

# **Yardstick Competition for Multi-product Hospitals**

## **An Analysis of the Proposed Dutch Yardstick Mechanism**

Agrell, Per J.; Bogetoft, Peter; Halbersma, Rein; Mikkers, Misja C.

### *Document Version*

Final published version

### *Publication date:*

2007

### *License*

CC BY-NC-ND

### *Citation for published version (APA):*

Agrell, P. J., Bogetoft, P., Halbersma, R., & Mikkers, M. C. (2007). *Yardstick Competition for Multi-product Hospitals: An Analysis of the Proposed Dutch Yardstick Mechanism*. Nederlandse Zorgautoriteit. Research Paper Series No. 2007-01

[Link to publication in CBS Research Portal](#)

### **General rights**

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

### **Take down policy**

If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 18. Jun. 2025





Nederlandse  
Zorgautoriteit

# **Yardstick competition for multi-product hospitals**

An analysis of the proposed  
Dutch yardstick mechanism

2007

01

# **Yardstick competition for multi-product hospitals**

An analysis of the proposed Dutch yardstick mechanism

Per J. Agrell, Peter Bogetoft  
SUMICSID AB  
Rein Halbersma, Misja C. Mikkers  
NZa (Dutch Healthcare Authority)



# Table of Contents

Preface	5
Samenvatting in het Nederlands	7
Abstract	11
<b>1. Introduction</b>	<b>13</b>
1.1 Background	13
1.2 Objectives	13
1.3 Contribution	13
1.4 Outline	14
<b>2. Health care in the Netherlands</b>	<b>15</b>
2.1 The health insurance market	15
2.2 Hospital budgeting	16
2.3 Liberalising hospital-insurer contracting	17
2.4 The need for transitory regulation	17
<b>3. Base model</b>	<b>19</b>
3.1 Economic entities	19
3.2 Production, costs, and revenues	20
3.3 Cost based ex ante regulation	21
3.4 Objectives	23
3.5 Revenue based yardsticks	28
3.6 Bargaining and competition	29
<b>4. Analysis</b>	<b>35</b>
4.1 Markets, bargaining and regulation	35
4.2 Cost efficiency effects	37
4.3 Volume effects	37
4.4 Mix effects	41
4.5 Reallocation effects	42
4.6 Demand inducement and moral hazard	43
4.7 Convergence	44
4.8 Robustness to diverse firm preferences	45
4.9 Robustness to heterogeneous cost functions	47
4.10 Information requirements	47
4.11 Discussion	49

<b>5. Quality and innovation incentives</b>	<b>51</b>
5.1 Incentives for innovation	51
5.2 Quality provision incentives	54
5.3 Steering mechanisms	59
5.4 Summary on R&D and quality	65
<b>6. Conclusions</b>	<b>67</b>
6.1 References	70
<b>7. Appendix A: Numerical estimates of distortions</b>	<b>73</b>

## **Preface**

The paper 'Yardstick competition for multi-product hospitals', is the first paper in the Research Paper Series by the Dutch Healthcare Authority (NZA). The Research Paper Series aims at the enhancement of the knowledge en expertise in the regulation of and competition in health care markets. The papers in this series are written by invited authors and/or NZa staff.

In 2005, the segment of uncomplicated, elective outpatient care has been deregulated. Prices in this segment are subject to bargaining between insurers and hospitals. However, the major part of hospital production is currently still regulated a budget system. The Ministry of Health is proposing to deregulate the remainder of elective hospital care (including inpatient elective care). Within the current competitive domain, insurers are unable to use their countervailing power in reducing inefficiency and profit margins and moving bargaining outcomes toward the competitive equilibrium. Therefore, regulation will be introduced as a transitory element on the road towards a market oriented health care market.

The paper 'Yardstick competition for multi-product hospitals' analyzes the properties of possible regulation schemes for hospitals. The paper provides a formalized description and analysis of the proposed regulation of Dutch hospitals by NZa. The paper analyses the consequences of two proposed regulation models on cost reduction incentives, the production level and mix of hospitals, cost of regulation, innovation and quality.

The paper shows that yardstick competition gives clear cost reduction incentives, as well incentives to undertake cost reduction innovations in the transitory phase. The two analyzed models differ in the limitation of potential hospital bargaining power, robustness to collusion and administrative cost.

A theoretical drawback is that the regulation may distort the production level and production mix by hospitals. However, the paper argues that these effects are not harmful in practice and that the regulation will lead to a reallocation of production from inefficient hospitals to efficient hospitals. This reallocation will be beneficial for obtaining a more competitive outcome.

The paper is written by two invited authors: Per Agrell (professor at the Louvain School of Management, Université catholique de Louvain in Belgium) and Peter Bogetoft (professor of Economics at Royal Veterinary and Agricultural University KVL in Denmark), in combination with two NZa employees: Misja Mikkers and Rein Halbersma, both of the unit Economic Analysis.

The paper has been discussed with economists of the NZa Council of Advisors. The authors would like to thank professors Henk Don, Sweder van Wijnbergen, Erik Schut and Jan-Willem Velthuisen for comments that lead to many improvements.

Frank de Grave  
Chairman of the Executive Board  
Dutch Healthcare Authority



## Samenvatting in het Nederlands

In 2005 en 2006 zijn er ingrijpende hervormingen in de Nederlandse gezondheidszorg doorgevoerd. In 2005 is een voorzichtig begin gemaakt met de introductie van vrije prijsonderhandelingen tussen ziekenhuizen en zorgverzekeraars voor 10% van de ziekenhuisproductie. Voor de overige 90% van de ziekenhuisproductie is er ook in 2007 nog geen prijsconcurrentie en worden de prijzen gereguleerd. De minister van VWS stelt voor vanaf 2008 het resterende deel van de electieve zorg verder te dereguleren zodat vanaf 2011 voor het leeuwendeel van de electieve zorg vrije prijsconcurrentie mogelijk wordt.

Op dit moment hebben verzekeraars nog onvoldoende onderhandelingsmacht om een betere prijs/kwaliteit verhouding af te dwingen. Daarom stelt de minister van VWS voor een tijdelijk prijsplafond te introduceren. Het tijdelijk prijsplafond dat de NZa daartoe heeft ontworpen, is een competitief mechanisme geïnspireerd op het maatstafstelsel van Shleifer (1985). Het idee achter het maatstafstelsel van Shleifer is dat prijzen per product op de sectorgemiddelde kosten per product worden gebaseerd. Dit geeft ondernemingen een individuele prikkel om de kosten te verlagen. Met deze stimulans tot kostenverlaging wordt alvast gesimuleerd wat onder vrije prijsconcurrentie ook zal gaan ontstaan: een doelmatige en kwalitatief goede productie van zorg.

In dit artikel worden verschillende reguleringsmodellen geanalyseerd die gebaseerd zijn op het genoemde maatstafstelsel, maar dan toegepast op de Nederlandse situatie waarin zorgaanbieders in een complexe onderhandelingsomgeving opereren met onderhandelingen meerdere verzekeraars over duizenden producten (DBC's). Vanuit een oogpunt van administratieve lastenverlichting is gekozen om de regulering te laten aangrijpen op de gehele bundel van DBC's, en niet per DBC afzonderlijk. Dergelijke maatstafmodellen waren tot voor kort niet of nauwelijks bekend in de economische literatuur. Daarom heeft de unit Economische Analyse van de NZa in samenwerking met de academische onderzoekers Per Agrell en Peter Bogetoft de prikkelwerking en waarschijnlijke markttuitkomsten van het voorgestelde prijsplafond onderzocht. Dit artikel doet verslag van de bevindingen van dit onderzoek.

In een van de geanalyseerde varianten van het prijsplafondmodel wordt de maatstaf gebaseerd op de gemiddelde kosten per geproduceerde productbundel. De herverdeling van de totale sectorkosten geschiedt in dit model op basis van het aandeel van ieder ziekenhuis in de gewogen

totale sectorproductie. In dit model worden geen prijzen van individuele producten en geen afzonderlijke verzekeraar-zorgaanbieder relaties gereguleerd: alleen het overall prijsniveau van een zorgaanbieder dient onder het prijsplafond te blijven.

Zorgaanbieders worden uiteraard wel gecompenseerd voor niet-beïnvloedbare verschillen in de mix van de producten die ze produceren. Anders zouden instellingen die relatief dure producten leveren immers benadeeld worden. Door het prijsplafond op te hogen (of te verlagen) met een zogeheten casemixindex, wordt dit probleem omzeild en wordt tevens specialisatie lonend. De casemixindex wordt berekend als een gewogen volume van de verschillende DBC's. De gewichten waarmee de verschillende DBC's bij elkaar worden opgeteld, worden bepaald op basis van de relatieve zwaarte en schaarste. Productie van relatieve dure en complexe DBC's leidt dus tot een hoge casemixindex en een navenante verhoging van het prijsplafond. De casemixgewichten worden periodiek aangepast op basis van nieuwe producten, onderhandelde prijzen en wachtlijstgegevens.

Deze kostengeoriënteerde variant van het prijsplafondmodel geeft individuele zorgaanbieders een prikkel om efficiënter te opereren, onder de veronderstelling dat zorgaanbieders hun nut op rationele wijze maximaliseren. Deze prikkel bestaat er voor zowel de korte als de langere termijn. Tevens wordt aangetoond dat het model de onderhandelingsmacht van zorgaanbieders verlaagt. In het artikel wordt bovendien aangetoond onder welke voorwaarden (de mate van vraagelasticiteit en de functionele vorm van de kostenfunctie) het model tot verstoringen in de productie leidt. Een deel van deze verstoringen is vanuit welvaarts-oogpunt wenselijk in de zin dat er een herallocatie van productie plaatsvindt van relatief inefficiënte zorgaanbieders naar meer efficiënte aanbieders. Wanneer er verondersteld wordt dat de vraag betrekkelijk inelastisch is en in aanmerking genomen dat de overige, ongewenste, verstoringen betrekkelijk klein zijn, kan geconcludeerd worden dat de totale verstoringen geen grote negatieve gevolgen hebben. Het herallocatie-effect is gunstig leidt tot een leidt tot een meer competitieve uitkomst.

Een tweede geanalyseerde variant van het prijsplafondmodel is sterk vergelijkbaar met het eerste model, met dit verschil dat de maatstaf wordt gebaseerd op de onderhandelde prijzen in plaats van op de gerealiseerde kosten. Het model leidt tot vergelijkbare prikkels voor zorgaanbieders, maar de mate waarin de efficiëntieprikkels zich manifesteren in lagere onderhandelingsmacht van zorgaanbieders en prijzen is onzeker. Historische prijzen kunnen makkelijk als richtprijzen

worden gehanteerd, waardoor een transitie naar meer competitieve prijzen uit kan blijven. In dit model is collusie daardoor makkelijker dan de kostengebaseerde maatstaf. Daarbij geeft het mechanisme pas duidelijke prikkels wanneer de maatstaf achteraf wordt vastgesteld. Dit leidt echter weer tot meer onzekerheid bij zorgaanbieders.

De kostengebaseerde maatstaf leidt tot hogere reguleringskosten dan de prijsgeoriënteerde variant, doordat de toezichthouder informatie nodig heeft over de gestandaardiseerde kosten per zorgaanbieder. Deze administratieve lasten kunnen sterk verminderd te worden door aansluiting te zoeken bij informatie die reeds om andere doeleinden worden verzameld. Bovendien is de kostengebaseerde variant nog altijd veel minder ingrijpend dan het huidige systeem van afzonderlijke tarieven per DBC.

Elk model met sterke efficiëntieprikkels kan leiden tot een onderinvestering in innovatie. Het effect van deze onderinvestering wordt gemitigeerd doordat de prikkels voor zorgaanbieders om van innovatie te leren en over te nemen sterker worden. Het effect van onderinvestering kan verder worden verminderd door aparte innovatiecontracten met bijvoorbeeld academische ziekenhuizen te sluiten, zoals dat nu al gebruikelijk is in Nederland.



## **Abstract**

Health care provision is undergoing major reforms in Europe as a reaction to rapidly increasing expenditure and lowered political acceptance to commit public finance to cover the deficits. The Dutch government will decide in 2007 if the current budget system will be replaced by a more competitive mechanism, based on the yardstick regulation principle by Shleifer (1985).

One of the proposed systems for the reform can be compared to a revenue-cap implementation of a multi-product cost-yardstick mechanism. The redistribution of the sector's relevant cost is made in proportion to the individual hospitals share of total weighted output. No regulation is made of the multi-lateral contractual relations between users, insurers and hospitals. The scaling weights are updated periodically for new services and as a function of observed excess demand (waiting lists).

The mechanism is shown to provide cost-reducing (effort-inducing) incentives for profit-maximizing rational agents in a single-period bargaining game. The game also shows that the regulation acts as a countervailing power for the insurers to reinforce bargaining power. The local distortion of the output profile induced by the regulation is a function of demand elasticity and cost function convexity. In case the revenue target is not binding, no welfare loss is incurred. The regime also provides incentives for cost-reducing investments in the short and the long run. These incentives manifest themselves in local reallocations of output to more efficient producers.

The analysis shows that care should be taken in the updating of the weights as to provide incentives for service innovation, as well as to guarantee convergence of the price system. Although this requires econometric analyses, the alternative with a cost-based yardstick is considered superior to a revenue-based ditto for this application.



# 1. Introduction

## 1.1 Background

The Dutch health care sector is undergoing major reforms. In reforming the Dutch health care sector from an incumbent low-powered budgeting regime towards a more high-powered incentive regime, a number of instruments are deployed in the Health Market Design Act (WMG), mainly structural as to promote both supply-side and demand-side competition.

The Dutch Healthcare Authority (NZa) has elaborated and documented a proposed regulation regime for hospital reimbursement based on an adjusted service bundle price yardstick. In all, four projects aim at the theoretical, behavioral and practical analysis of the specific proposal as to investigate its feasibility and appropriateness for a potential implementation in 2008. The current paper is one of the research projects in the reform process, aiming at a theoretical investigation of the economic soundness of the chosen instrument.

## 1.2 Objectives

The goals this paper are to investigate if the proposed regulations with respect to:

- *Feasibility* in a complex multi-product industry
- *Economic soundness* of the motivation and coordination features of the proposal
- *Existence of an equilibrium* for some reasonable parameter settings within the proposal

## 1.3 Contribution

The paper extends the earlier literature on yardstick competition in healthcare provision by treating the multi-product dimension through a weighted output index. Whereas earlier work mainly has been focused on either detailed price-cap approaches per DRG with periodic updating or aggregate frontier estimates, the paper proposes an implementable solution that limits regulatory burden for suppliers in the pricing game to observing negotiated contracts, realized output, audited total costs and the excess demand function (waiting lists).

The paper also provides new insights into the state of current European healthcare reform programs under institutional diversity.

