doctors services. Section 2 sets out the basic model in the context of a private system in which patients pay the capitation fee to their GP. In section 3 fees, quality and entry are controlled by a self-regulating profession. Section 4 considers the first and second best welfare properties of these private markets.

Section 5 presents a public capitation system in which the capitation fee is paid by the state and financed by taxation. The section examines the quality implications of increasing the fee and the optimal regulatory policy. Section 6 outlines a public choice model of the capitation fee in a public capitation system and examines the choice of fee when voters consider the tax implications of increasing the fee. It also considers the Nash Bargaining Solution when the voters bargain with a monopoly union of doctors.

# 2 Model

The model is an extension of the Salop (1979) horizontal product differentiation model. A patient derives benefit  $q - td$  from belonging to the list of a GP located at a distance  $d$ .  $d$  is best interpreted as the geographical distance from the surgery to the patient's home. It may also be interpreted as the difference between the level of some horizontally differentiated service characteristic of the practice and the level which would maximise the utility of the particular patient.  $q$  is the maximum amount the patient would be willing

---

<sup>1</sup>The doctor may also be paid on a fee per item of service basis for some of the services provided. This makes no essential difference to the analysis in this paper, which is concerned with the implications for quality, rather than quantity, of the package of services provided.

$p$	payment to GP per patient on list
$d$	patient distance to practice
$t$	patient cost per unit of distance
$q$	quality of service provided to patient at practice
$q - td$	benefit to patient at $d$
$M$	number of patients uniformly distributed around circle of circumference $K$
$m = M/K$	uniform density of patients
$\delta$	market segment of practice
$D_i = m\delta$	demand (list) of practice $i$
$C_i = \alpha D_i q_i^2$	variable cost function of practice $i$
$F_i$	fixed cost of practice (reservation wage)
$\pi_i = pD_i - C_i$	gross profit of practice $i$
$N$	number of practices

Table 1: Notation

to pay to belong to this practice, rather than belong to no practice, if its distance from her preferred location was zero. It measures the vertical quality of the service provided by the practice.<sup>2</sup>

Patients have the same valuation of vertical quality characteristics and incur the same marginal distance cost  $t$ . They differ only in their location or in their taste for horizontal quality. The  $M$  identical consumers are distributed uniformly around a circle of circumference  $K$  with density  $m = M/K$ . The  $N$  profit maximizing GP firms are located around the circle.

If a patient located at distance  $d$  from GP firm  $i$  pays a capitation fee  $p_i$  to join its list he gets a net benefit of  $q_i - p_i - td$ . As the practice reduces its price below  $q$  it initially attracts patients who would not have joined the list of any other firm because they were too far away. In this monopoly range the practice does not attract patients from other lists. As the price is reduced further the practice's additional patients are those who switch from neighbouring practices. In this competitive range of prices the demand curve for the practice is less elastic since the alternative for the new patients is the list of the other practice, rather than receiving no care. The demand curve for the practice has a kink at the point where it starts to attract patients from neighbouring practices and at this point the its marginal revenue curve is discontinuous.

Depending on the parameters of the model, there are three possible equilibria: on the monopoly segment of the practice demand curves, at the kink and on the less elastic competitive segment of the demand curve where practices are affected by the prices and qualities of neighbours.<sup>3</sup>

<sup>2</sup>In an appendix (not included but available on request) it is shown that the results for the private capitation model are unchanged if  $q$  is the patient's value of the contract with the service provider. It makes no difference if the practice provides insurance to risk averse patients and sells services to them on a fee per item of service basis when they are ill. The premium paid is the capitation fee and competition leads the GP firm to offer an efficient insurance contract, including copayments to regulate patient demand. See Sheldon (1990) for an example of an optimal insurance contract with capitation features.

<sup>3</sup>There is a fourth super-competitive segment of the demand curve where the firm reduces its price sufficiently to take all of the demand from its neighbour and attract patients from practices located on the far side of its immediate neighbour. Its demand curve is discontinuous at the price where it just captures

For each equilibrium we first examine the Nash equilibrium price and quantities with a given number of practices. These prices and quantities produce a profit for each practice which is non-increasing in the number of firms. We then close the model by assuming free entry and an increasing supply price for GP firms.

## 2.1 Competitive equilibrium

In this case practice  $i$  will get patients to its right up to the distance  $d_{i+1}$  defined by

$$q_i - p_i - td_{i+1} = q_{i+1} - p_{i+1} - t[\Delta_{i+1} - d_{i+1}] \quad (1)$$

where  $\Delta_{i+1}$  is the distance between  $i$  and its rival to its right. (I label firms so that  $i + 1$  is the next firm to right and  $i - 1$  the next firm to the left.) Solving

$$d_{i+1} = \frac{(q_i - p_i) - (q_{i+1} - p_{i+1}) + t\Delta_{i+1}}{2t} \quad (2)$$

The firm's market area to its left  $d_{i-1}$  is similarly defined. Total demand for the list of firm  $i$  is

$$\begin{aligned} D_i &= md_{i+1} + md_{i-1} \\ &= \frac{m}{t}(q_i - p_i) - \frac{m}{2t}[(q_{i+1} - p_{i+1}) + (q_{i-1} - p_{i-1}) - t(\Delta_{i+1} + \Delta_{i-1})] \\ &= D_i(p_i, q_i; p_{i+1}, p_{i-1}, q_{i+1}, q_{i-1}) \end{aligned} \quad (3)$$

The demand for firm  $i$  is increasing in its quality and the capitation fees of its immediate rivals and decreasing in its fee and the quality of its rivals.

GP firms have identical cost functions. The total cost of firm  $i$  is  $C_i = \alpha D_i q_i^2 + F_i$  where  $F_i$  is fixed cost.<sup>4</sup> We interpret  $F_i$  as a reservation wage for the GP and assume that the distribution of reservation wages across potential GPs is such that there is a positively sloped inverse supply curve

$$F = F(N), \quad F'(N) > 0 \quad (4)$$

A practice's vertical quality is the same for all its patients and each patient costs the practice  $\alpha q_i^2$ . Gross profit is

$$\pi_i = (p_i - \alpha q_i^2) D_i(p_i, q_i; p_{i+1}, p_{i-1}, q_{i+1}, q_{i-1}) \quad (5)$$

There is three stage game: firms choose whether to enter; given that they enter they choose a location (horizontal differentiation); having chosen a location they compete in vertical quality and capitation fee. With identical firms and patients it seems reasonable

---

its neighbours entire list. Equilibria cannot occur on super-competitive segments because a firm which has lost its entire market will leave the industry and the demand discontinuity will disappear.

<sup>4</sup>For an interesting welfare problem it is necessary to assume either that GP costs are a convex function of quality or that patient valuation is concave function of vertical quality. Our assumption that quality is measured by willingness to pay means that patient quality valuation is linear.

Economides (1993) also examines vertical quality in a Salopian framework. In his model consumers have different valuations of vertical quality but equilibrium quality depends on the average valuation. This is not essentially different from the model of this paper where all consumers have the same valuation of vertical quality. However, he also assumes that the cost function is separable in quality and quantity produced. In the GP context this would correspond to quality being a public good for all patients in the practice. For example a GP may spend more money on better software for organising patient records. In the model in the paper increasing quality increases the average and marginal cost to the practice of each patient. One example would be the GP spending longer with each patient or providing more diagnostic tests.

to assume that a symmetric Nash equilibrium will exist in which firms locate equidistantly around the circle and choose the same quality and capitation fee.<sup>5</sup>

Given the number of firms and their location, GP firm  $i$  chooses its capitation fee and quality to maximise its gross profit.<sup>6</sup> These satisfy

$$\pi_{ip_i} = D_i + [p_i - \alpha q_i^2] D_{ip_i} = 0 \quad (6)$$

$$\pi_{iq_i} = [p_i - \alpha q_i^2] D_{iq_i} - \alpha D_i 2q_i = 0 \quad (7)$$

taking the fees and qualities of its rivals as given.

Using the fact that  $D_{iq_i} = -D_{ip_i}$  gives the competitive market equilibrium quality as

$$q^c = \frac{1}{2\alpha} \quad (8)$$

Note that even without imposing symmetry *the firm's quality is independent of its rival's qualities and capitation fees, the density of patients and the number of firms.*

The intuition, which also applies in the kink and the monopoly case, is that an increase of £1 in quality supplied to all patients enables the firm to raise its fee to them by £1 and to keep its list size constant. The additional revenue per patient thus obtained is £1. The additional cost incurred per patient is  $2\alpha q_i$ . Since the marginal revenue and cost per patient from raising quality and adjusting the fee are unaffected by the number of patients none of the other factors affecting demand have any effect on the profit maximising quality.<sup>7</sup>

Rearranging (6) and using the equidistant location assumption ( $\delta_{i+1} = \delta_{i-1} = K/N$ ,  $\forall i$ ), the solution for quality ( $q_i = q^c = 1/2\alpha$ ,  $\forall i$ ) and the symmetry assumption ( $p_i = p^c$ ,  $\forall i$ ), we get the competitive Nash equilibrium capitation fee:

$$p^c = \frac{1}{4\alpha} + \frac{tK}{N} \quad (9)$$

The equilibrium capitation fee exceeds the marginal cost of an additional patient ( $\alpha q_i^2 = 1/4\alpha$ ).

Using (8) and (9) we can calculate the gross profit  $\pi^c(N)$  of each firm at the competitive Nash equilibrium for a given number of firms. Table 2 summarises the fee, quality, list size, and gross profit.

## 2.2 Monopoly equilibrium

In the monopoly case some potential patients choose not to join the list of any practice because their distance cost  $td$  to the nearest practice exceeds their net valuation  $q - p$

---

<sup>5</sup>See Economides (1989, 1993) for demonstrations that this is so for a similar model.

<sup>6</sup>A full account of quality determination would include more complex doctor objectives and information asymmetry. We neglect these aspects to keep the analysis of those aspects of doctor decision making which are our main concern tractable.

<sup>7</sup>In Economides (1993) the equilibrium quality is a decreasing function of the number of firms. The reason is that in his model the cost of quality is fixed with respect to the number of the firm's customers. With more firms each firm has fewer customers and therefore must charge each a higher price to finance any given level of quality.

		Monopoly	Kink	Competition
Price	$p$	$\frac{3}{8\alpha}$	$\frac{1}{2\alpha} - \frac{tK}{2N}$	$\frac{1}{4\alpha} + \frac{tK}{N}$
Quality	$q$	$\frac{1}{2\alpha}$	$\frac{1}{2\alpha}$	$\frac{1}{2\alpha}$
List size	$D$	$\frac{M}{4\alpha tK}$	$\frac{M}{N}$	$\frac{M}{N}$
Gross Profit	$\pi$	$\frac{M}{32\alpha^2 tK}$	$\frac{M}{4\alpha N} - \frac{tKM}{2N^2}$	$\frac{tKM}{N^2}$
Marginal surplus	$S_N^*$	$\frac{3M}{64\alpha^2 tK}$	$\frac{tKM}{4N^2}$	$\frac{tKM}{4N^2}$
Range of $N$	$N$	$[1, 4\alpha tK]$	$[4\alpha tK, 6\alpha tK]$	$[6\alpha tK, \infty)$

Table 2: Equilibria of private capitation system

of vertical quality. The number of patients to the right of practice  $i$  who join its list is  $d_{i+1} = m(q_i - p_i)/t$  and so

$$D_i = 2m(q_i - p_i)/t. \quad (10)$$

Substituting the monopoly demand function into the profit function (5) and maximising yields the monopoly fee, quality, and maximised profit  $\pi^m(N)$  shown in Table 2. Price, quality and profit are unaffected by the number of firms since there are gaps in the coverage of the market and each firm's profit depends only on its decisions.

### 2.3 Kink equilibrium

The slopes of the monopoly demand curve ( $-2m/t$ ) and the competitive demand curves ( $-m/t$ ) differ (see (3),(10)). In the kink case the equilibrium is at the kink in the demand curve. At this point the market is just covered. If the price is raised some consumers will not join the practice, whilst if it is reduced some patients will be attracted from neighbouring practices. Each practice's marginal revenue curve is discontinuous. At a kink equilibrium the price and quantity must be such that the monopoly demand is just equal to the competitive demand which is, since the market is covered,  $M/N$ . Hence  $2m(q - p)/t = M/N$ . Maximizing profit by choice of price and quality subject to this constraint gives the kink price, quality and maximum profit shown in Table 2.

### 2.4 Prices, profit and number of practices

Practice profit varies with the number of firms. Since there is free entry of GP firms the equilibrium number of firms is determined by the condition that the marginal GP earns a gross profit just equal to her reservation wage:

$$\pi(N) = F(N) \quad (11)$$

where  $\pi(N)$  is the Nash equilibrium gross profit. Which of the three equilibria outlined above occurs depends on the parameters of the model which determine the shape of the

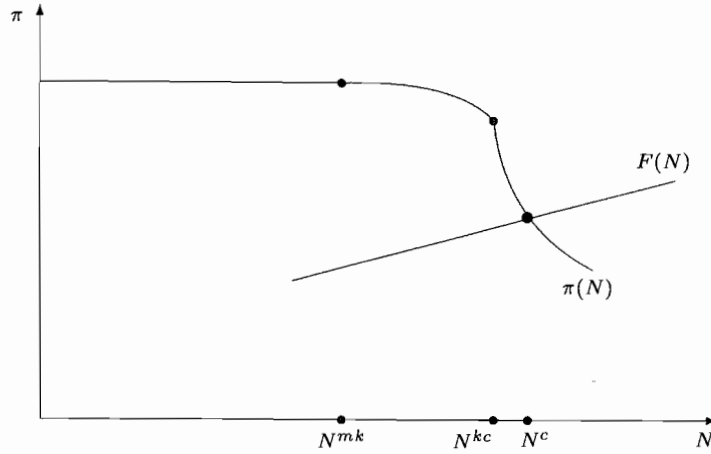


Figure 1: Equilibrium in competitive range

gross profit function and the inverse GP supply function.<sup>8</sup>

Figure 1 plots gross practice profit  $\pi$  against the number of practices. Increases in the number of practices reduce profit except in the monopoly equilibrium where profit is unaffected. The equilibrium number of practices is  $N^c$  where the inverse supply curve of GPs cuts the gross profit curve. In Figure 1 the equilibrium is competitive, but with an upward shift in the supply curve or downward shift in the profit curve either of the other two types could result.