## **1.4 Outline**

The outline of the paper is as follows. Chapter 2 gives a short introduction to the health care sector in the Netherlands, in particular to the hospital sector and the incumbent regulation. Chapter 3 formalizes the relationships in a base model and it captures the combined usage of regulation and bilateral negotiations by a constrained bargaining framework. The basic properties of the proposals are derived in Chapter 4. Chapter 5 provides a model-independent discussion on incentives for innovation and quality incentives in the new regime. Chapter 6 closes the paper with some conclusions.



## **2. Health care in the Netherlands**

In this section, we present some key details about the Dutch health care sector and the current reforms.

Since the last few years, the Dutch health care system is being reorganized. For this paper, the introduction of competition in the hospital sector in 2005 and the reorganization of the health insurance system in 2006 are of particular importance. The reforms are based on a model of managed competition with mandatory insurance for all Dutch citizens. The basic idea is that hospitals compete for contracts with insurers and insurers compete for contracts with consumers.

A major reason for the market oriented reforms is the public dissatisfaction with the negative effects of rationing (such as waiting lists) induced by the current cost containment policies (see Schut and Van de Ven, 2005).

### **2.1 The health insurance market**

The coverage of the mandatory insurance is a basic package defined by law, that includes all essential medical care. There is an optional additional insurance covering all health services not included in the basic package (De Jong and Mosca, 2006). The mandatory insurance aims to guarantee universal access to the health care system.

Insurers have to accept any client regardless of gender, state of health, age etc., without room for price discrimination and risk screening. A refined risk adjustment system is in place to compensate insurance companies for cost differences induced by socio-demographic factors, such as age, sex, income, location and prior health care consumption (chronic pharmaceutical dependencies and prior hospitalization) (see Schut and Van de Ven, 2005).

The premium for the new insurance consists of two components: a nominal premium of around 1,000 Euro and an income related premium of 6.5% of the income up to an income ceiling of 30,015 Euro (De Jong and Mosca, 2006). The income related premiums are collected through payroll and income taxes and are redistributed through the risk-adjustment system.

The risk adjustment system levels the playing field for health insurers by enabling price competition on the nominal premium rates (see Schut and Van de Ven, 2005).

Another important characteristic is that insurers are obliged to make sure that their enrollees get the necessary treatments within a certain time and geographic area. In other words, insurers are obliged to contract sufficient care, given the demand of their enrollees.

Health insurers are privately owned (both for profit and not for profit). Many are part of larger holding companies within the financial sector, with access to capital markets. Following the introduction of the new insurance competition, a few mergers between the largest players have reduced the market to four large national insurers (with about 80% of the entire market) and a handful of smaller national or regional players.

## **2.2 Hospital budgeting**

The Dutch government plans a step by step introduction of price competition between hospitals. The introduction of competition in the health care sector has been long debated. For a thorough discussion of the political decision making process leading to the recent reforms, we refer to Helderma et al. (2005). In this section, we will shortly describe the most relevant changes for this paper.

The current budget system (which is called 'Functional Budget System') was introduced in 1983 to contain hospital cost. The budget was set by the legal precursor of the NZA, the National Health Tariff Authority ('CTG'). Hospital budgets were based on a small number of 'parameters', such as the number of inhabitants in the surrounding area, the number of beds, specialists and assistant-doctors and three broad measures of production: the number of outpatient visits, inpatient visits and overnight stays. All these parameters were expressed in monetary units.

Within their budget constraint, hospitals bargained with the collective of insurers on local production plans. As the National Health Tariff Authority set administrative tariffs on hospital products, only the production volumes and quality could be freely negotiated. Furthermore, hospitals and insurers were mutually obliged to contract with each other.

## 2.3 Liberalising hospital-insurer contracting

In 2005 two major changes were introduced. First, a new hospital payment system was implemented based on Diagnosis Treatment Combinations (DBC's). A DBC is a product definition based on a medical description and contains the whole inpatient and outpatient activities (Haeck, 2005).

Second, since 2005, the segment of uncomplicated, elective outpatient care has been deregulated. First, prices in the competitive segment are no longer set by the National Health Tariff Authority, but are subject to bargaining. Second, the bargaining process is no longer multilateral but bilateral, with no mutual obligation to close any contract.

This competitive segment consists of 1376 different DBC's. Since some DBC's are almost identical, the group of DBC's in the competitive segment can be clustered to 145 different products (See CTG/ZAio, 2005). The DBC's in the competitive segment cover 15 (out of 24) different medical specializations and belong to 28 different diagnoses.

The revenue of the competitive segment is approximately 1.1 billion Euro, which is about 8% of the total expenditure on hospital care. To eliminate the revenue associated with the competitive segment from the prospective budgets for the regulated segment, the Dutch Healthcare Authority estimated cost prices (i.e. average unit cost) for the products, based on a survey of a group of 12 hospitals and multiplied these cost prices with the estimated volumes.

Apart from hospitals, there are also so-called Independent Treatment Centers (ZBC's) active in the market for hospital care. These ZBC's are small outpatient treatment centers which have been allowed in the market since 1998. However, ZBC's are only allowed to provide treatments that do not require an overnight stay in the hospitals, so they do not compete on the whole range of elective care products.

## 2.4 The need for transitory regulation

The major part of the hospital production is currently still regulated by the old budget system. Currently, the Ministry of Health is proposing to deregulate the remainder of elective hospital care (including inpatient elective care), estimated to be 70% of hospital production. The goal is to introduce more incentives for efficiency in order to guarantee sustainable health care expenditures for the long run.

The results from first two years of market incentives in the small competitive segment have been mixed, however. The National Healthcare Authority (CTG-Zaio 2005 and CTG-Zaio 2006) monitored the price developments for the elective outpatient hospital care and established that there was a net real price increase over the two year period, albeit with a small real price decrease in the second year

A more detailed analysis by the NZa (CTG-Zaio, 2006) established that the underlying bargaining process currently relatively favors hospitals over insurers. The econometric evidence points towards supernormal profit margins for hospitals and Nash bargaining outcomes significantly different from the competitive equilibrium.

Within the current competitive domain, insurers are unable to use their countervailing power in reducing profit margins and moving bargaining outcomes towards the competitive equilibrium. This is most obvious from the fact that almost every insurer contracts with almost every hospital, even though they are under no obligation to do so. This undermines the credibility of threatening to move patients towards hospitals with the best price-quality ratios.

Insurers indicate that they cannot yet credibly threaten to move patients towards their preferred providers for several reasons. First, the lack of transparent hospital quality information gives patients little incentive to overcome additional travel time. Second, insurers are legally restricted to significantly differentiate reimbursements across providers. Finally, patients currently seem to value freedom of choice over lower premiums.

Finally, almost the entire hospital sector has been without incentive regulation for several decades. Market parties have had little time to reach some sort of long run competitive equilibrium. The new proposed hospital regulation system in the Netherlands is meant to be a transitory element on the road towards a more fully market oriented healthcare market.

### 3. Base model

To get a better idea of the proposed regulation, we start by formalizing the basic relationship between production, costs and price-caps. We shall next combine this with an outline of likely objectives of the parties involved, viz. the users (insurance companies), the hospitals and the regulator. Lastly, the base formula and the objectives are developed into a more refined constrained bargaining model.

#### 3.1 Economic entities

The framework involves three types of economic actors, the regulator, hospitals and users. A user here involves a combined entity of patients that use a given insurance company. The entity indirectly involves general practitioners (GP) as well since we presume that the GPs remain responsible for the visitation of patients to the hospitals, cf. below. We shall not model the possible internal conflicts and informational asymmetries within the user group. Likewise, the hospital entity here involves both a hospital and the physicians involved. Bogetoft and Mikkers (2006) show that from the point of view of incentivizing the hospital and physicians, it is often attractive to look at them as an integrated entity – or at least to contract with the physicians via the hospitals.

The regulator serves as a market maker that puts restriction on the contracts and payment plans that individual hospitals and users can choose.

The three entities are illustrated in Figure 3-1 below.

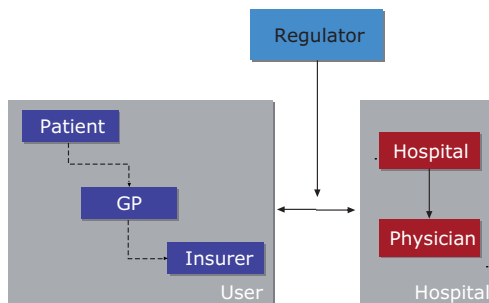


Figure 3-1 Economic entities

### 3.2 Production, costs, and revenues

Let there be  $H$  hospitals each involved in the production of up to  $D$  DBCs. The production at hospital  $h$ ,  $h=1,\dots,H$ , of DBCs  $d$ ,  $d=1,\dots,D$ , is denoted  $y_{hd}$  and the vector of outputs from hospital  $h$  is denoted  $Y_h$ , i.e.

$$Y_h = (y_{h1}, \dots, y_{hD})$$

Since not all hospitals offer the full range of possible treatments, some of these dimensions are typically zero. Also, the period of production shall be suppressed but it will be introduced when we discuss lagged models below. We shall think of the productions  $y_h$ ,  $h=1,\dots,H$  as being realized as opposed to planned, and we shall assume that they are verifiable.

The total revenue of hospital  $h$  is denoted  $R_h$  and the total cost is  $C_h$ . Both are assumed to be verifiable through normal auditing procedures. The total hospital cost include all operational and capital costs used in providing the produced DBCs. In particular, operational costs include (the currently separately accounted) physicians' fees and the capital costs include not only depreciation of fixed assets but also standardized cost of equity and debt.

Since hospitals potentially produce services other than DBCs (e.g. research, teaching), and since not all DBCs will fall under the proposed regulation (e.g. experimental care), it will be necessary to develop uniform and unambiguous accounting rules to split hospital revenues and costs across the several regulatory segments.

We shall in general *not* assume that the any more specific revenue elements are verifiable. One rather obvious extension is to assume that the revenue of hospital  $h$  can be split according to the user  $u=1,\dots,U$  paying it

$$R_h = \sum_u R_{uh}$$

Here  $R_{uh}$  is shorthand for the revenue paid by user  $u$  to hospital  $h$ .

We will not assume that the regulator can observe the contracts signed between users and hospitals. In consequence, the regulator is not assumed to use information about the principles according to which payments are settled. The revenue  $R_{uh}$  can for example result from a contract with a fixed payment for a given capacity, a payment per DBC or some combination.

### 3.3 Cost based ex ante regulation

The proposed regulation, the cost based ex ante regulation, shall now be formalized. Some obvious variants, cost based ex post as well as revenue based ex ante or ex post regulations will be discussed at the end of this chapter.

#### Weights

The proposed regulation is based on DBC weights

$$w=(w_1,\dots,w_D)$$

The weights are set by the regulator and intended to reflect relative scarcity or cost of different DBCs.

#### Average DBC-adjusted price-cap

The proposed scheme suggests that the average DBC price in any given hospital cannot exceed the average DBC price in the sector when corrected for case mix. Using the notation above, this can be formulated into the following requirement for hospital  $h$

$$\frac{R_{h^*}}{\sum_d y_{h^*d}} \leq \frac{\sum_{h \neq h^*} C_h}{\sum_{h \neq h^*} \sum_d y_{hd}} \square \frac{\sum_d w_d y_{h^*d} / \sum_d y_{h^*d}}{\sum_{h \neq h^*} \sum_d w_d y_{hd} / \sum_{h \neq h^*} \sum_d y_{hd}}$$

The left hand side is the revenue of hospital  $h^*$  per DBC volume. The DBC volume is a simple summation of the treatments provided at a given hospital and it therefore does not reflect the hospital's degree of specialization. On the right hand side, the first fraction is the industry wide cost per DBC volume, while the last fraction, the *case mix index*, is intended to correct for the fact that the case mix of hospital  $h^*$  may be more or less resource demanding than the industry mix.

Quality of treatments is not mentioned explicitly in the above formulation. At this point, however, this is without loss of generality since one can simply differentiate treatments according to quality, i.e. assume that two DBCs may represent the same treatment at different quality levels. This presumes of course that quality is verifiable at the treatment level. If this is not the case, alternative adjustments may be needed.

Observe also that we have excluded the hospital in question,  $h^*$ , from the determination of the industry wide average cost and volume calculations. This is in accordance with a basic idea in the yardstick literature, namely

that a unit should not be able to affect its own benchmark. In reality and with a reasonable number of hospitals, this refinement may not matter much<sup>1</sup>.

### **Associated revenue cap function**

Although the yardstick scheme is formulated as a price cap requirement, it can be convenient analytically to think of it as a revenue cap system in which the revenue cap is a (linear) function of the ex post realized volume. This is radically different from a fixed budgeting system where a revenue cap has no volume dependence. Reformulating the above, we see that the associated revenue cap function for hospital  $h^*$  is effectively

$$R_{h^*}^{CAP}(y_{h^*}) = \frac{\sum_{h \neq h^*} C_h}{\sum_{h \neq h^*} y_h w} y_{h^*} w$$

where the revenue cap function for hospital  $h^*$ ,  $R_{h^*}^{CAP}(\cdot)$  is the maximal sum of charges it can make as a function of the services it provides. In the following we will loosely denote the associated revenue cap function as a revenue cap, but this is not to be understood as a fixed budget imposed on the hospital.

The revenue cap function can be rewritten into

$$R_{h^*}^{CAP}(y_{h^*}) = y_{h^*} w \left( \frac{\sum_{h \neq h^*} C_h}{\sum_{h \neq h^*} y_h w} \right)$$

This suggests the following interpretation of the system: The regulator sets the weights,  $w$ , equal to the relative values of the different services, and the actual costs level in the industry serves to calibrate the absolute level. If actual costs  $\sum_{h \neq h^*} C_h$  exceeds the costs stipulated by the regulator,  $\sum_{h \neq h^*} y_h w$ , then the allowed charges are increased with the same percentage and vice versa.

We note also that the revenue cap function is linear. This means that there it increases in direct proportion to the number of DBCs, i.e. it is a constant return to scale scheme, and that the rate of substitution between DBCs is constant (independent of specialization etc), i.e. that the iso-revenue curves are linear.

<sup>1</sup> The current number of full-scale hospitals is around 100, with an additional 130 outpatient treatment centers.



### 3.4 Objectives

To predict the effects of using the above revenue cap formula, we need to make assumptions about the motives of the parties involved. We shall now discuss these motives.

#### Hospital objectives

Let us assume that the hospital can affect the resulting costs according to the cost function  $C_h^*(y_h^*, e_h)$ . Here, the variable  $e_h$  represents the non-verifiable workload in activity coordination and local incentive alignment among categories of staff within the hospital. It is traditionally referred to as "effort" for simplicity. Also, let  $v_{h*}(e)$  be the cost of effort in monetary equivalents.

The behavior of the hospital will depend on its overall motives. The classical assumption in the yardstick literature is that firms are profit maximizing. In reality this is far from always the situation, and, especially in Dutch hospitals, we can image other objectives such as revenue maximization or volume maximization.