## 2.5 Comparative statics of market equilibrium

Table 3 summarises the comparative static properties of the private capitation system.<sup>9</sup> Equilibrium quality is the same in all equilibria and depends only on the cost parameter  $\alpha$ . Consequently many of the comparative static responses examined in Salop (1979) (where quality is fixed) carry over. However, because we have not normalised the circumference (market length  $K$ ) and introduced the production cost parameter  $\alpha$  we can examine the effects of these parameters on the equilibrium. In addition because the upward sloping GP supply curve it is now possible to investigate the comparative static properties of the monopoly equilibrium.

*An increase in the parameter  $\alpha$  which increases the marginal cost of quality and of an increased list size reduces quality and the GP firm's capitation fee in all cases. This is so*

<sup>8</sup>Salop (1979) assumes that  $F$  does not vary with the number of firms which simplifies the analysis in the competitive and kink cases. However, the assumption makes it impossible to conduct comparative static analysis of the monopoly case since practice gross profits do not vary with the number of firms under monopoly. With free entry and a positive monopoly profit, the number of practices would increase until the market is covered, leading to one of the other solutions. If the profit is negative the number of practices will fall to zero. Other papers developing the Salop model also retain the assumption that there is a perfectly elastic supply of firms. See, for example, Economides (1989, 1993) and Norman (1989).

<sup>9</sup>The boundaries of the ranges of  $N$  for the three types of equilibrium are found by noting that the profit (or price) functions are continuous in  $N$ . Thus, for example, at the boundary  $N^{mk}$  between the monopoly and kink equilibria  $\pi^m(N^{mk}) = \pi^k(N^{mk})$ . Solving gives  $N^{mk} = 4\alpha tK$ .

Effect on endogenous variables												
Equilibrium type	Price			Quality			List size			Number of GPs		
	$p$			$q$			$D$			$N$		
	m	k	c	m	k	c	m	k	c	m	k	c
Increase in												
cost of quality $\alpha$	-	-	-	-	-	-	-	+	0	-	-	0
distance cost $t$	0	+	+	0	0	0	-	+	-	-	-	+
market area $K$	0	+	+	0	0	0	-	+	-	-	-	+
population $M$	0	+	-	0	0	0	+	+	+	+	+	+
area, $M/K$ constant	0	-	+	0	0	0	0	+	+	0	+	+

Table 3: Comparative statics of private capitation system

even in the kink case where the number of firms affects the price and  $\alpha$  also alters gross profit and thus the number of firms. (See the appendix.)

This is in contrast with the usual result for a firm facing a downward sloping demand curve where an increase in marginal cost will always increase price. Here an increase in  $\alpha$  raises the marginal cost of quality and thereby reduces quality. The reduction in quality reduces demand and marginal revenue at given  $p$  and leads to a reduction in the fee.

Equilibrium list size increases with the population  $M$  in all cases even after allowing for the induced increase in  $N$ . List size falls with the market length  $K$  and patient distance cost  $t$  under monopolistic competition and monopoly but increases with  $K$  and  $t$  in the kink case. Inspection of Table 3 shows that population  $M$  and market “length”  $K$  have effects which are in addition to the changes they induce in the population density  $M/K$ . Increases in  $M$  and decreases in  $K$ , both of which increase population density, sometimes work in the same direction, for example in increasing list size or reducing price in the competitive case. However, the effect of an increase in population density is to increase the number of doctors if the increase in density follows from an increase in population but to reduce it if the increase in density follows from a reduction in market length. This suggests that empirical studies need to be based on careful specification of the measures of variables intended to capture demand side characteristics.

Finally, note that *there is more likely to be a monopoly equilibrium, and thus gaps in the market with some patients not registering, the greater are distance costs, market area and the GP cost parameter  $\alpha$* . As these increase practice gross profit is reduced and the boundary between monopoly and kink equilibria is shifted to the right. Hence the supply curve is more likely to intersect  $\pi(N)$  in its monopoly segment.

## 2.6 Geographical distribution of GPs

Although this is not the main focus of the current paper we note that the model can easily be extended to examine the geographical distribution of GPs, list sizes and quality. We can consider different regions as Salopian circles with different parameter (population, market size, cost conditions etc). If doctors can locate in any region equilibrium will require that gross profit is equalised across regions. We can then, for example, use the model to examine the cross-sectional variation in doctors' fees and the doctor population ratio (Pauly and Satterthwaite, 1981; Satterthwaite, 1985; Harris, 1985). In the competitive case Table 2 suggests that areas with a greater number of doctors have lower capitation fees. However we must take account of the endogeneity of the number of doctors and the reason why the number of doctors is higher in one area than another. Thus suppose that  $K$  varies across areas. Areas with larger  $K$  will, *ceteris paribus*, have greater higher profits and thus attract more doctors. The increase in the number of doctors reduces the capitation fee but the increase in  $K$  increases it. The Appendix shows that the overall effect is to increase  $p$ .<sup>10</sup> Hence, as Table 3 indicates, there is a positive cross-sectional correlation of the number of doctors and the capitation fees. On the other hand if different regions are of the same size but differ in population  $M$  or distance cost  $t$  there will be a negative correlation between the number of doctors and capitation fees.

The lesson to be drawn is that cross-sectional analysis of the determinants of capitation fees needs to be careful in allowing for the endogeneity of fees and the number of doctors and in capturing the important exogenous variables.

## 3 Private capitation with a self-regulating profession

The medical profession is often able to persuade policy makers that it should be permitted to regulate entry, capitation fees or quality. The self-regulatory equilibrium depends on what the profession can control and on its objectives. There are seven possible subsets of  $\{p, q, N\}$  and a large number of possible objective functions.<sup>11</sup> We examine only two possible models in both of which the profession seeks to maximize total gross profits  $N\pi$ .

### 3.1 Entry control

Consider first the case in which the profession only has the ability to control entry and assume that the free entry unregulated equilibrium would be competitive. In terms of Figure 2 the profession seeks a point on the  $\pi(N)$  curve where  $\pi(N)N$  is maximized. The profession's indifference curves in  $(\pi, N)$  space are rectangular hyperbola. Clearly the solution will not lie on the monopoly segment of the profit function since over this range additional entrants increase total profit at a constant rate. The solution must lie on the downward sloping portion of the profit curve. By comparing the slope of the profit curve over the kink and competitive segments with the slope of the indifference curve we see that *the total profit maximising number of practices is at the boundary between the kink and monopolistic competition equilibrium ranges.* (See the Appendix.)

The self-regulating profession restricts the number of practices, driving up the capitation fee and practice profit until the market is just covered. The marginal patient gets a

---

<sup>10</sup>Intuitively:  $N$  increases less than proportionally with  $K$  to equalise profit, hence price, which varies proportionally with the ratio  $K/N$ , must increase with  $K$ . See the appendix.

<sup>11</sup>See the literature on the objectives of trade unions, for example Booth (1994). And this ignores the possibility that the profession has quasi-altruistic concerns with patient wellbeing.

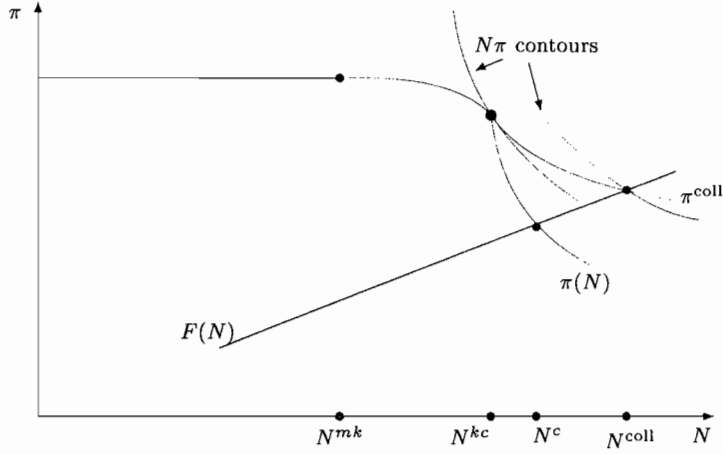


Figure 2: Self-regulating profession

zero surplus and is indifferent between joining a list and not joining.

### 3.2 Collusion over fee and quality

Assume now that, as well as restricting entry, the members of the profession can collude, rather than compete, on quality and capitation fee. For any given  $N$  they aim to maximise practice profit. They do so by choosing a combination of fee and quality so that the marginal patient is left with no surplus. Suppose not and the marginal patient has a positive surplus. If all practices increase their fees no individual practice will lose market share because the difference between practice fees is maintained. Since demand is unaltered and the fee has increased all practices are better off.

The collusive  $(p, q)$  must be at the kink and satisfy  $2(q-p)/t = K/N$  or  $q = p + tK/2N$ . Substituting this into the profit function  $D(p - \alpha q^2) = M(p - \alpha q^2)/N$  and maximising with respect to the fee gives  $p = \frac{1}{2\alpha} - \frac{tK}{2N}$  and  $q = 1/2\alpha$ . Collusion enables the profession to establish the kink fee and quality with numbers of doctors which would otherwise lead to the competitive capitation fee. For levels of  $N$  which sustain the monopoly solution there is no need for collusion since there is no competition between practices.

The feasible profit function under full collusion coincides with the non-collusive monopoly profit and kink profit function but lies above the non-collusive competitive profit function for levels of  $N$  which would lead to competition in the absence of collusion. Figure 2 illustrates.

We can show (see Appendix) that total gross profit  $N\pi^{\text{coll}}$  is increasing in  $N$ . The collusive solution is at the point where the collusive profit function cuts the reservation wage function. This is true whether the profession can control entry (in which case it chooses the largest number of practices consistent with the marginal practice not making a profit less than its reservation wage) or whether it has no control of entry.

These two examples show that the extent of the profession's control over the variables determining practice profit affect not only the profit but also the number of practices. *If practices can collude over capitation fee and quality there will be more practices and smaller list sizes than if the profession can only control entry. If the unregulated non-collusive equilibrium is competitive, then in the former case collusion increases the number of practices whilst in the latter case the number of practices is reduced.*

## 4 Welfare and private capitation regimes

Welfare is the unweighted<sup>12</sup> sum of consumer surplus, practice rents and taxpayer surplus or patients' willingness to pay less the firms' production costs, including their reservation wages:

$$\begin{aligned} W(p, q, N) &= N\delta \frac{M}{K} \left( q - \frac{t\delta}{4} - \alpha q^2 \right) - \int_0^N F(\tilde{N}) d\tilde{N} \\ &= S(p, q, N) - \int_0^N F(\tilde{N}) d\tilde{N} \end{aligned} \quad (12)$$

where  $\delta$  is the segment of the market served by one firm. When there are gaps in coverage  $\delta = 2(q - p)/t$  but when all patients are registered with some GP  $\delta = K/N$ .  $S$  is the gross surplus: the sum of consumer surplus, taxpayer surplus and gross practice profits. The expression in parenthesis in the second equation in (12) is welfare (excluding gp reservation wages) per patient. Every patient on a list has quality of  $q$ , which costs  $\alpha q^2$  per patient, and the average patient distance cost is  $t\delta/4$ .

The first best allocation results from maximizing  $W$  with respect to price, quality and the number of firms. It could be interpreted as the allocation resulting from an ideal public system in which the planner can levy lump sum taxes to finance losses, fix prices paid by patients and received by GPs, control quality directly and control the number of GPs. As we will see the first best may have gaps in coverage, so that some patients are not on the list of any practice or it may require that all patients are on some list. We investigate the circumstances in which these types of first best apply. We consider whether the unregulated Nash equilibria, of any type, coincide with the first best allocation. We also compare the equilibria with a second best optimal allocation in which the regulator can only control entry, leaving quality and price to be determined by the market.

### 4.1 Welfare with incomplete coverage

Consider first the situation in which each practice has  $\delta = (q - p)/t$ . The first order conditions on  $p$  and  $q$  for maximising welfare at given  $N$  imply  $q^* = 1/2\alpha$  and  $p^* = \alpha q^2 = 1/4\alpha$ . Thus *in a monopoly equilibrium the quality chosen by GPs is first best optimal*. The reason is that the practice can appropriate all the patients gains from improving the quality of service by increasing the fee. Hence it will internalise the gains from improved quality and weigh them against the cost of improved quality. Patients are exploited but efficiently. As we will see, this is also true in the other market equilibria.

However, *the monopoly capitation fee  $3/8\alpha$  (see Table 1) is too high*. With  $q = 1/2\alpha$ , the net social benefit from admitting to the list an additional patient at distance  $\delta/2$  and charging him  $p$  is

$$q - \frac{t\delta}{2} - \alpha q^2 = q - \frac{2(q - p)t}{t} \frac{t}{2} - \alpha q^2 = p - \alpha q^2 = p - \frac{1}{4\alpha}$$

<sup>12</sup>If different weights are given to consumers, taxpayers and firms the welfare analysis, which is intended to focus on issues of quality and horizontal product differentiation, is dominated by the implications of the distributional value judgements. In the competitive and kinked equilibria all patients register. Thus changing the fee has no efficiency implications and it can be used solely to redistribute, leading to uninteresting corner solutions. For example if taxpayer surplus had a greater weight than practice rent or patient surplus a tax should be levied on the capitation fee and should be raised until the market is only just covered.