Currently, all hospitals in the Netherlands are private not-for-profit by law. The Healthcare Providers Entry Act (WTZi) allows the introduction of the profit motive from 2012 onwards. Nevertheless, even in market oriented countries as the U.S., the vast majority of hospitals are either publicly operated or private not-for-profit. The health economics literature (Sloan, 2000) suggests that private ownership matters more than the profit motive as far as performance and quality are concerned. To facilitate further analysis, we first impose the profit motive and will relax this assumption later on.

In general, not-for-profit hospitals may not want to cut costs, but rather maximize "slack", i.e. usage of resources on non-market goods or usage of resources at uncompetitive prices<sup>2</sup>. This could for example be better working conditions for the employees or extra non-verifiable quality of treatments. Also, they may only be partially concerned about cost reduction and profit maximization, namely to minimize bankruptcy risk. Despite a multiplicity of goals, a hospital must in the end care about the continuity of operations.

<sup>2</sup> E.g., a not-for-profit hospital maximizing a linear combination of profit and revenue can be regarded as a profit maximizing hospital with an artificially lowered internal cost function (Lakdawalla and Philipson, 1998).

We start by assuming profit maximizing behavior, i.e. the hospital  $h^*$  maximizes the net gains, i.e. it solves

$$\max_{Y_{h^*}, e_{h^*}} \Pi_{h^*} = R_{h^*}(Y_{h^*}) - C_{h^*}(Y_{h^*}, e_{h^*}) - v_{h^*}(e_{h^*})$$

Here  $R_{h^*}(Y_{h^*})$  is the revenue function faced by the hospital. It will depend on the demand side and the regulation. If we assume that the hospital has all the bargaining power and that there is no uncertainty, we could assume it to equal  $R_h^{CAP}(Y_{h^*})$ . We shall return to this below.

If we anchor the cost function at maximal effort (minimal slack), and work with a simple slack model where  $r$  is the value of 1 Euro slack, the hospital will solve

$$\max_{Y_{h^*}, S_{h^*}} \Pi_{h^*} = R_{h^*}(Y_{h^*}) + \rho S_{h^*} (C_{h^*}(Y_{h^*}) + S_{h^*}) = R_{h^*}^{cap}(Y_{h^*}) - C_{h^*}(Y_{h^*}) + (\rho - 1) S_{h^*}$$

### User objectives

Let  $B(y_u)$  be the benefits to user (group)  $u$  from treatment profile  $y_u$ . Also, let  $R_u(y_u)$  be the payment of user group  $u$  – via insurance, taxes etc. We assume that user  $u$  tries to maximize net benefits, i.e.

$$\max_{y_u} B_u(y_u) - R_u(y_u)$$

### Regulator objectives

We may in general assume that the regulator strives to maximize social welfare, which here corresponds to a weighted sum of the users' surplus and the producers' surplus, i.e. (Baron and Myerson, 1982)

$$W(y, e) = \sum_u \{B_u(y_u) - R_u(y_u)\} + \eta \sum_h \Pi_h(y_h, e_h) - \lambda \sum_u R_u(y_u)$$

Here,  $\eta \in [0, 1]$  is the weight associated with the hospital surplus and  $\lambda \geq 0$  is a parameter corresponding to the (tax distortions) cost of public funds. In the Netherlands, half of hospital costs are covered by insurance premiums and half by pay-roll taxes.

As a benchmark, one may consider a *first best* situation, where the regulator has perfect information and is able to instruct the hospitals effort  $e$ , production  $y$  and payment  $R$ . In reality, the regulator is restricted by asymmetric information, and must settle with a *second-best solution*. In a second best solution,  $W$  is maximized subject to both *individual rationality* (IR) and *incentive compatibility* (IC) constraints ensuring that all hospitals are willing to participate and choose the planned efforts and productions and that the clients are willing to pay for the services.

Let us assume that payment schemes are fixed as  $R_h(y_h)$  for hospital  $h$  and  $R_{uh}(y_{uh})$  from user  $u$  to hospital  $h$ , i.e. we have

$$R_h(y_h) = \sum_u R_{uh}(y_{uh})$$

We can therefore formulate the IR and IC conditions as

$$\begin{aligned} \Pi_h(y_h, e_h) &\geq 0 & \forall h \\ (y_h, e_h) &\in \operatorname{argmax}_{(y_h, e_h)} \Pi_h(y_h, e_h) & \forall h \\ B_u(y_u) - \sum_h R_{uh}(y_{uh}) &\geq 0 & \forall h \end{aligned}$$

Observe that in the proposed regulation, the regulator does not intend to regulate the payment plans in all details, only to influence the (price-level of the) overall payment  $R_h$ . This means that we shall ideally include IR and IC constraints corresponding to the choice of the exact payment schemes between individual users and hospitals, the  $R_{uh}(\cdot)$  plans. The contemplated  $R_{uh}$  - plans must be individually rational and incentive compatible as well.

In the mechanism design literature, it is customary to maximize the regulator's objective subject to the IR and IC constraints like above. This may lead to important insights in simplified settings, and in turn this may give useful benchmarks. In applied regulation, however, the stress is more on the complexity of the real setting, like the multiplicity of outputs in health, and the objective of maximizing social welfare becomes analytically too ambitious. We therefore focus on the economic soundness of some specific approaches, i.e. we investigate if the more specific mechanisms proposed in the Dutch context lead to reasonable coordination of activities (production level, production mix etc) and reasonable motivation of the entities involved (incentives to minimize costs, make innovations etc)..

### The regulation procedure

We close this section by refining the description of the proposed system. In addition to the basic price-cap formula, we need to understand the procedural usage of this. We shall therefore work with time-conditioned variables in this section.

The idea of the scheme is to work with price caps set ex ante based on past experience. This means that the associated revenue cap function would be known ex ante. Moreover, since the hospitals will only be held responsible for their total charges, the price caps only serves to calculate the revenue cap function which is the maximal allowed charge.

The system in this way resembles an ex ante regulation with ex post volume adjustments. The ex ante known revenue cap for period  $t$  is given by

$$R_{h^*t}^{CAP}(y_{h^*t}) = y_{h^*t} \left( w \square \frac{\sum_{h \neq h^*} C_{ht-1}}{\sum_{h \neq h^*} y_{ht-1} w} \right)$$

The idea of the scheme is to impose an upper bound on the overall price level, rather than to impose a fixed revenue target. That is, (1) if a hospital charges less than the revenue cap, nothing happens, and (2) if a hospital produces more volume it is allowed to earn more revenue. If, however, it charges more for its produced DBCs than the allowed price cap, and hence earns more than the associated revenue cap, it must repay the surcharges, possibly with some extra penalty  $a$  (in addition to interest), i.e. to repay in net present value  $(1+\alpha)[R_{h^*t} - R_{h^*t}^{CAP}(y_{h^*t})]$

It is clear, however, that unless penalties are very high, a hospital will indeed charge above the revenue cap, if it has enough market power over local insurers to do so. In consequence, the system would work like an ex ante revenue cap system with ex post adjustments for volume. It would therefore be similar to the standard regulation systems used to regulate for example distribution system operators in electricity. The main difference is that to cope with the many output dimensions, the system work with a simple weighting system to aggregate the outputs into one simple output.

The fact that surcharges may be repaid to the risk-adjustment fund and not necessarily to the users (the insurance companies) means a transfer from one insurer that pays too much to all insurers. More specifically, the charges  $R_h(y_{ht}) = \sum_u R_h(y_{uht})$  are compared to allowed charges  $R_{ht}^{CAP}(y_{ht})$

and modified revenue  $R_h^{MOD}(y_{ht})$  and user payments  $R_u^{MOD}(y_{ut})$  are calculated as

$$R_h^{MOD}(y_{ht}) = R_h(y_{ht}) - \max \left\{ 0, (1 + \alpha) [R_h(y_{ht}) - R_{ht}^{CAP}(y_{ht})] \right\}$$

$$R_u^{MOD}(y_{ht}) = R_u(y_{ht}) - \sum_h \beta_u \max \left\{ 0, (1 + \alpha) [R_h(y_{ht}) - R_{ht}^{CAP}(y_{ht})] \right\}$$

Here,  $\beta_u$  is a measure of user  $u$ 's size compared to other users.

This will encourage the insurance companies to limit a hospitals payment to the yardstick level. One problem however is that a given hospital may contract with several insurance companies and that individually they may free ride in the bargaining.

The above redistribution scheme collects the total surcharges of all hospitals and distributes the aggregate surcharges across all insurers proportional to their size (as measured by e.g. their aggregate demands). An alternative is to redistribute surcharges per hospital across all locally involved insurers proportional to their size *in that hospital*. The user payments would become

$$R_u^{MOD}(y_{ht}) = R_u(y_{ht}) - \sum_h \beta_{uh} \max\{0, (1 + \alpha)[R_h(y_{ht}) - R_{ht}^{CAP}(y_{ht})]\}$$

Here,  $\beta_{uh}$  is a measure of user  $u$ 's size in hospital  $h$  compared to other local users.

### The full game

The *timing of the resulting game* in a given period is as follows:

- The regulator announces price caps for period  $t$  (or equivalently, the revenue cap functions  $R_{ht}^{CAP}(y_{ht})$ ) for the different hospitals.
- The hospital and users negotiate payment schemes  $R_{uht}(y_{uht})$  for the different users and hospitals.
- Users choose service levels,  $y_{uht}$ , and hospital effort levels  $e_{ht}$ .
- Realized charges  $R_h(y_{ht}) = \sum_u R_{uh}(y_{uht})$  are compared to allowed

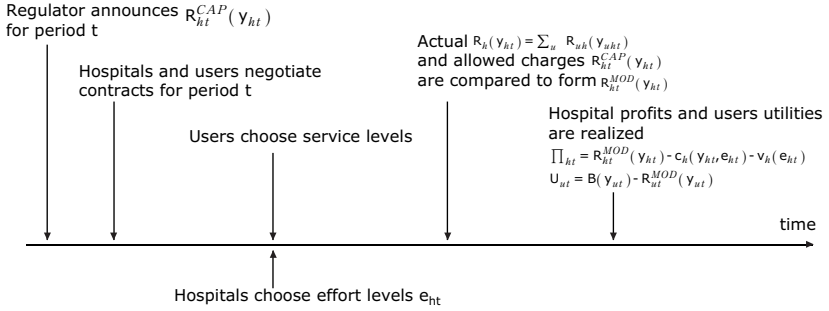
charges  $R_{ht}^{CAP}(y_{ht})$  and modified revenue  $R_h^{MOD}(y_{ht})$  and user

payments  $R_u^{MOD}(y_{ut})$  are calculated

- Hospital profits and user utilities are realized:

$$\begin{aligned}\Pi_{ht} &= R_h^{MOD}(y_{ht}) - c_h(y_{ht}, e_h) - v_h(e_h) \\ U_{ut} &= B(y_{ut}) - R_u^{MOD}(y_{ut})\end{aligned}$$

The timing of the ex ante cost based price-cap procedure is illustrated in Figure 3-2 below.



*Figure 3-2 Ex ante cost-based price-cap procedure*

### 3.5 Revenue based yardsticks

An alternative formulation of the model above that we shall investigate is to make a redistributive game with respect to sector revenues rather than costs. That is, the price caps or the associated revenue caps are calculated from observed revenues rather than from regulatory audited costs.

In its simplest possible form, the *ex ante revenue based scheme*, this means a replication of the above except that the previous revenue function  $R_{ht}^{CAP}(y_{ht})$  is changed into a new revenue cap function

$$G_{h^*t}^{CAPex\,ante}(y_{h^*,t}) = \frac{wy_{h^*,t}}{\sum_{h \neq h^*} wy_{ht-1}} \sum_{h \neq h^*} R_{ht}(y_{ht-1})$$

The only difference is therefore that we use revenues rather than costs to determine the general payment level. The system is similar in terms of ex ante commitment to the allowed revenue levels and ex post adjustments for mix and volume changes.

Now, it is quite clear that, except for expansions in volume, the ex ante based revenue system will not allow for any increases in the price-level of spending, irrespectively of the actual development in the cost structure. This is clearly not a viable approach and we shall therefore focus the discussion on a less naive revenue based yardstick scheme, namely one based on ex post revenues rather than ex ante ones.

In the *ex post revenue based scheme*, present revenues and volumes are substituted for historical ones corresponding to a revenue cap function

$$G_{h^*t}^{CAPex\ post}(y_{h^*,t}) = \frac{wy_{h^*,t}}{\sum_{h \neq h^*} wy_{ht}} \sum_{h \neq h^*} R_{ht}(y_{ht})$$

In the same way as one may consider ex post revenue based systems, one can of course consider ex post cost based systems. In some cases, this may be useful, e.g. to avoid arbitrary adjustments for price changes, cf. Agrell and Bogetoft (2005).

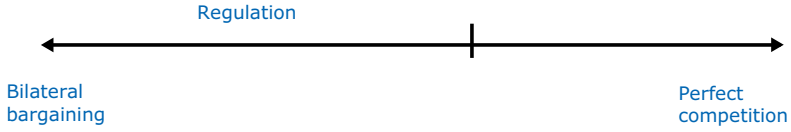
In terms of the timing and the full game, it does not alter the general description above. The so-called ex ante and ex post systems both have ex ante and ex post elements. They both commit ex ante to the approach for determining allowed revenues ex post and they both allow for ex post adjustments to volume and mix. The difference is only if the payment level can be fixed ex ante using historical data or whether it should be settled ex post using the most recent data on cost or revenue levels. In the case of revenue based caps, it is necessary to use an ex post approach to get sensible results. In the case of cost based schemes, both an ex ante and an ex post approach may have advantages. Lastly, we note that with a cost based approach the difference between ex ante and ex post may be limited when the ex ante commitments are not for too long periods, i.e. with fast updating, the two systems are quite similar.

### 3.6 Bargaining and competition

The likely impact of the proposed regulation depends on the details of the Dutch context. In particular, it depends on the degree to which hospitals and users (insurers) can exercise market power. In the transition phase, for which the regulation is intended, the situation can best be characterized as one of bilateral bargaining among a multiplicity of hospitals and users.

As a worst case scenario in terms of social welfare, we shall therefore investigate how the regulation would impact the bilateral bargaining among one representative hospital and one representative user.

This corresponds to the extreme outcome in Figure 3-3 below.



*Figure 3-3 Between bargaining and competition*

### **Nash bargaining**

The bargaining between hospitals and insurance companies can be modeled in many different ways.

An important question here is the role of asymmetric information and the possible losses from strategic behavior. Given the multiplicity of users and hospitals, the long history of collective negotiations on production plans and the detailed expenditure analyses for the risk adjust system, we do not believe asymmetric information to be a dominant factor. Still, we shall discuss some of the impacts of strategic behavior and the interaction with the proposed regulations in the next chapter.

In terms of formal modeling, we shall instead use a convenient and relatively general approach to bargaining under perfect information, namely generalized Nash bargaining, i.e. Nash bargaining with possibly non-symmetric bargaining power.