An additional patient registered increases social cost by  $\alpha q^2 = 1/4\alpha$ . Patients should be charged the marginal cost of their joining a list, rather than not joining any list. The monopoly equilibrium with  $p^m = 3/8\alpha > 1/4\alpha$  results in too few patients joining lists because the monopoly fee exceeds marginal cost.

Now examine the number of GPs. Substituting the first best price  $p^* = \alpha q^2 = 1/4\alpha$  and quality  $q^* = 1/2\alpha$  gives the marginal gross social value of an additional practice as

$$S_N^* = S_N(p^*, q^*, N) = \frac{M}{16\alpha^2 tK} = 2\pi^m(N) \quad (13)$$

where  $\pi^m$  is the monopoly equilibrium gross profit. Since the private reward for entry  $\pi^m$  is less than the social gain  $S_N^*$  the monopoly equilibrium number of firms is smaller than the first best number.

Indeed, the monopoly equilibrium number of practices is too small compared with the second best optimal number which would be chosen by a regulator who had no control over quality and price. Substituting the monopoly equilibrium fee and quality into the gross surplus function gives the marginal gross value of a practice as

$$S_N(p^m, q^m, N) = \frac{3M}{64\alpha^2 tK} = \frac{3}{2}\pi^m(N) \quad (14)$$

and so monopoly provides a suboptimal incentive for entry.

## 4.2 Welfare in the kink and competitive cases

With all consumers registered with practices  $\delta = K/N$  and

$$W(p, q, N) = M \left( q - \frac{tK}{4N} - \alpha q^2 \right) - \int_0^N F(\tilde{N}) d\tilde{N} \quad (15)$$

Differentiating we get the socially optimal quality  $q^* = 1/2\alpha$  and, referring to Table 2, we see that the competitive and kink equilibria have first best quality.

Unlike the incomplete coverage case, the capitation fee does not affect the number of patients registered. Given the distributional value judgements, the level of the fee does not affect welfare. The only constraint on the fee is that the market is covered:  $p \leq q - tK/2N = 1/2\alpha - tK/2N$ . Hence the kink and monopolistic competition fees are first best optimal.

Because the equilibrium price and quality are socially optimal there is no difference between the first best and second best number of GPs, unlike the monopoly case where the price is too high. The marginal gross value of an additional practice at competitive or the kink equilibria is

$$S_N(p^*, q^*, N) = S_N(p^j, q^j, N) = \frac{tKM}{4(N^j)^2} < \pi^j(N), \quad j = c, k \quad (16)$$

The competitive and kink equilibria have too many firms because practice profit exceeds the marginal social value of an additional practice. The social value of an additional practice is less than in the monopoly case. Once the market is covered additional practices do not supply additional patients with care.

Figure 3 illustrates the relationship between the equilibria (determined by the intersection of  $\pi(N)$  with the inverse supply function  $F(N)$ ) and the first and second best social

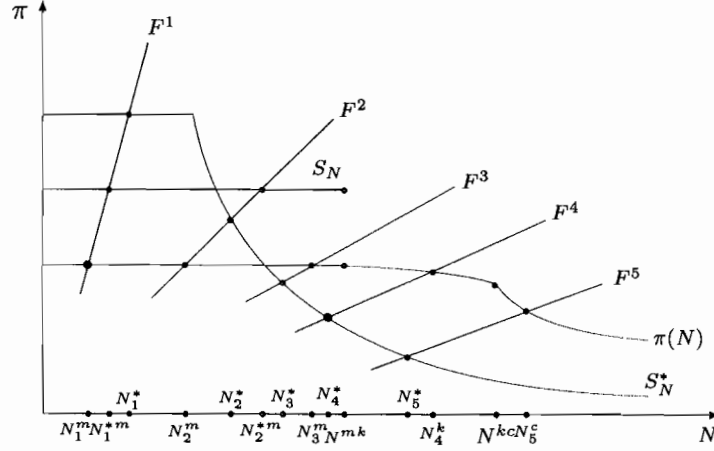


Figure 3: Welfare properties of equilibria

optimum (determined by the intersection of the marginal first best and second best social value curves  $S_N^*$  and  $S_N$  with  $F(N)$ ).<sup>13</sup>

In Figure 3 a variety of supply functions are used to illustrate the possible relationships between number of practices in the first best  $N^*$ , second best  $N^{*j}$ , and market equilibria  $N^j$ . Summarising:

$$\begin{aligned}
 F^1 &\implies N^* > N^{*m} > N^m \\
 F^2 &\implies N^{*m} > N^* > N^m \\
 F^3 &\implies N^{*m} > N^m > N^* \\
 F^4 &\implies N^k > N^{*k} = N^* \\
 F^5 &\implies N^c > N^{*c} = N^*
 \end{aligned}$$

When the market equilibrium has too many GPs policy is apparently very simple: restrict the number of firms to  $N^*$  either by direct entry control or by a tax on each practice. This leads to an increase in the capitation fee (see (9) and increases practice profits. Patients are worse off because they pay a higher capitation fee but this is exactly offset, in welfare terms, by the increased profits of the firms. Since quality does not vary with the number of firms the equilibrium quality continues to be socially optimal.<sup>14</sup>

In the case of  $F^5$  in Figure 3 a self-regulating profession which restricts entry to the boundary between the kink and competitive cases increases welfare because it reduces the

<sup>13</sup>The maximum number of practices compatible with incomplete coverage is smaller in the first best case ( $2\alpha tK$ ) than under monopoly ( $4\alpha tK$ ) because the first best price is less than the monopoly price. The second best marginal gross social value curve  $S_N(p^j, q^j, N)$  is discontinuous at the boundary between the incomplete and complete coverage cases.

<sup>14</sup>There is an interesting complication with direct control of entry: since there will be an excess supply of GPs the regulator will need to ration entry. Efficient rationing requires that she knows the reservation wage of individual entrants. If she does not some GPs admitted to practice will have higher reservation wages than some of those excluded. This is inefficient. One method of solving this problem is to auction off the right to practice. This is equivalent to an appropriately chosen tax.

See Mumy and Hanke (1975) for an analogous discussion of the implications of imperfect rationing of demand for optimal investment levels.

number practices to somewhere close to the first best level. However, if  $F^5$  was shifted further to the right the self-regulating profession could make reduce welfare because it would lead to too large a reduction in the number of practices.