Let  $\gamma_h$  be the (relative) bargaining power of a hospital  $h$  and  $\gamma_u=1-\gamma_h$  be the relative bargaining power of user  $u$ . Also, let the costs of hospital  $h$  be  $c_h(y_h)$  and the monetary equivalent benefits to the user of hospital  $h$  be  $b_h(y_h)$ . The "gains from trade" are therefore  $b_h(y_h)-c_h(y_h)$ . This benefit is transferable via the selected payment level  $R_h(y_h)$ . If there is only one hospital and user, the unconstrained Nash solution predicts a production level  $y_h$  and division of net benefits using  $R_h$  that solve

$$\max_{y_h, R_h} [b(y_h) - R_h]^{\gamma_u} [R_h - c_h(y_h)]^{\gamma_h}$$

This result has an axiomatic foundation. The classical symmetric Nash solution is the unique outcome that satisfy individual rationality, scale invariance (against affine transformations), symmetry, independence of irrelevant alternatives, and Pareto efficiency. It implies that the product of the agents' gains from trade shall be maximized.



In the case of a hospital and a user having to divide a net social value of 1 M€, this is illustrated in Figure 3-4 below.

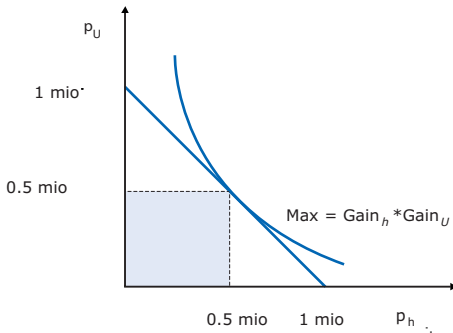


Figure 3-4 Unrestricted symmetric Nash bargaining

A simple consequence of the asymmetric (generalized) Nash bargaining solution is that if the hospital and user has to share a fixed net-surplus of say  $P$ , they will do so in proportion to the bargaining powers, i.e. the hospital will get  $\gamma_h P$  and the user  $\gamma_u P$ .

Now, taking into account the regulation, the hospital and user cannot freely choose the production level and transfer  $R_h$ . They must instead solve the following restricted bargaining problem:

$$\begin{aligned} \max_{y_{ht}, R_{ht}} & [b(y_{ht}) - R_{ht}]^{\gamma_u} [R_{ht} - c_{ht}(y_{ht})]^{\gamma_h} \\ \text{s.t.} & R_{ht} \leq y_{ht} p_t \\ & p_t = w \frac{\sum_{h' \neq h} c_{h't-1}}{\sum_{h' \neq h} y_{h't-1} w} \end{aligned}$$

Of course, instead of assuming that one hospital only bargains with one insurance company, we could extend the model to account for the fact that hospitals will typically bargain with several hospitals. A simple way to do so is to assume that the bargaining is basically independent except for the linkage through the revenue cap. This leads to

$$\begin{aligned} \max_{R_{uht}, y_{uht}} & \sum_u [b(y_{uht}) - R_{uht}]^{1-\gamma_{uh}} [R_{uht} - c(y_{uht})]^{\gamma_{uh}} \\ \text{s.t.} & \sum_u R_{uht} \leq u_{ht} p_t \\ & p_t = w \frac{\sum_{h' \neq h} c_{h't-1}}{\sum_{h' \neq h} y_{h't-1} w} \end{aligned}$$

We shall not pursue this any further at this stage since such an approach ignores the fact that production costs will typically depend on the total production as well. In the single user model, this is implicit in the interpretation of the cost function. We simply use  $c$  to represent cost given the production level negotiated with other firms.

### **Bargaining power and outside options**

By using the Nash bargaining solution, we basically assume *efficient bargaining*. That is, we abstain from any explicit modeling of the asymmetric information and the frictions and dead weight losses this may give in the negotiations. This seems a reasonable simplification – but as indicated, we shall discuss it further below.

By varying the *bargaining powers* we can get reduced form representations of alternative settings. Thus for example, if there are several hospitals serving a given area,  $g_h$  should fall and if there are several insurance companies competing for the services of the given hospital,  $g_h$  should increase. In practice, this is at least partially confirmed by studies of cost mark-ups in the competitive part of the hospital sector, cf. Halbersma et al. (2006).

In this way, we suggest that the details of the market conditions, the number of more or less substitute suppliers (hospitals) for a given user and the number of more or less substitute buyers (user) of a the capacity of a given hospital, can be roughly captured by the bargaining powers.

From a bargaining perspective, the regulation functions as an *outside option* or a credible threat to discard certain settlements. As it has been discussed in the literature, this can be modeled either by invoking alternative threat points – or by changing the bargaining set. We have taken the latter approach since it corresponds best to explicit modeling of such effects in for example a strategic bargaining model à la Rubinstein (1982).

The Rubinstein (1982) approach, i.e. the writing down of some particular sequence of offers and replies to be made over time in the course of negotiations, and then looking for a non-cooperative equilibrium in the game thus specified, has been used in numerous studies. For a partial overview, see e.g. Sutton (1986). One of the issues that has been addressed is the role of outside options, see also Binmore, Rubinstein, and Wolinsky (1986). An important insight, sometimes referred to as the outside option principle, is that having access to outside options does not necessarily influence the outcome because the threat of having recourse to these may not be credible (i.e. not in one's best interest when given

the option). If we modeled it via the threat-point, it would almost always influence the outcome which is not realistic – a very lax regulation will not influence the outcome.

The simplification from real bilateral bargaining between a multiplicity of players to bilateral Nash bargaining among a single user and hospital may also impact the solution in other ways. A user may for example have a total benefit function that is largely *price-inelastic* up to a fixed demand for services, but his demand from a single hospital may be at least somewhat elastic since by using the given hospital he also forgoes the option to use another hospital. This suggest that the case of inelastic demand is not entirely realistic in bilateral negotiations. Again, if we include the possibility of buying services from another hospital into the bargaining, this correspond to an outside option to the user, and if we allow this by varying the price elasticity of demand, we are once again modeling via the set of feasible outcomes.



## 4. Analysis

### 4.1 Markets, bargaining and regulation

The proposed regulation is intended to work in an environment that have features of both the most attractive of economic settings, that of a perfectly competitive market, and one of the least attractive economic settings, that of bilateral bargaining.

#### Market

In a perfectly competitive market, everyone act as price takers and there is *no strategic behavior*. Sellers and buyers alike reveal the relevant parts of their supply (marginal cost) and demand (marginal benefit) curves, and the competitive equilibrium is formed where the horizontal aggregation of demand and supply schedules meet. In this situation, the total gains from cooperation (trade) are realized. In particular, all marginal rate of substitution is the same for all buyers and equal the same marginal rate of technical transformation for the sellers.

In reality, this ideal is of course difficult to obtain in the exchange of numerous and complex products like DBCs. The multiplicity of products and the inability to freely handle all relevant information creates economic losses. Likewise, the possible preferences for local treatments and the possibilities of economies of scale and scope may hinder the ideal outcome and call for some sort of regulation.

#### Bilateral negotiations

In a bilateral trade regime, the individual users and hospitals search the market and negotiate price on a bilateral basis. This allows for some attractive arrangements, but suffers from several problems compared to the competitive outcome.

The first problem is a *matching problem* of finding the right partners and quantities. The direct search costs in terms of time, advertisements, etc. may be considerable. Likewise, the indirect search costs of for example ending up with a suboptimal quantity may be significant.

The second problem is that of *imperfect information and strategic behavior*. Bilateral trade has to be settled among users and hospitals with private information about the values and costs of services. Simple Bayesian bargaining models have demonstrated how sellers will overstate costs and buyers understate the values, often to the extent that no trade

is realized, even when the buyer in reality value the services above the costs of the sellers. As emphasized by Myerson and Satterthwaite (1983), delays, failures and social losses are inevitable in private bargaining when it is uncertain that the gains from trade can be realized.

The third problem is that of *possibly uneven bargaining power*. Experimental studies show that bilateral trade in a market with many buyers and sellers tends to empower the sellers, and result in higher prices than the competitive prices. This result was first shown in Ketcham et al. (1984). Non-competitive prices will result in inefficiency and most likely lower supplied quantities.

In conclusion, it is unlikely that bilateral trade will suffice to realize the full potential gains from reallocating the production of DBCs. Moreover, and just as importantly, it is likely that some parties, including the hospitals, may extract excessive rents.

### **Regulator as a market maker**

The role of the regulation is to limit some of the drawback of having a non-competitive market for DBCs.

It is important to understand that the regulator does not in this case – as it is typically the case in the regulation of (local) monopolies – act as a proxy consumer that acquires the necessary services at a low price. Rather, the regulator lets the parties, hospitals and users, negotiate individual agreements, but he does influence the agreements that they can chose by putting an overall constraint on the allowed revenue to a hospital for given services. The regulator in this way serves as a *market maker rather than as a proxy consumer*. The role of the suggested regulation is to limit the possibly supernormal profits to the hospital sector as a whole. It hereby affects the settlements that the insurance companies and hospitals can choose from.

There are several effects of this as we shall now demonstrate:

The most obvious effect is to limit the hospitals' *rent extraction*. The regulation improves the bargaining position of the users and hereby the benefit share that they retain. Intuitively, the regulation limits the use of market power that a hospital can exercise as a local monopolist by making it subject to implicit competition by imposing as a yardstick the average performance of the other hospitals.

Another and more indirect effect of the regulation is that it makes some production levels and product mixes more attractive than others. The

regulation is aiming at imposing normal profits for the hospital sector as a whole, not at constraining the profits of individual hospitals and one way to improve profits under a price cap system is to adjust the production level. The regulation hereby has *volume, mix and allocative effects*. We shall develop these in more details below.

It is important to understand also that within the regulation, i.e. for payment below the associated revenue cap function, the market participants are free to negotiate payment plans. There is no reason to restrict the choice of these plans to the one implicitly suggested by the regulator – in fact, there are reasons to avoid using the simple price caps as contracted unit prices. We shall demonstrate this as well.

## 4.2 Cost efficiency effects

Hospital  $h^*$  will supply effort until

$$\frac{\partial v_{h^*}(e_{h^*})}{\partial e_{h^*}} = \frac{\partial c_{h^*}(y_{h^*}, e_{h^*})}{\partial e_{h^*}}$$

i.e. until the marginal disutility of effort equals the marginal cost savings. In the slack model, we see similarly that slack and hereby costs will be minimized, i.e.

$$s=0$$

will be selected as long as  $r$  does not exceed 1.

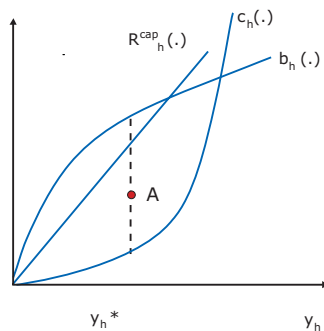
From the point of view of cost minimization, the scheme works well. This is a simple consequence of the high-powered yardstick property – with fixed output, the hospital cannot affect its yardstick and any cost reduction is retained leading to optimal cost reduction incentives. In the sequel, we can therefore assume that  $e$  is set optimally, and we can therefore suppress it and work with costs of a hospital  $h^*$  equal to  $c_{h^*}(y_{h^*}) = c_{h^*}(y_{h^*}, e_{h^*})$ .

## 4.3 Volume effects

In general, i.e. without any further assumptions about costs and benefits, the regulatory constraint may impact not only the profit sharing, but also the production level and production mix that a hospital and an insurance

company agrees on. We shall illustrate this in some details for the case of the production level and briefly discuss the mix effect next. As we will see, the possible distortions will be small in the likely situation of an almost price inelastic demand schedule. Moreover, the distortions will in general be dampened by the bargaining among hospitals and insurance companies.

The constrained bargaining problem from above is illustrated in the symmetric case in Figure 4-1 below and Figure 4-2 on the next page. In Figure 4-1, the unconstrained Nash bargaining leads to a choice of output level and payments corresponding to A that maximize the surplus and a division that do not collide with the regulation. In this case, the regulation will not impact the outcome.



*Figure 4-1 Constrained bargaining with inactive constraint (redundant regulation)*

The exact payment plan in A is not regulated. The regulation does not prescribe how the hospital and the insurance company should implement the payment and the control plan illustrated by A. Needless to say, there are many possibilities. They may for example decided that the insurance company buys a fixed capacity for a fixed price. Alternatively, they may decide that the insurance company pays a fixed fee plus a unit fee for all treatments corresponding to the marginal benefit at the production level in A. In Chapter 5 below, we sketch some classical contracts and discuss their pros and cons in a situation with uncertainty about costs and benefits. A more theoretical discussion is provided in Bogetoft and Olesen (2004).

In Figure 4-2, however, the allowed revenue is less than the unconstrained bargaining outcome A. The regulation would therefore intervene and require the excess charges to be repaid – possibly even with a penalty. Of course, in equilibrium this will never happen.



The parties know the constraints ex ante and will therefore settle for a production level and payment that respect the constraint. One possibility is to simply lower the payment to the hospital by choosing a contract equal to B. That is, the regulation would implicitly give the user added bargaining power. This however is not the only impact. If the parties agreed on a production level and payment like in C, where marginal costs equals the price cap, the profit of the hospital could increase without conflicting with the regulation. The user may still profit from the regulation – although less than in B. Depending on the relative bargaining power of the parties, we the outcome will therefore be between B and C. If the hospital has considerable bargaining power, we will be close to C and vice versa. This is the distortion of production level that we hinted at above.

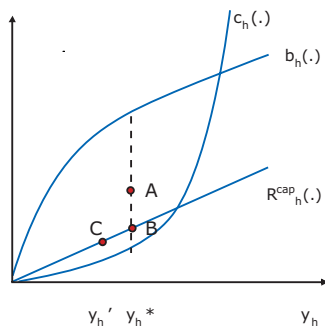
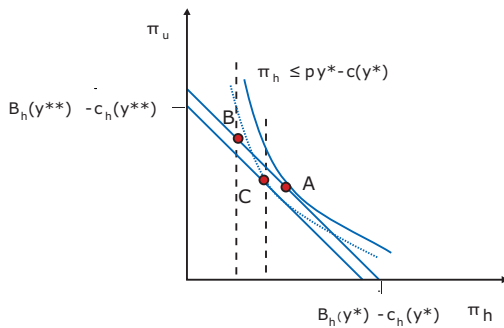


Figure 4-2 Constrained bargaining with active constraint (regulation).

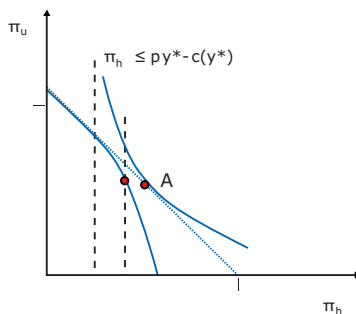
To understand the impact of the regulation, it is convenient to represent the bargaining set under regulation as in Figure 4-3 on the next page. Without regulation, the parties would settle for A corresponding to production level  $y^*$ . This however breaks the price cap constraint. If the parties were still to settle for the socially optimal production level  $y^*$ , it would lead to the point B which make the hospital considerable worse off. An alternative is to adjust production to the socially suboptimal level  $y^{**}$ . The total benefits to be distributed here,  $B_h(y^{**}) - c_h(y^{**})$ , is less – but the hospital gains a larger share in this case. The interesting feature is therefore that some of the inefficient outcomes (in the sense that the total gains from trade are not maximized) does form part of the Pareto efficient frontier and that the optimal solution may therefore involve suboptimal production levels.



*Figure 4-3 Constrained bargaining.*

Figure 4-4 below gives a reduced form representation of the effects of introducing the regulation.

It is clear that the distortions in production levels depend on several factors. The concavity (elasticity) of demand, the convexity of the cost function, and the strength of the regulation affects the shape of the reduced form bargaining set. Also, the choice of solution in the bargaining set is affected by the bargaining powers. It is intuitively clear that distortions are increasing in (i) the hospitals bargaining power, and (ii) the power of the regulation (more high powered = lower required prices). In addition, we conjecture that the distortions will be more severe the more elastic the demand and the less convex the cost function are. The numerical example below illustrates this and give some idea about the effects of the regulation on the quantities supplied.



*Figure 4-4 Constrained bargaining in reduced form.*

Price inelastic demand reduces the distortions and in some cases eliminates them. Figure 4-5 below illustrates the possible impact. For some modest reductions in allowed revenue, say from the unconstrained A to D, there may not be any volume effects since the gains in profits may not outweigh the loss in user benefits. Only when hospital profits becomes even lower, a change in volume may result from the Nash bargaining. Distortions will show up as waiting lists giving information about excess demand.

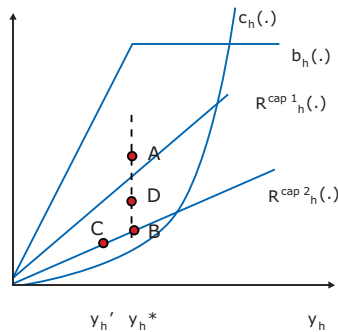


Figure 4-5 Regulation under inelastic demand.

#### 4.4 Mix effects

Another important issue is to ensure that the hospitals provides the mix of treatments that maximizes social welfare. In principle this can be analyzed in a way parallel to the production level above, but it would require the introduction of an extended and somewhat cumbersome notation. Since the conclusions are furthermore trivial to derive given the insight from the analyses of the production level above, we shall abstain from this.

We simply conclude that with price-elastic demand, a distortions in mix is possible if the weights  $w$  are not equal to the marginal costs in optimum. With entirely inelastic demand, however, this is not necessarily the case, i. e. the regulation will not always distort the mix in this case (Sloan and Steinwald, 1980).

It is worthwhile to add that the level and mix distortions are second order effects. In the socially optimal production plan, the marginal user benefits  $B'$  equal marginal hospital costs  $C'$ . This means that the loss from deviating is also marginal. This is a consequence of the so-called envelop

theorem, cf. chapter 5 below. It says that first order deviations in the economic choices only have second order economic impact.

## 4.5 Reallocation effects

It is worthwhile to observe that the level effect may not always be to lower production to keep up the profit of the constrained hospital. The response will depend on the relationship between marginal costs  $MC$  at the unconstrained outcome and the price in the price cap system, i.e. the marginal revenue cap,  $MR$ . If  $MC > MR$ , the effect will be to lower production and if  $MC < MR$ , the effect will be to increase production. This is attractive from an allocative perspective since we hereby reallocate production from the more costly to the less costly hospitals much like in a perfect market. This is illustrated in Figure 4-6 below.

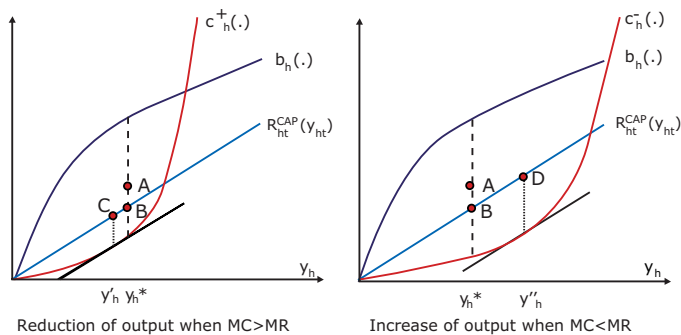


Figure 4-6 Allocation effect

The reallocation process above may lead to overproduction in presence of fixed costs since  $MC < AC = MR$  due to the linear pricing scheme. Setting  $MR = MC$  would violate the participation constraint for the hospitals with fixed costs. The standard solution to this problem is to use two-part tariffs  $R'(y) = K + py$ , where  $K = C(y^*) - py^*$  for a given set of (allowed) prices  $p$  at production  $y^*$ . For the case where the true marginal costs are represented by the allowed prices ( $MC = MR = p$ ) the fixed term is positive  $K > 0$ . If the prices sufficiently overestimate the marginal cost ( $MR > MC$ ) and the demand is exogenously imposed, then the fixed term can be negative ( $K < 0$ ) as in a franchising agreement. However, the two-part tariff faces an important implementation problems in the actual setting. As one of the explicit objectives of the reform is to promote market entry by new operators, arrangements with  $K > 0$  would require an

ex ante statement on optimal scale for any hospital and some guarantee (the ugly princess problem) that the franchisee does not collect the fixed fee without performing the intended services. Under limited liability, when no more than the relationship specific rents can be redistributed, there would no equilibrium in the market structure as the entry/exit incentives are skewed towards hit-and-run. Moreover, as the model is based on the assumption that the aggregate demand for any user (read: insurer) is bounded, the distortion effects due to output expansion are also bounded upwards. Nevertheless, the implementation problem for two-part tariffs does pose a danger for supply-induced demand that must be countered by other measures, e.g. safeguards of the vertical separation between primary diagnosis givers and health care providers as well as some aggregate analyses on the development of overall expenditure.

## 4.6 Demand inducement and moral hazard

Given the previously mentioned local volume and mix distortions, an important question is to what extent the global demand facing the hospital sector is exogenous from the regulation scheme as well as the incentives and market environment that physicians and patients face.

When hospitals and physicians face a high-powered incentive scheme (such as a fee-for-service regulation), *demand inducement* is widely thought to be a prevalent phenomenon. Various theories have been proposed, such as the target income hypothesis, income effects (inspired by the labor economics literature) and revenue targeting from a participation constraint.

The empirical health economics literature (McGuire, 2000), however, shows mixed results about the magnitude of demand inducement. There is a large volume of research that could support demand inducement, but which do not discriminate between the inducement hypothesis and theories with fixed patient preferences. Moreover, most studies are about the marginal effects on demand following income effects on the supply-side. The inconclusive effect on the magnitude of marginal demand inducement therefore does not rule out infra-marginal demand inducement. The no-effect finding in the literature therefore does not rule out the existence of demand inducement. In summary, whatever the current level of demand inducement, the literature is not supportive of large marginal demand inducement following changes in incentives, such as those imposed by the regulatory system described in this paper.

The effects of patient incentives on demands are much clearer (Zweifel and Manning, 2000). The Rand Health Insurance Experiment (Newhouse et al., 1993) showed that health care demand is responsive to cost-sharing arrangements across a broad range of health care. The experiment also showed no significant or appreciable differences on measures of health status (including self-reported health, physical health and mental health) for participants with different health care consumption.

## 4.7 Convergence

An important question regarding the regime is naturally whether the dynamic game between insurers and hospitals, freely establishing contracts and prices for services, will converge to a stable and unique equilibrium.

Beginning with a competitive setting on a perfect market, denote by  $z(p)$  the vector of excess demand at a price system  $p$  at any given time. Varian (1992) states that any price system  $p$  defined over a set of product converges in a finite steps to an equilibrium  $p^*$  if the adjustment rule is defined as

$$p_d = z_d(p) \quad \forall d$$

and the excess demand function holds the weak axiom of revealed preference, i.e.

$$p^* z(p) > 0 \text{ for all } p \neq p^*$$

Departing from the global stability theorem, one may investigate if all of the assumptions are satisfied for the proposed regime.

First, the existence of an excess demand function is based on an axiom with insatiable and monotonous utility (demand) functions. In the case of the health care demand, this is normally approximated with the observation of waiting lists/times for specific services. Although trivially true (the waiting lists would disappear if there was a bidding of each treatment slot), there are a number of complications with the analogy. The existence of a waiting list is already an instrument to manage demand by the provider, not the market maker. Thus, the excess demand may be imperfectly observed and recorded.

Second, the tatonnement process of gradual adjustments presumes that the value of excess demand  $\mathbf{pz}(\mathbf{p})$  represents an incentive to provide additional services or adjust prices. However, under regulation this is not necessarily true. Under price regulation, the convergence obviously is left to the regulator, often in absence of a reliable excess demand function as of above. Under revenue regulation, the regulated enterprise has the right to earn  $\mathbf{R} = \mathbf{py}(\mathbf{p})$  and could theoretically adjust the prices until it fully exploits the demand curve, either to maximize net profit as a constrained monopolist, or minimize effort in the case profit is extracted.

Although the two invoked assumptions are not fulfilled, the yardstick regime has additional features that warrant for faster convergence. On the one hand the revenue regulation is not regulating the prices or conditions between insurers and hospitals. I.e., a hospital cannot apply the monopolist logic referred to above, as there may be other hospitals underbidding it on a set of services. On the other hand, waiting lists,  $\mathbf{z}(\mathbf{p}) > 0$ , are here signals of inefficient pricing that indeed could be addressed by adjusting prices. In addition to firm-level incentives to manage the price adjustment process, such as an improved utilization of resources and higher marginal revenue per service, there is also a second dimension to the problem: the regulator and the weight adjustment process. It follows from the convergence conditions that the regulator should take into account sector-wide wide proxies for excess demand (waiting lists, artificially long treatment times, etc) in the revision of the DBC weights as to facilitate the capacity adjustment process and the incentive value of the weights.

## 4.8 Robustness to diverse firm preferences

We have so far assumed that the hospitals are profit maximizers. We note however, that this is not necessary for most of the properties derived above. What is important is

- Do they seek to maximize the charges they make  $R_n$  ?
- Do they seek to minimize costs?
- Do they weigh cost reductions and revenue increases equally?

The maximization of charges hold quite generally. A hospital driven by other motives, like slack maximization, effort minimization, maximization of resource to be spend on non-verifiable quality etc will – like a profit maximizing firm - seek to get the highest possible payment for its given services.

It is of course possible to have hospitals that not seek to maximize the charges. Cooperatives are the obvious theoretical exception to the charge maximizing behavior. A cooperative owned by the clients should in principle minimize its charges. In reality, however, they may nevertheless charge much the same – by simply substituting slack or excessive costs for compensations to its investors. For many types of hospitals, we therefore suggest that it is reasonable to assume that the bargaining between the hospitals and the insurance companies as well as the discipline hereon from the price cap will be analogous to the case of profit maximizing firms.

A desire to minimize costs is also reasonable to assume in many case. Profit maximizing firms are the prime example, but a user owned cooperative should in principle behave similarly on the cost side.

In summary, we suggest that hospitals in general, and not only profit maximizing hospitals, will try to increase charges and to lower its costs of given procedures. This is the crucial behavioral characteristics focused at above.

The last behavioral characteristic, assigning equal importance to increases in revenue and decreases in costs, is important for the levels of production chosen and for the speed by which adjustments towards a competitive outcome is attained. We suggest that these effects are second order effects compared to the main properties of the system. This is also the behavioral characteristic that is hardest to justify across ownership types etc. A non-for-profit organization may not want to restrain production to a level, where profit is maximized. Likewise, a cooperative may end up overproducing or under producing depending on the details of the situation, cf. e.g. Bogetoft and Olesen (2006).

Finally, as long as all firms are perfectly profit maximizing with an identical cost function, an ex ante cost-based revenue cap with ex post volume adjustment is functionally equivalent to the ex post DBC-adjusted revenue yardstick. If, however, the set of firms contains firms with other objectives, such as revenue maximization or maximum capacity utilization, their pricing decisions may not correspond to those of a rational profit maximizer. In particular, the fact that most hospitals currently are non-for-profit is likely to have an impact on their way of pricing actual and future services. The revenue logic is then superimposing the objectives of the least (current) cost provider upon the sector, although this may not necessarily be the long-term least cost policy. Moreover, without knowing the cost structure, it can also be difficult to identify which agents are profit-maximizing and which are not



in order to possibly correct the revenues. As will be discussed below on innovation incentives, the lowest current revenue may not be enough to guarantee sufficient process development in the future.

## 4.9 Robustness to heterogeneous cost functions

Although the model assumes an identical sector cost function, the scheme does not deploy an information revelation mechanism that crucially depends on this assumption. In particular, one may extend the scenario to local cost functions  $C_h(\cdot)$  that depend on local factor prices and varying fixed costs related to the assets (e.g. buildings). If the factor prices follow some regularity with respect to the input basket, the regime could be implemented with price indexes  $K_h$  to represent local cost differences (e.g. related to exogenous data on salary costs, land prices, etc). An alternative would be to estimate perform a stochastic production function estimation and to regress the detected cost inefficiency on some indicators or location, environment and market. The latter approach cannot be directly used in the regime, as it is only partially capable of detecting the direction of the causality involved (high cost areas or pockets of inefficiency). However, it could be instrumental to detect the relevance of, and plausible factors for, local cost function correction.

## 4.10 Information requirements

The advantage with a revenue-based yardstick mechanism compared to a cost-based lies in the lower information requirements, since the regime does not need to determine costs for the activity. In a competitive market, the final price, not the accounting operating cost, is the pivotal signaling device towards clients. Thus, firms may choose different ways of internalizing investments, innovations, managerial incentives and financial structure that manifest themselves as different accounting costs, although the user value, i.e. the price, may be the same.

However, the autonomy of the revenue-based scheme from cost-level data is not as perfect in a practical implementation. This is due to primarily three problems: the price adjustment mechanism, the diversity in firm-level objectives and capital valuation.

First, a revenue-based yardstick can be implemented using either fixed (or only downwardly revisable) or free service prices.

In the first case, the viability of the system relies on the absence of systematic cost shocks (that would violate the IR conditions, i.e. non-negative expected profit) and/or a frequently applied updating system (which will reintroduce the cost aspect).

In the latter case, the prices of last period are used in the updating. If now there is no penalty for exceeding the revenue cap, the pricing adjustment does not converge for a situation of inelastic demand since there is no incentive to decrease prices. In the case of sufficiently large repayment penalty (cf. Agrell and Bogetoft, 2005), there exists an equilibrium such that the prices adjust to the same level as for an equivalent cost-based yardstick. The management of the penalties must be careful as both too low and too high values will give rise to distortions.

Moreover, the price adjustment mechanism of the revenue-based yardstick is more sensitive to collusion than the cost-based yardstick. In the revenue-based system, the historical price level and the relative DBC weights act as focal points in the price setting game. Given this focal point, individual hospitals have no incentive to deviate their pricing behavior (though they might adjust their costs, this has no consequences for the update the yardstick). For the cost-based yardstick, the analogous focal point is unstable against individual costs adjustments (for which there is ample incentive).