## 5 A public capitation system

In a public system, like the NHS, GPs are paid a tax financed capitation fee  $p$  per patient on their list. Patients pay nothing when joining a practice list. There are again three types of equilibria but for reasons of space we here only consider the competitive case. The practice list is found from (3) but with the price paid by patients set to zero:  $D_i = D_i(0, q_i; 0, 0, q_{i+1}, q_{i-1})$ .

Practices can compete only via quality. Substituting the demand function into the first order condition on quality (7) and assuming symmetry we get the competitive Nash equilibrium quality for each firm as

$$q(p, N) = \left[ \left( \frac{tK}{N} \right)^2 + \frac{p}{\alpha} \right]^{\frac{1}{2}} - \frac{tK}{N} = \beta(p, N) - \frac{tK}{N} \quad (17)$$

Although  $N$  reduces the first term and increases the second we show in the Appendix that *quality is increasing in the number of practices under public capitation*. Increasing the number practices increases competition for patients and since practices cannot compete on price they have greater incentives to increase quality to attract additional patients.

Even when quality is not observable by the regulator she has an effective instrument for controlling it since *quality is increasing in the capitation fee in a public system*. Raising the fee makes marginal patients more valuable and induces practices to compete for them by raising quality.<sup>15</sup>

By setting the public capitation fee equal to the private market equilibrium fee  $\frac{1}{4\alpha} + \frac{tK}{N}$  the regulator can induce firms to supply the optimal quality  $1/2\alpha$ . The regulator mimics the outcome of unregulated private market to induce optimal quality and does not need to monitor the quality level of any practice.

Unfortunately, with free entry the equilibrium number of firms would be too large. The planner has two targets (optimal quality and number of firms) and requires two instruments to achieve a first best. *The first best optimal capitation remuneration scheme has a negative lump sum component to control the number of practices and a capitation component to control quality*. An alternative to such a two part non-linear scheme is a proportional capitation fee coupled with direct control of entry.

If only the fee can be controlled and the regulator is restricted to a linear scheme, the optimal second best fee reflects the influence of the fee on both numbers of firms and quality. The second best public capitation fee is less than required to induce optimal quality but higher than required to induce the optimal number of firms.

## 6 Public choice models

Instead comparing the welfare properties of the private capitation outcome with a public capitation scheme run by a welfare maximising regulator it may be more appropriate to adopt a comparative institutional approach based on a positive model of a public capitation

---

<sup>15</sup>The equilibrium qualities in the monopoly and kink cases are  $q^m = p/3 + \gamma/3\alpha$  (where  $\gamma = [(p\alpha)^2 + 3p\alpha]^{1/2}$ ) and  $q^k = p + tK/2N$ . Increases in the capitation fee increase quality in these cases as well.

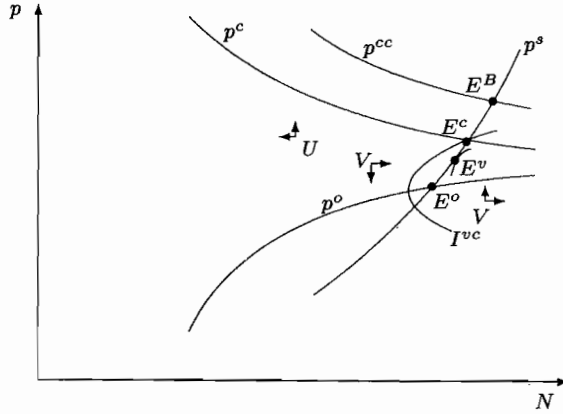


Figure 4: Voting and bargaining

system (Demsetz, 1969). Suppose that the public capitation fee is decided by majority voting, rather than by a welfare maximizing regulator. Assume for simplicity that the public system will be in equilibrium in the competitive range. Each voter has utility  $q - td - p$  where  $p$  is paid as poll tax rather than directly as a fee for joining the list of a particular practice. Given our assumption of full coverage this is equivalent to a model in which the patient pays the fee directly to her GP and the fee is determined by voting of voter-patients. We assume that there are so few GPs that their votes have no effect on the outcome.

Voters realise that a higher fee leads to a higher quality via (17):  $q = q(p, N)$ . We assume that policy is chosen to maximise the voter objective function

$$V = V(p, N) = q(p, N) - p - \frac{tk}{4N}. \quad (18)$$

We can interpret (18) by supposing that voters, when choosing  $p, N$ , do not know where additional practices will be located. They therefore evaluate the effect of the increases in the number of practices by their effect on the expected distance cost.

The voters' choice of  $p, N$  is constrained by the supply function of practices. Given  $p$  each practice chooses its quality taking the number of practices as given and the resulting profit is

$$\pi(p, N) = [p - \alpha q(p, N)]^2 \frac{M}{N} = \frac{2M\alpha t K q(p, N)}{N^2} \quad (19)$$

where we get the last expression by using (17). The voters' choice of  $p, N$  is constrained by  $\pi(p, N) \geq F(N)$ . Writing this as an equality yields the inverse supply function for practices  $p^s(N)$  shown in Figure 4.

### 6.1 Myopic voters

If voters regard the number of practices as fixed and are concerned only about the effect of the fee on the quality of care, the fee chosen for given  $N$  will satisfy

$$\frac{\partial q(p, N)}{\partial p} - 1 = 0 \quad (20)$$

where  $q(p, N)$  is given by (17). Solving for the myopic voters' optimal fee gives:

$$p^o(N) = \frac{1}{4\alpha} - \frac{\alpha(Kt)^2}{N^2} < \frac{1}{4\alpha} + \frac{tK}{N} \quad (21)$$

and the resulting quality level is

$$q^0(N) = \frac{1}{2\alpha} - \frac{tK}{N}. \quad (22)$$

If there is free entry equilibrium will be established by adjustments in the number of practices until  $\pi(p^o(N), N) = F(N)$  or  $p^o(N) = p^s(N)$ .

In Figure 4 the myopic voting equilibrium is at  $E^0$  where  $p^o$  and  $p^s$  intersect.<sup>16</sup> The curve  $p^c$  plots the competitive equilibrium fee (9) and the competitive equilibrium is at  $E^c$ . We see that *the myopic voting equilibrium has a lower price, lower quality and fewer practices than the unregulated competitive equilibrium.*

In terms of the welfare function (12) which takes account of the welfare of patients and doctors,  $p^c$  induces the welfare maximising quality at any given  $N$ . Thus *the myopic voting equilibrium has too low a level of quality but since the competitive equilibrium has too many practices the overall effect on welfare is ambiguous.*

The voters are concerned only with the effect of the fee on quality and place no weight on the profit of the GPs. Because some of the gains from higher fees accrue to firms in the form of higher profit the voters' marginal gain from higher fees is less than the social gain.