Second, the previously discussed existence of heterogeneous firm-level objectives, including distorted capital costs from weak political principals, could jeopardize the sustainability of the ex post revenue-based yardstick under mixed ownership in the sector. Since distorted capital may be scarce and the principals unwilling to expand to cover the shortfall in capacity due to exiting firms with full capital costs, the excess demand signal may be insufficient to reach a feasible coverage. To counter this effect, the regulator may monitor the revenues or the financial structure of the firms as to detect non-competitive behavior, in essence a return to a heavier information system requirement.

Third, the institutional settings may differ in terms of capital valuation and financing. Assuming a budget/cost-based pricing scheme, the incumbent capital valuation will have a direct influence on the revenues. However, it is well known in regulation that accounting-based capital measures suffer from serious drawbacks for pricing related to their temporal and fiscal limitations. Historic acquisition values may be a poor prediction of future capital intensity in an activity, also for reinvestments and/or associated staff costs. The capital cost may also be influenced with elements from the prior point, e.g. institutional and ownership

involvement. A hospital that is integrated in a larger structure of publicly owned enterprises may enjoy artificially low capital costs due to passive owners and low risk rating due to soft budget constraints at group level.

## **4.11 Discussion**

In summary, the yardstick principle assumes a homogenous task description, a standardized cost definition and homogenous firm objectives. The revenue oriented yardstick mechanism is an alternative for sectors with long tradition of private and dispersed ownership, where the competitive behavior is well established. In a setting like the incumbent Dutch situation, the avoidance of cost estimation might jeopardize the fundamentals of the regime, i.e. the feasibility of continued operation at the given level of reimbursement for a privately owned market-oriented operator.

The regime is analogous to similar approaches used in regulation of complex multi-product services with high societal importance, e.g. the ECOM+ model utilized by Austria, Denmark, the Netherlands, Norway and Portugal in the regulation of transmission system operators (cf. Agrell and Bogetoft, 2004). Both models use a detailed output space to capture the diversified services of the providers and simplifying assumptions such as constant relative prices and constant returns to scale in order to arrive at a scalar efficiency norm. Experience shows that the ECOM+ norm due to the rich output description is robust to errors in the weights of individual services and to technical assumptions related to e.g. life times of assets (cf. Agrell and Bogetoft, 2006). An interesting empirical finding is also that the operators under regulation with ex ante relative output prices have shown interest to reveal private information with respect to output valuation in order to obtain relative advantages. This reduces the information acquisition burden of the regulator and contributes under some mild assumptions on the information structure to faster convergence in the price space.



## 5. Quality and innovation incentives

In the main analysis, we have focused on the proposed mechanism as presented in preliminary documents. However, as the contractual and institutional framework is likely to evolve in parallel with the economic reform process, we also present some comments on alternative contract forms between clients, insurers and hospitals with respect to a particularly central issue: the incentives for innovation and quality provision. We begin with the most straight-forward issue related to innovations in cost-reductions and service enhancements, later to proceed to a more in-depth discussion of quality provision and its regulation.

### 5.1 Incentives for innovation

In a dynamic setting over several periods, the incentives for innovation become important for the welfare effects of the mechanism. To the extent an optimal rate of innovation and research are promoted (demoted) by the regulation, the welfare effects will be positive (negative). Not to complicate the presentation with a full dynamic model, we will limit the discussion here to the qualitative findings regarding information generation. Initially, we will distinguish two dimensions of analysis and then proceed to look at them systematically.

First, only innovations related to process or service improvements can be taken into account in the analysis. Their effect over some horizon is to either lower the cost of a given set of services, or to extend the set of eligible services.

Second, the main characteristics of information economics related to innovation is whether the information is generated as a *common goods* property or not. In our context, this would correspond to the capacity of the agents to protect the generated processes and services from competing agents. In a competitive setting, the common goods property would induce underinvestment in research, as the benefits cannot be fully internalized by the agents. However, in a non-competitive setting (such as regulation), offering extended property rights to information may actually lower welfare if markets for information are imperfect and firms are prohibited to extend their services.

### **Cost-reducing innovations**

Assume that a hospital may undertake an investment  $I$  in research that is expected to lower the annual costs for bundle  $y$  from  $C(y)$  to  $C'(y)$ . Irrespective of revenue cap  $R(y)$ , the decision to undertake or not the investment is then purely based on the rule

$$I < (C(y) - C'(y)) \frac{\delta}{1 - \delta}$$

where  $d$  is an annuity factor to represent the value of time. (The specific risk related to R&D, i.e. the probability of failure or no effects is here incorporated in the cost effects). Next, assume that the sector will be learn about and be entitled to use the information from year  $N$ . Under yardstick regulation, this actually induces endogeneity into the mechanism, since the firm will have to take into account the effects on the revenue side by their disclosed information before undertaking the investment. Assume that the expected costs for the other agents will decrease from  $C_h(y)$  to  $C'_h(y)$  from year  $N$ . Ceteris paribus, the investment rule is now:

$$I < (C(y) - C'(y)) \frac{\delta}{1 - \delta} - \left[ \frac{w y}{\sum_{h \neq h^*} w y_h} \right] \sum_{h \neq h^*} \{C_h(y_h) - C'_h(y_h)\} \frac{\delta^N}{1 - \delta}$$

where the second term lowers the threshold for the investment value, i.e., increases the required payoff during the proprietary period of investment. If the information has common goods properties, i.e., the innovation is instantly perfectly observable and costlessly imitable by the other agents ( $N=1$ ), then there is no incentive to undertake innovation, as the gains are washed out in the first period. Adding a cost for adoption of rival innovations changes the conclusions only if it is higher than the investment cost for the investment, net of any protected period earnings.

The yardstick logic mimics the competitive markets with respect to cost-reducing innovations: their frequency depends on the period for harvesting the benefits and the ease by which existing innovations can be copied. In the hospital setting in the Netherlands, where the managerial focus so far has been less cost-driven, one might expect positive initial incentives for innovation in a redistributive setting. Hospitals that are actively pursuing both process innovation and process learning (e.g. copying in our context) are likely to achieve temporal gains, i.e. a proportional higher share of the sector revenues.

Under a low-powered regulation regime, there is no incentive to undertake cost-reducing investments. Although  $C(y)$  is refinanced by the

users, the investment I may be subject to either equity financing or capital rationing, which naturally causes a distortion towards maintaining the expenditure. If the investment I is interpreted as the effort (coordination etc) involved in information acquisition activities related to learning, we see also that the incumbent regime does not provide any incentives to refrain from internal activities and to engage in learning best practice.

### Service innovations

The second class of innovations aims at product/service differentiation, by which the product space is extended. As opposed to the earlier scenario with respect to cost-reducing innovations, the regulatory weights  $w$  will have an impact on the incentives in this case. The innovation decision is here driven by the revenue side, viz. the output characterization. Assume that a firm can undertake an investment I in R&D as to achieve an expected differentiation in output from  $y$  to  $(y', y'')$ , with cost  $C(y)$  and  $C(y', y'')$  respectively. The decision rule for investment in isolation is

$$I < \left\{ \left( \frac{w' y' + w'' y''}{\sum_{h \neq h^*} w' y y_h} - \frac{w y}{\sum_{h \neq h^*} w y_h} \right) \sum_{h \neq h^*} c_h(y_h) - c(y', y'') + c(y) \right\} - \frac{\delta}{1 - \delta}$$

where the new weights are denoted  $(w', w'')$  after innovation. Analogous to the previous case, the time of exclusivity (barriers for imitation) will have an effect on the investment threshold. The longer time (higher barriers), the lower requirements put at the regulator to reward the investment through the weights. Note that the incentives here are primarily redistributive: the innovating firm hops to get a higher share of the sector's cost by extending the output vector before the competing firms. However, the analysis also requires item-line rents, i.e., the any possible substitution between existing services  $y'$  and the new services  $y''$  must be offset by the increase in value-added:  $w'y' + w''y'' - wy - C(y', y'') + C(y)$ .

The conclusions in this respect reveal differences from competitive markets as firms are constrained in their overall revenues. Rather than collectively expanding the output space, the service-innovating firms in healthcare are acting preemptively to safeguard their share of the cake. As research and development not only are costly, but also risky, the weights should take into consideration the necessary premium for innovation, given the specific barriers of imitation for the service. If the weight is set based on the cost for realized investments, the costs for (invisible) failed investments will not be recovered, which lowers the incentives below the competitive (unconstrained) level.

For comparison, contrast the yardstick regime with the incentives for service innovations under a low-powered regime. As long as capital is rationed or subject to equity financing, the innovation in new services is disincentivized by the cost-recovery regime since it may lead to lower costs and/or higher private costs for the same potential output. Clearly, the cost-plus regulation makes the individual agent independent of possible learning by other agents and the subsequent impact on sector costs. The incentives for the service innovation process are skewed towards introduction of new services that either are more expensive (more advanced) than previous products or that are complements (output expansion). Development of new cost-efficient substitute treatments are directly demoted. We note that the new service incentives under low-powered regimes are far from those of competitive markets and that there are risks of “overtreatment” rather than the opposite.

## **5.2 Quality provision incentives**

Quality has traditionally been handled through the imposition of a system of compulsory and a system of suggestive minimum standards. Coupled with a tendency to rely on medical reasoning, this has led to very high quality standards in Western Europe. This, however, may neither be optimal nor the long run competitive equilibrium. First, the cost of ensuring the present high quality level may exceed the benefits and the present quality level, although certainly high enough, may actually be too high and too costly. Second, any change in the regulatory approach will change the behavior of the agents. In particular, a movement towards a more high powered arms-length incentive regulation will induce the hospital to focus more on cost minimization with a possible adverse effect on quality.

### **Optimal quality level**

The basic underlying problem is to determine the optimal trade-off between the costs of producing higher quality and the benefits derived from it. This problem is illustrated in Figure 5-1 on the next page. Here the cost function is the cost to the firm of changing its present quality level. The benefit function  $B(q)$  is similarly the gains to the consumers from changes in the quality level.



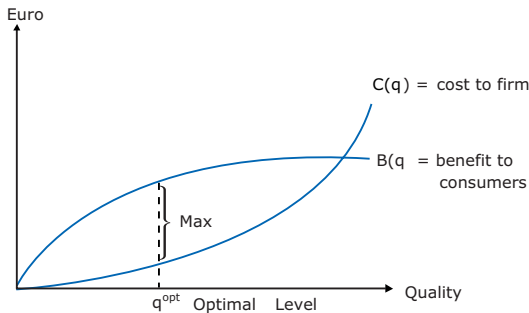


Figure 5-1 Optimal quality level

The optimal level leads to the largest difference between costs and benefits (i.e. marginal benefits equal to marginal costs).

### Information and Strategic Behavior

In reality, the regulator knows neither costs nor benefits a priori. He must therefore try to reveal information about these aspects from the firms and the customers. This raises the problem of *asymmetric information and strategic behavior* since the firms may want to exaggerate costs to get a higher compensation and the customers may want to underplay their true values to pay less. Faced with these problems, the regulator should not strive for so-called first-best solutions as illustrated above. Rather, he must settle with second best solutions or – if he takes into account broader systems costs like administrative costs - third best solutions.

In an incumbent low-powered regime, the hospital has incentives to increase the cost of quality above the socially optimal level  $C(q^{\text{opt}})$ . As this is valid in aggregate too, the provider has greater incentives to increase the cost for relatively more expensive treatments, irrespective of the possible decreasing returns in welfare. Since the situation has no equilibrium under soft budget constraints, the regulator must then impose ad hoc budgets to the hospitals to allocate over a set of services and quality levels. A full modeling of the behavior of the firm under the low-powered regulation is beyond the scope of this paper, but it suffices to hypothesize that the firm in a one-period game would select an allocation policy that would minimize effort while making the budget constraint binding. In a two-period model with possible revisions upwards for detected shortages, one can also show how hospitals can allocate their budgets as to achieve too high quality levels for subsets of the output

space while rationing in areas that are likely to increase the probability for further increases in funding. Hence, although quality as such is promoted under low-powered regimes, the net effects on social welfare are ambiguous, even when only considering the output mix and quality level decisions.

### **Collective or individual qualities**

From a regulatory point of view, a crucial question is if it is possible and / or desirable to have different quality levels across insurers and hospitals. This leads to the distinction between:

- Collective regimes
- Individual regimes

In the *collective regime*, all customers enjoy the same quality level – or at least the same minimal level. The regulator works as a proxy customer and he imposes quality standards with respect to all insurers as well.

In the *individual regime*, the regulator allows the insurers to demand and the hospitals to supply different qualities to different customer groups. The terms may be settled through bilateral negotiations among the firms and the insurers. In a secondary market, the insurers may then promote their conditions through differentiated enrollment fees.

Whether to regulate a given quality dimension in an individual or a collective scheme depends on *technical aspects*, *economic implications* as well *regulatory preferences*.

Based on fairness criteria etc, the regulator may prefer the universal provision of quality even though it may be possible to provide different quality levels to different users.

From the point of designing quality regulation, an important first step is to choose which aspects of quality should be governed by a collective or an individual regime. To do so in a systematical way, it may be useful to establish even a rough evaluation of the cost of differentiated quality and the preferences for such differentiation. While the former depends on primarily technical aspects the latter depends on both the possible differences in the consumers demand functions and the general preferences of the regulator. Table 5-1 on the next page illustrates parts of such an evaluation.

Table 5-1 Evaluation of quality dimensions wrt cost, value and mode

Quality dimension	Cost of individual regime	Value of individual regime	Choice of regime
Mortality	Prohibitively high	Some	Collective
Case information	Limited	Some	Individual

Figure 5-2 below illustrates an instance where there are significant differences in the demand for quality from different consumer groups and where it may therefore be worthwhile to introduce such variations depending of course on the costs.

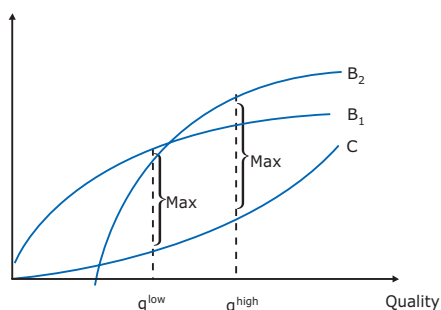
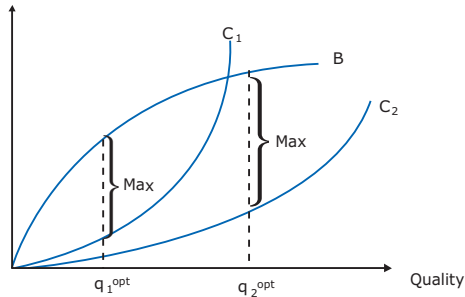


Figure 5-2 Different quality to different consumers

All of what has been said above can – with obvious modifications – be repeated in an inquiry as to the desirability of allowing differentiated quality levels from different hospitals. An instance suggesting high social benefits of differentiation provided the regulator does not have a strong desire for fairness or equality at all costs, is illustrated in Figure 5-3 on the next page.



*Figure 5-3 Different qualities at different firms*

### Equality Concepts

A common objective among stakeholders seems to be that different consumers should be treated equally and fairly. This is also related to the universal service obligations. Unfortunately, it is not clear what fairness means.

To illustrate the problems, we note that an equal treatment of different consumers may mean at least three different things:

- All are given the same services for the same price
- All have the same possibilities to choose among different service-price combinations
- All get the same improvements

To understand the differences a restaurant analogy may be helpful. The first interpretation corresponds to the restaurant quests being served the same meal and paying the same price. The second fairness notion refers to a situation where all customers are handed the same menu. The third situation could refer to a situation with income differentiated prices and products.

It is beyond the scope of this report to define exactly what fairness means. We do suggest however that an important question in the design of a regulatory scheme, including the regulation of quality and the fee structure, is to analyze and agree on some more precise notions of fairness.

### Bundled or Separated Regulation

Another important general question in the design of a regulatory framework for quality is the extent to which the regulator should implement different dimensions jointly as a package deal or separately as a series of individual regulation problems. Given the information problem

and the likely strategic behavior of firms and consumers, the joint implementation has advantages. By bundling the dimensions, the parties can be induced more easily to reveal the underlying costs and benefits, cf. e.g. Antle, Bogetoft and Stark(1999).

#### *Comprehensive or Partial Regulation*

A principal question facing the regulator is whether to integrate the quality dimension into the price regulation framework to form a comprehensive model of the costs of providing different levels of different qualities of output. Theoretically, this would be the ideal solution but practically, this may lead to dimensionality problems in the estimation of the resulting complex and detailed benchmark model.

A more realistic approach is probably to think of the price regulation as being conditioned on certain minimal standards and then to allow the regulation of quality to be undertaken via one or more partial add-on models of the cost increases (decreases) that will be allowed for certain increases (decreases) in quality. This is the approach we shall discuss here. What is forgone by this approach is the possible interaction of quality and quantity and the possible gains from bundling quality and quantity signals.

#### **Implementation**

Given a reasonable amount of information about costs and benefits, the (near) optimal quality level can be determined. The natural next question is how the regulator can steer the firms (or consumers) to choose these levels. There are several such ways and in this chapter we outline some important ones and discuss their pros and cons in the context of uncertainty and asymmetric information. The methods can be used in an individual as well as in a collective scheme. We emphasize the steering of the firms but we note also, as we shall return to, that similar steering of the customers is possible via demand management schemes.

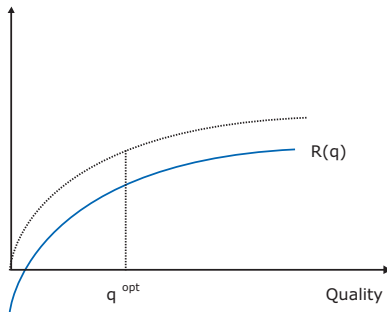
### **5.3 Steering mechanisms**

We shall now outline four principal ways in which the hospital may be incentivized to adjust quality towards the socially optimal level, the qopt in Figure 5-1.

One possibility is to use a *generalized price plan* where the firm is reimbursed an amount  $R(q)$  equal to the consumers benefit  $B(q)$  minus a lump sum (quality independent) payment  $A$ :

$$R(q) = -A + B(q)$$

The lump sum amount  $A$  can be chosen as any value between 0 and  $B(q^{opt}) - C(q^{opt})$ . High values means that all the gains from adoption to optimal quality goes to the consumers and low values means that the gains go the firm. This scheme is illustrated in Figure 5-4 below.



*Figure 5-4 Generalized price plan*

The generalized scheme is advantageous by leading to optimal quality levels for all possible cost functions. The regulator do not need to know and constantly track changes in the costs function except to determine the exact range in which  $A$  can be chosen. On the other hand, the regulator needs considerable information about the benefit function. To collect such information, the regulator may undertake willingness to pay and consumer choice studies where a number of consumers are asked how much they are willing to pay for improved quality and how they would choose in some hypothetical choice experiments. There is a considerable literature on the design of such studies and a large body of practical experience, in part from the marketing science. Still, the collection of information about  $B(\cdot)$  may be a non-trivial task. Moreover, it may be difficult to communicate especially in the multiple dimensional case.

A second possibility is to use a so-called *two-price scheme* where the firm is paying a lump sum amount  $A$  for the right to make quality decisions plus a relative high price for quality improvements, when quality is low, and a small price for quality improvements when quality is higher:

$$R(q) = -A + p_1 q - p_2 \max\{q - q^{opt}, 0\}$$

This scheme is illustrated in Figure 5-5 below.

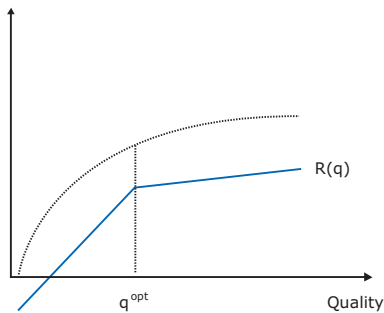


Figure 5-5 Two-price plan

The advantage of this scheme is its relative simplicity making it easy to communicate and to adapt to. Also, the outcome is less sensitive to changes in costs and benefits than the restriction based approach.

A third possibility is to use a so-called marginal-price scheme where the firm is paid a lump sum amount  $A$  plus a relative small price for quality improvements equal to the marginal value to the consumers in optimum:

$$R(q) = A + pq$$

This scheme is illustrated in Figure 5-6 below.

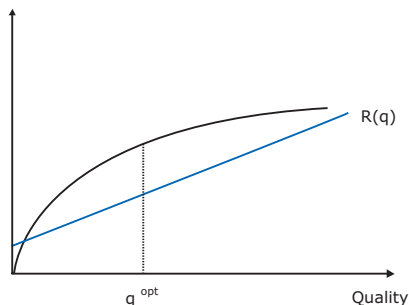


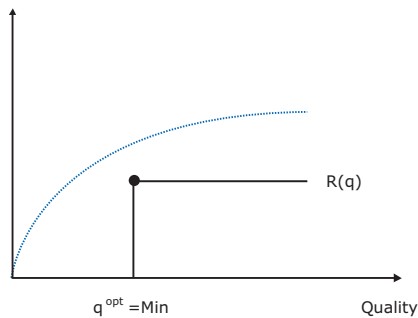
Figure 5-6 Marginal price scheme

The advantage of this scheme is its relative simplicity making it easy to communicate and to adapt to. Also, the outcome is not too sensitive to changes in costs and benefits. On the other hand, the estimation of marginal value in optimum must be rather precise.

The final possibility we will consider here is to use a *restriction based plan* similar to the familiar minimal quality requirement approach in health care provision. In this scheme, the reimbursement to the firm equals A if it comply with minimal standards and the penalty otherwise is very large

$$R(q) = A \text{ if } q \geq 0 \text{ and very negative otherwise}$$

Again, the lump sum amount A can be chosen as any value between 0 and  $B(q_{opt}) - C(q_{opt})$ . High values mean that all the gains from adoption to optimal quality go to the firm and low values that the gains go the consumers. This scheme is illustrated in Figure 5-7 below.



*Figure 5-7 Restriction based scheme*

The advantage of this scheme is its simplicity making it easy to communicate and to adapt to. On the other hand, its optimality is extremely sensitive to variations in the cost and benefits function. It is therefore primarily useful in those cases, where the benefit or cost curves are linked with a sharp decrease in marginal value or a sharp increase in marginal costs at  $q_{opt}$ .

All the schemes sketched above involved some *lump sum* payment, denoted A. The size of this payment depends both on the way the non-quality revenue model is calibrated and on the way the gains from quality adjustments shall be distributed between the hospitals, the insurers and the consumers. By and large, however, the incentive effects are not dependent on A and we shall therefore leave the problem of setting A for future more detailed studies.



To illustrate the idea here, it suffices to note that if, for example, we assume that the reimbursement for the non-quality dimensions presumes a given minimal quality level, the quality payment schemes shall ideally be interpreted as penalty or bonuses for deviations from these minimum levels. This means that  $A$  shall be chosen such that the quality payments are 0 at the minimal levels.

### Robustness to changes in costs and benefits

All the schemes above require information about benefits and –except for the generalized payment plan – costs. Since such information is noisy at best, it is important in the choice of regime to consider the impact of having misspecified costs and benefits – or to have changes in costs and benefits over time. We have already indicated that the generalized scheme is the most robust to changes in cost structure and the restriction based among the least robust schemes in this respect. Two general economic results may shed further light over this question.

The first, sometimes known as the *envelope theorem*, suggests that first order deviations in the estimation of economic choices may only have a second order economic impact. In the present case, let

$$N(q) = B(q) - C(q)$$

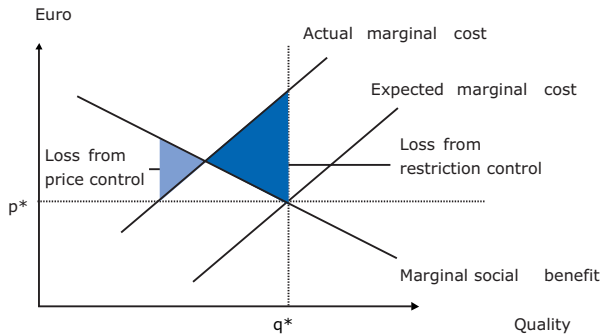
denote the net benefit and let us assume that we have estimated the optimal  $q$  to  $q^*$  rather than  $q^{opt}$ . Assuming differentiability and making a so-called Taylor approximation of  $N(q)$  we get

$$N(q^*) = N(q^{opt}) + N'(q^{opt})(q^* - q^{opt}) + 0.5N''(q)(q^* - q^{opt})^2$$

for some  $q$  between  $q^*$  and  $q^{opt}$ . Since  $N'(q^{opt}) = 0$ , we see that the difference between  $N(q^*)$  and  $N(q^{opt})$  will not be too large unless the net-benefit function is strongly curved. See also Akerlof and Yellen (1985).

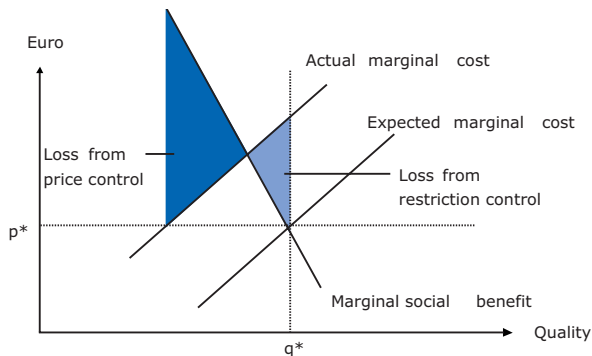
The second set of results concern the relative merits of the marginal price and the restriction based methods. Varying the benefit function has the same impact in all regimes since the signal sent to the firm and therefore its behavior is fixed. Quality adaptations to changes in the cost function, however, are severely affected by the regimes. In the generalized regime, optimal adaptation is obtained – at least as long as the lump-sum payment is set to 0. The two price system works reasonably as well, although of course not as well as the generalized scheme. To choose among the price and restriction based approaches, we need to consider the elasticity of supply and demand. When the demand is rather elastic, i.e. the marginal benefit curve is relatively flat compared to the supply

curve, i.e. marginal cost curve, price regulation leads to the smallest losses. This is intuitively natural since in this case it is of particular importance to take into account the costs. Figure 5-8 illustrates this.



*Figure 5-8 High elasticity of demand.*

Similarly, when the demand is rather in-elastic, i.e. the marginal benefit curve is relatively steep compared to the supply curve, the adaptation to costs should play a smaller role and therefore minimal standards are superior. Figure 5-9 below illustrates this case.



*Figure 5-9 Low elasticity of demand control*

### **Decision rights to the best informed**

Traditionally, quality decisions are delegated to the hospitals and indeed this is the perspective we have used in the discussion of implementation above. This is particularly relevant when we consider common regimes

where all users are going to enjoy the same quality level. For other quality attributes however it is possible to let the users decide. The four implementation mechanisms above can conceptually be turned around to cover user based implementations. In such cases, the regulator should set up payment schemes or price plans stating what the consumers should pay for different level of required quality. Even the restriction based method can be used in this case. It would be a maximal allowed quality requirement below which the consumers can choose.

A key question in the allocation of decision rights is who has the best information. If the costs are relatively stable and foreseeable but the benefit structure is hard to elicit, the consumers should be allocated the decision rights and they should pay a lump sum for this right. If on the other hand benefits are relatively well described but costs are complicated and likely to vary over time, the firm based regime is preferable.

## 5.4 Summary on R&D and quality

The incentives for research and development activities under the proposed regime are different with respect to the objective of the innovation. For cost-reducing innovations, the regime is a robust pseudo-competitive instrument that provides both carrots (frontier-shift) and sticks (catch-up) for firm-level R&D. In this regard, it is the high-powered quality of the regime that provides the correct incentives, which are likely to provide positive behavioral signals to the decision makers. For the service or quality innovations, the conclusions depend on the exact updating procedure applied for new services and the policy adopted for potentially protecting proprietary methods. As innovation inherently is risky, a retrospective output-valuation based on actual costs in combination with an absence of information property rights is likely only to promote imitation (adoption) and to demote innovation. Adequate risk premiums in the weight settings, potentially limited in time, could mitigate this effect.

Quality provision incentives can be modulated in several ways, each of which pose specific requirements on the contractual partners involved: the hospitals, the insurers and the clients.

In theory, it would be ideal to integrate the quality and quantity regulation into a comprehensive regulation. However, the complexities involved related to information access on client status and treatment outcome are likely to be prohibitive. Thus, a separated regime could be more effective

as long as it provides a credible threat to penalize quality deviations either through direct adjustments of the revenue cap or through reclassifications of the services rendered.

In healthcare provision, one can also argue that an ex post yardstick regime is superior to an ex ante regime. Treatment outcomes and procedures are observable ex post and this information can be used to incentivize current behavior. We therefore believe that the current proposal stands well with respect to quality regulation.

Lastly, considering the specific choice of contractual form, we have sketched four possible schemes above, viz. the generalized price plan, the two-price plan, the marginal price plan and the restriction based plan. In a world of perfect information, it really does not matter which of the approaches one chooses. In a world of uncertainty and asymmetric information about production costs, however, the simple marginal cost scheme seems to have particular attractive properties. The incentives are not too dependent on precise and reliable information about both costs and benefits. Moreover, it leaves the adjustments to the insurer-hospital negotiation that is presumably better informed about the costs of quality. Lastly it is simple and easy to communicate.

## 6. Conclusions

In this paper, we have provided a formalized description and analysis of the proposed yardstick regulation for Dutch hospitals. The model clarifies how the regulation is envisaged to work – and how this interacts with the objectives of users, insurance companies, hospitals and the regulator.