## 6.2 Non-myopic voters

When voters take account of the number of practices  $p$  and  $N$  are chosen to maximise  $V(p, N)$  subject to the practice supply constraint. The first order conditions are<sup>17</sup>

$$q_p(p, N) - 1 + \lambda\pi_p(p, N) = 0 \quad (23)$$

$$q_N(p, N) + \frac{tK}{4N^2} + \lambda[\pi_N(p, N)F'(N)] \quad (24)$$

$$\pi(p, N) - F(N) \geq 0, \quad \lambda \geq 0, \quad \lambda[\pi(p, N) - F(N)] = 0 \quad (25)$$

Since an increase in the number of practices always increases quality the constraint on the number of practices binds and the solution  $E^v$  is where the voter indifference curve  $I^v$  is tangent to the supply function in Figure 4.

We show in the Appendix that the voter indifference curve  $I^{vc}$  through the unregulated competitive equilibrium  $E^c$  is flatter than the supply curve at this point. Thus *the voting equilibrium has a lower fee, lower quality and fewer practices than the competitive equilibrium.* The welfare comparison is similar to the case in which voters are myopic: *quality is less than the welfare maximising level but since there are too many practices in the competitive equilibrium the overall effect on welfare is ambiguous.*

<sup>16</sup>An equilibrium in which there is full coverage and practices compete via quality, at a given fee, can only occur for combinations of  $N$  and  $p$  satisfying  $2M[q(p, N) - p]/tK > M/N$ . In what follows we assume that supply function of practices lies, in part, in this region.

<sup>17</sup>The assumptions about preferences, technology and the supply function ensure that these are also sufficient if we further assume that the supply of practices intersects  $p^o$  in the region in which there is an equilibrium in which practices compete in quality at the given price. (See footnote 16.)

### 6.3 Voter-Union Bargain

In the presence of a powerful professional union we can depict policy as determined by the outcome of bargaining between the union and the voters. As before suppose that the union's objective function is total gross practice profits:

$$U(p, N) = N\pi(p, N) = 2\alpha t K M q(p, N)/N. \quad (26)$$

Consider the cooperative Nash Bargaining Solution (NBS) where the policies  $p, N$  are chosen to maximize the Nash product

$$B = V(p, N)U(p, N) \quad (27)$$

subject to the supply constraint that  $\pi(p, N) \geq F(N)$ .

The NBS lies on the contract curve defined by the tangency of the voter and union indifference curves.  $U$  is clearly increasing in  $p$  since  $q_p = 1/2\alpha\beta > 0$ . We can also show that the union prefers a smaller number of practices given the capitation fee:<sup>18</sup>

$$U_N = \frac{M2\alpha t K}{N^2}(Nq_N - q) = -\frac{q^2 M2\alpha t K}{N^2\beta} < 0 \quad (28)$$

The union's indifference curves are positively sloped, as in Figure 4.

As we have seen voters always prefer more practices to fewer so that, since their indifference curves must be positively sloped along the contract curve the curve must lie above the  $p^0$  locus where  $V_p = 0$ . In fact we show in the Appendix that *the contract curve locus  $p^{cc}$  lies above the competitive locus  $p^c$  and is negatively sloped*. Hence *the NBS has a higher price and quality than the competitive equilibrium*. It is not surprising that the profession's bargaining power enables it to increase the capitation fee above the level which would be chosen by voters who are constrained only by the practice supply function. The fact that the union-voter bargaining increases the capitation fee above the competitive level is less obvious.

Given that the positive slope of the supply curve  $p^s$  and the negative slope of the contract curve  $p^{cc}$  we see that *if the supply constraint binds the NBS has a larger number of practices than the competitive equilibrium*.<sup>19</sup>

Recalling that the competitive equilibrium has optimal quality but too many firms, *if the supply constraint binds welfare is lower at the NBS than at the competitive equilibrium since quality is greater and the number of firms higher*. If the supply constraint does not bind the welfare implications are ambiguous: the NBS has greater than optimal quality but it may have a smaller number of firms than the competitive equilibrium.

## 7 Conclusions

The quality and accessibility of primary care are major concerns of policy makers. This paper is an initial attempt to examine the implications of public and private capitation

<sup>18</sup>Use the fact that  $q_N = tKq/N^2\beta$  to get

$$Nq_N - q = -\frac{q^2}{N^2\beta} < 0$$

<sup>19</sup>The NBS satisfies independence of irrelevant alternatives (Osborne and Rubinstein, 1990) so that shifts in the supply function affect the NBS only if they render the initial NBS infeasible.

systems for quality of care and the number of practices. At least some of the aspects of quality with which policy makers are concerned are observable by patients. The paper has shown that in this case a private capitation market will lead to an efficient choice of quality by doctors without the need for regulation. Even when there are gaps in the coverage of patients, so that a practice is not affected by the price and quality of other practices, it will choose an efficient level of quality. The reason is that the practice can recoup all the gains to patients from higher quality by raising its capitation fee. The practice's monopoly power enables it to exploit patients but it will do so efficiently. The same is true when practices are affected by the price and quality of rivals. Because patients observe and are willing to pay more for better quality, practices internalise both the costs and the benefits of higher quality.

An increase in the number of practices is of social benefit since it improves patient access but it is also costly because doctors have an opportunity cost (even if only in working in other sectors of the health service). Under monopoly the social gain from an additional practice is greater than when the market is covered. If there are gaps in coverage, as there are under monopoly, an additional practice results in some additional patients joining a list as well as reducing access costs for those already on a list. When the market is covered only the latter benefit arises. Thus the marginal social benefit from an increase in the number of practices is greater under monopoly than when the market is covered. Practice entry decisions are determined by profit rather than marginal social benefit. Under monopoly profit is less than marginal benefit and there are too few practices. When the market is covered profit exceeds marginal benefit and there are too many practices. The policy maker will therefore have to pay a subsidy to practices when there is incomplete coverage and tax or otherwise constrain entry when the market is covered.

Under a public capitation regime the twin policy targets of welfare maximizing quality and accessibility require two policy instruments. The policy maker can induce practices to choose the appropriate quality level by a suitably chosen capitation fee. Practices will then compete for patients by increasing quality. The number of practices can be controlled by a lump sum tax or subsidy unrelated to number of patients.

A proper comparison of private and public capitation regimes must recognise that in many health care systems doctors' professional organisations exercise some influence over policy, usually with regard to numbers of practices. Unfortunately welfare comparison of public and private systems in the presence of such organisations are often ambiguous in the absence of very finely detailed specification of their preferences.

Draft: 2e; Sept 30 1996, Printed: December 2, 1996

## References

- Booth, A. *The Economics of the Trade Union*, Cambridge University Press, 1995.
- Demsetz, H. "Information and efficiency: another viewpoint", *Journal of Law and Economics*, 1969, 12, 1-22.
- Economides, N. "Symmetric equilibrium existence and optimality in a differentiated product market", *Journal of Economic Theory*, 1989, 47, 178-184.
- Economides, N. "Quality variations in the circular model of variety-differentiated products", *Regional Science and Urban Economics*, 1993, 23, 235-257.
- Harris, J. "Comment: competition and equilibrium as driving force in the health services sector", in Newman, R. P. (ed.), *Managing the Service Economy: Perspectives and Problems*, 1985, Oxford University Press.
- Mumy, G. E. and Hanke, S. H. "Public investment criteria for underpriced public products", *American Economic Review*, 65, 1975, 712-721.
- Norman, G. "Monopolistic competition: some extensions from spatial competition", *Regional Science and Urban Economics*, 1989, 19, 31-53.
- Osborne, M. J. and Rubinstein, A. *Bargaining and Markets*, Academic Press, 1990.
- Pauly, M. V. and Satterthwaite, M. A. "The pricing of primary care physicians' services: a test of the role of consumer information", *Bell Journal of Economics*, 1981, 12, 488-506.
- Salop, S. C. "Monopolistic competition with outside goods", *Bell Journal of Economics*, 1979, 10, 141-156.
- Satterthwaite, M. A. "Competition and equilibrium as a driving force in the health services sector", in Newman, R. P. (ed.), *Managing the Service Economy: Perspectives and Problems*, 1985, Oxford University Press.
- Shelden, T. M. "A model of capitation", *Journal of Health Economics*, 9, 1990, 397-409.

## Appendix: proofs

1. *Effect of  $\alpha$  on price in the kink case.* Assume, to give the greatest possible offsetting effect via  $N$  that the supply function is horizontal and  $dN/d\alpha = -\pi_a/\pi_N$ . Then

$$\frac{dp}{d\alpha} = -\frac{1}{2\alpha^2} + \frac{tk}{2\alpha} \frac{1}{(4\alpha tK - N)}$$

which is negative since, if the equilibrium is at the kink,  $N \in [4\alpha tK, 6\alpha tK]$ .

2. *Effect of  $K$  on price in the competitive case.* The supply of practices is determined by  $\pi(N, K) - F(N) = 0$  so that

$$\frac{dN}{dK} = -\frac{\pi_K}{\pi_N - F'} = \frac{\frac{tM}{N^2}}{\frac{2tKM}{N^3} + F'}$$

which is positive. From Table 2  $dp/dK$  is positive since it has the same sign as

$$\begin{aligned} N - K \frac{dN}{dp} &= N - \frac{\frac{tKM}{N^2}}{\frac{2tKM}{N^3} + F'} \\ &= \left( \frac{2tKM}{N^2} + F'N - \frac{tKM}{N^2} \right) \left( \frac{2tKM}{N^3 + F'} \right)^{-1} > 0 \end{aligned}$$

3. *Slopes of union indifference curves and the profit function.* At  $N = 6\alpha tK$  we have

$$\frac{d\pi^k}{dN} = -\frac{tKM}{2N^3} > -\frac{tKM}{N^3} = \frac{d\pi}{dN} \Big|_{\pi_N} = -\frac{\pi}{N} > \frac{d\pi^c}{dN} = -\frac{2tKM}{N^3}.$$

4. *Slopes of the collusive kink profit function and union objective function.* The collusive kink profit function  $\pi^{\text{coll}} = M/4\alpha N - tKM/2N^2$  has slope  $tKM/N^3 - M/4\alpha N^2$  and is flatter (absolutely) than the contour of  $\pi N$  which has slope  $-\pi/N = tKM/2N^3 - M/4\alpha N^2$ .

5. *Effects of  $N$  and  $p$  on quality under public capitation.* Using (17)

$$q_N = -\frac{(tK)^2}{N^3} \beta^{-1} + \frac{tK}{N^2} = \frac{tK}{N^2} \left( -\frac{tK}{N} \beta^{-1} + 1 \right) = \frac{tK}{N^2} q \beta^{-1} > 0.$$

$$q_p = \frac{1}{2\alpha\beta} > 0$$

6. *Slope of the supply function.* The equation defining the supply function is  $\pi(p, N) - F(N) = 0$ . Using (19), multiplying through by  $N$  gives  $2\alpha M tK q(p, N) - F(N)N = 0$  and so

$$\frac{dp^s}{dN} = \frac{2\alpha M tK q_N + N F' + F}{(1 - 2\alpha q q_p) M}$$

Now, from part 5 of the appendix,  $1 - 2\alpha q q_p = 1 - q/\beta = tK/N\beta > 0$  and so the supply curve is positively sloped in  $(p, N)$  space.

7. *Relative slopes of  $I^v$  and  $p^s$  at  $E^c$ .* We have

$$\begin{aligned} \frac{dp}{dN} \Big|_{I^v} &= \frac{q_N + \frac{tK}{4N^2}}{q_p - 1} \begin{matrix} \geq \\ < \end{matrix} \frac{2\alpha q q_N M + NF' + F}{(1 - 2\alpha q q_p)M} = \frac{dp^s}{dN} \\ \iff q_N M(1 - 2\alpha q) + \frac{tKM}{4N^2}(1 - 2\alpha q q_p) &\begin{matrix} \geq \\ < \end{matrix} (1 - q_p)(NF' + F) \\ \iff \frac{F}{4}(1 - 2\alpha q q_p) &\begin{matrix} \geq \\ < \end{matrix} (NF' + F)(1 - q_p) \end{aligned}$$

where the last step uses the fact that  $q = 1/2\alpha$  and  $F = \pi = tKM/N^2$  at the competitive equilibrium  $E^c$ . Since  $q_p - 1 = V_p < 0$  above the  $p^o$  locus

$$\begin{aligned} F(1 - 2\alpha q q_p) &< (NF' + F)(1 - q_p) \\ \iff F q_p(1 - 2\alpha q) &< NF'(1 - q_p) \\ \iff 0 &< NF'(1 - q_p) \end{aligned}$$

and we have established that  $I^v$  is flatter than  $p^s$  at  $E^c$ .

8. *Properties of the contract curve.* The contract curve is defined by

$$\frac{V_N}{V_p} = \frac{q_N + \frac{tK}{4N^2}}{q_p - 1} = \frac{Nq_N - q}{Nq_p} = \frac{U_N}{U_p}$$

Multiply through to remove the denominators, cancel terms, recall from (17) that  $q_p = 1/2\alpha\beta$ ,  $Nq_N - q = -q^2/\beta = -q^2 q_p 2\alpha$ , divide by  $q_p$  and rearrange to get

$$2\alpha q^2 - q - \frac{tK}{4N} = 0$$

Solve for the quality generated by the  $p, N$  on the contract curve as

$$q^{cc} = \frac{1 + \left(1 + \frac{2\alpha tK}{N}\right)^{\frac{1}{2}}}{4\alpha} > \frac{1}{2\alpha}$$

Since  $q^{cc} > 1/2\alpha$ , the contract curve price  $p^{cc} > p^c$  for given  $N$ . Note also that  $q^{cc}$  declines with  $N$  and so therefore must the  $p^{cc}$  which induces the individual GPs to choose  $q^{cc}$ . Hence the contract curve lies above the competitive locus  $p^c$  and is negatively sloped.