The overall findings on the investigated properties are summarized in Table 6-1 on the next page. The main advantages, the clear cost reduction incentives as well as the incentives to undertake cost reduction innovations, are common for the two alternative formulations, being consequences of the high-powered character of the regulation. However, the transfer of rents from hospitals to users and limitation of the bargaining power of the hospital sector which is important in the transition period, are more clearly identified for the cost-based regime. The revenue-based yardstick leads to multiple equilibria for the rent sharing with the efficiency depending on whether past prices are focal points and the responsiveness of the pricing management to changes in the cost structure.

We have identified also some potential drawbacks of the systems. Most importantly, the system may in theory distort production level and production mix. In the Dutch context however, we suggest that total demand will effectively be inelastic, and the local distortions will enable a reallocation of production to more efficient producers. In that case, and taking into account also that these distortions are second order effects, we conclude that the level and mix effects will not be too harmful and that the reallocation effect will be beneficial for obtaining a competitive market outcome. The volume and mix distortions occur under both formulations, but the uncertainty for the firms in the revenue-based yardstick may make the reallocation process less effective.

The innovation incentives will approach those of competitive markets under imperfect information rights, with a consequent undersupply of innovation. However, the incentives for learning and catch-up are high, which in practical settings when there are externalities between innovation, testing and development activities will mitigate the problem. A further solution to the service innovation supply would be to offer separate contracts to able provider-hospitals (e.g. academic research hospitals) to engage in new product development without retaining the information rights. Combined with fast learning, the resulting welfare effects might be higher than for dispersed innovation and low learning in

an incumbent low-powered regulation.

As concerns robustness to collusion, the cost-based formulation has the advantage of using ex post final audited cost, which is more difficult to coordinate in a collusive agreement. Revenues and prices are naturally more prone to ex ante rigging, particularly in a situation where there exists a set of relative DBC weights from the regulator that can act as a focal point of stipulated “prices”. The revenue-based regime is also prone to distortions from diverging firm preferences, such as benevolent labor, public capital and cooperatives. Firms with diverging objectives may lower the compensations for investor-owned firms below their participation constraints, at least blocking entry to the sector. The cost-based ex post regime is also more robust to systematic cost shocks for the sector, being picked up and corrected with a lag of one period. The revenue-based regime is here dependent on the consistency of pricing decisions to correctly identify the systematic from the idiosyncratic cost drivers in the annual adjustments, which may be hard in a real situation.

In terms of information and administration, finally, the two formulations do rely on a fairly heavy investment in the DBC reporting system and the periodic updating of weights. The cost-based regime has in addition the calculation of standardized costs for the hospitals, which the revenue-based regime does not need. However, the administrative footprint of the cost-based yardstick scheme might be substantially mitigated depending on whether the regulatory audited costs are also necessary for other tasks of the NZa, or for information provision to stakeholders and shareholders of hospitals (e.g. compliance with cost-level data requirements of the security exchange regulator for stock-listed companies).

Table 6-1 Evaluation of the properties for the price-cap and the revenue yardstick

	Type of revenue-cap	
Criterion	Cost-based price cap Ex ante	Revenue yardstick Ex post
Cost-efficiency	High incentives	High incentives
	Full efficiency	Full efficiency
Production level and mix distortions	Distortions depend on cost efficiency and strictness of regulation	Distortions depend on cost efficiency and strictness of regulation
Rent extraction	Rent reduction driven by profit	Multiple equilibria
Incentives: process innovation	High incentives for learning	High incentives for learning
Incentives: new services	Undersupply compared to first-best	Undersupply compared to first-best
Incentives: quality provision	High if redefined product, low otherwise	High if redefined product, low otherwise
Robustness to collusion	High (ex post verifiable cost)	Low (revenues and price)
Robustness for diverse hospital objectives	Mitigated by standardized costs	Low (deviant pricing will affect the market)
Robustness to systematic cost shocks	Compensation ex post	Distortion downwards unless (unlikely) common change
Administrative cost	High (cost data + DBC+waiting lists)	Medium (revenue + DBC+waiting lists)

The convergence of the pricing information in the model relies upon some assumptions regarding the exogeneity of demand and the observability of excess demand. It is important that subsequent reforms of this and the adjacent sectors respect and honor these principles, e.g. in terms of the vertical separation between the gatekeepers, the insurers and the care givers.

## 6.1 References

- Agrell, Per J. and P. Bogetoft (2004) *Note on Methodology: ECOM+ Project, Report*, SUMICSID AB.
- Agrell, Per J. and P. Bogetoft (2005) *Regulatory mechanism design: NEMESYS Subproject B*, Final Report, SUMICSID AB.
- Agrell, Per J. and P. Bogetoft (2006) *Results 2005: ECOM+ Project*, Final Report, SUMICSID AB.
- Akerlof, G.A., and J.L. Yellen (1985) Can Small Deviations from Rationality Make Significant Differences to Economic Equilibria, *American Economic Review*, 75 (3), 708-720.
- Antle, R., P. Bogetoft and A. Stark (1999) Selection From Many Investments With Managerial Private Information, *Contemporary Accounting Research*, 16, 397-418.
- Baron, D., R. Myerson (1982) Regulating a Monopolist with Unknown Costs. *Econometrica*, 50, pp. 911-930.
- Binmore, K., A. Rubinstein, and A. Wolinsky (1986), The Nash Bargaining Solution in Economic Modelling, *The Rand Journal of Economics*, 17, pp. 176-188.
- Bogetoft, P. and M. Mikkers (2006), Organizing and regulating hospital and physician substitutes, Working Paper.
- Bogetoft, P. and H. Olesen (2004) *Design of Production Contracts*, CBS Press, Copenhagen.
- Bogetoft, P. and H. Olesen (2006), *Cooperatives and Payment Schemes*, to appear, CBS Press, Copenhagen
- CTG-Zaio (2005), Oriënterende monitor Ziekenhuiszorg, Analyse van de Marktentwikkelingen in het B-segment in 2005, <http://www.nza.nl/7113/10083/10089>
- CTG-Zaio (2006), Monitor Ziekenhuiszorg, Analyse van de Marktentwikkelingen in het B-segment in 2006, <http://www.nza.nl/7113/10083/10095>
- Haeck (2005), Health Care Reform in the Netherlands, OECD working party on competition and regulation round table oct. 17th 2005, Written submission
- Halbersma, R., M.C. Mikkers, E. Motchenkova and I. Seinen (2006) Market power and hospital-insurer bargaining in the Dutch hospital sector, Research paper to appear, NZa.
- Helderman, J. F.T. Schut T.E. D. van der Grinten and W.P. M. M. van de Ven (2005), *Market-Oriented Health Care Reforms and Policy Learning in the Netherlands*, Journal of Health Politics, Policy and Law, Vol. 30, Nos. 1-2



Jong, Ph. R. de and I. Mosca (2006) *Changes and Challenges of the New Health Care Reform in the Netherlands: What Should the Dutch Be Aware Of?* TILEC Working Paper DP-2006-026, Tilburg University.

Ketcham, J., V.L. Smith and A.W. Williams (1984). A Comparison of Posted-Offer and Double-Auction Pricing Institutions. *Review of Economic Studies*, 51(4): 595-614

Lakdawalla, Darius and Philipson, Tomas J. (1998) *Nonprofit Production and Competition*. NBER Working Paper No. W6377.

McGuire, Th. (2000), *Physician agency*, in Culver A., *Handbook of Health Economics* vol 1, Amsterdam, Elsevier, 2000, pp. 462-528.

Myerson, R. B. and M. A. Satterthwaith (1983). Efficient Mechanisms for Bilateral Trading. *Journal of Economic Theory* 1(29): 265-281.

Joseph P. Newhouse, *Free for All?: Lessons from the RAND Health Insurance Experiment*, Harvard University Press, reprint 1996.

Rubinstein, A. (1982), Perfect Equilibrium in a Bargaining Model, *Econometrica*, 50, pp. 97-109.

Shleifer, A. (1985). A Theory of Yardstick Competition. *Rand Journal of Economics* 16, pp. 319-327.

Schut, F.T. and W.P.M.M. van de Ven (2005) *Rationing and Competition in the Dutch Health-care System*, *Health Economics*, 14, pp. 59-74.

Sloan, F (2000), *Not-for-profit ownership and hospital behavior*. In Culyer AJ and Newhouse JP (eds), *Handbook of Health Economics*, Amsterdam, North Holland, Ch. 27., pp. 1141-1172.

Sloan, F. and A. Steinwald (1980) Effects of Regulation on Hospital Costs and Input Use, *Journal of Law and Economics*, 23(1), pp. 81-109.

Sutton, J. (1986) Non-Cooperative Bargaining Theory: An Introduction, *Review of Economic Studies*, LIII, pp. 709-724.

Varian, H. R. (1992) *Microeconomic Analysis*. Norton, NY.

Zweifel, P and W.G. Manning, *Physician agency*, in Culver A., *Handbook of Health Economics* vol 1, Amsterdam, Elsevier, 2000, pp. 462-528.



## 7. Appendix A: Numerical estimates of distortions

Consider a single-product case with a hospital cost function parameterized as  $C(y) = ay^\alpha$  where the parameter  $\alpha$  expresses the convexity of the production function and a scaling parameter  $a \geq 0$  it's level. Let the benefit function  $B(y) = y^\beta$  with  $0 < \beta \leq 1$  be defined as the envelope of pseudo-concave utility functions. Initially, we ignore the competitive effects of the regulation and simply assume an allowable payment level  $R(y) = yf(w)$  where  $f(w)$  is the potential outcome of the yardstick mechanism using relative prices  $w$ , with  $f(w) = 1$  for a fully efficient hospital.

In Figure 7-1 below and Figure 7-2 on the next page, we show the unconstrained and constrained solutions for the case of  $a = 0.5$ ,  $\alpha = 2$ ,  $\beta = 0.75$ ,  $w = 0.5$  and  $\gamma = 0.5$  and  $0.8$ , respectively.

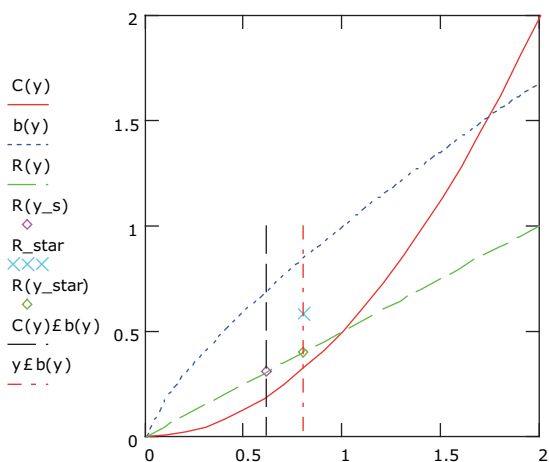


Figure 7-1 Unconstrained and constrained bargaining with  $g = 0.5$ .

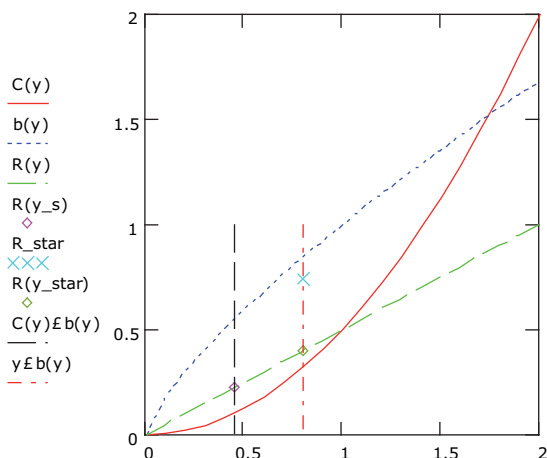
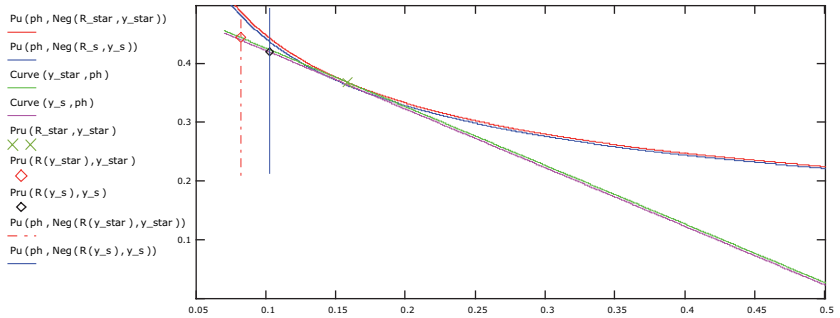
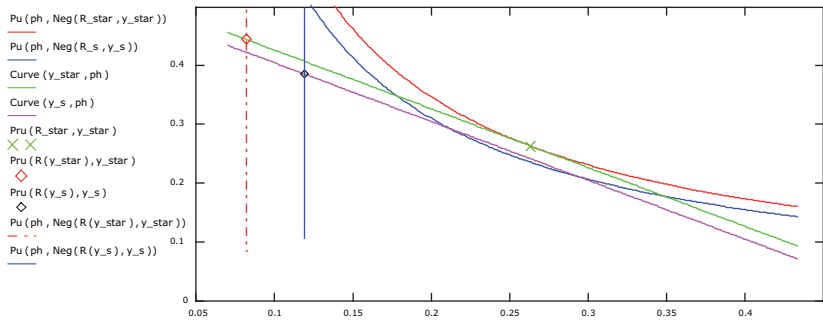
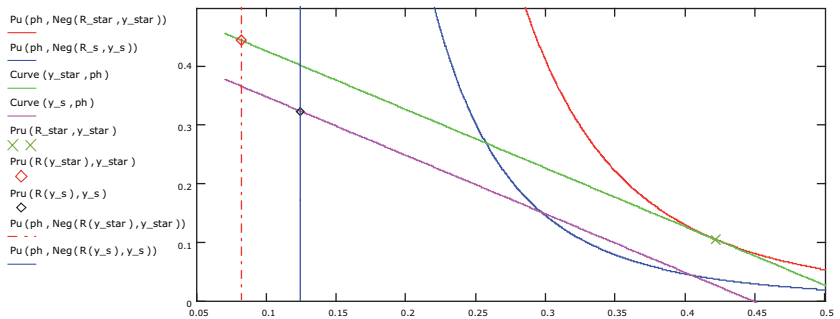


Figure 7-2 Unconstrained and constrained bargaining with  $g = 0.8$ .

Table 7-1 Numerical example: unconstrained and constrained production levels and surplus.

	Uncon- trained $\gamma = 0.5$	Adjus- ted $\gamma = 0.5$	Cons- trained $\gamma = 0.5$	Uncon- trained $\gamma = 0.8$	Adjus- ted $\gamma = 0.8$	Cons- trained $\gamma = 0.8$
$y^*$	0.794	0.794	0.610	0.794	0.794	0.449
$\pi_h(y^*)$	0.263	0.082	0.119	0.421	0.082	0.124
$\pi_u(y^*)$	0.263	0.444	0.385	0.105	0.444	0.324
$R^*$	0.579			0.579		
$R(y^*)$		0.397	0.305		0.397	0.224

The effect of the bargaining power is clearly illustrated in Figure 7-3 through Figure 7-5, where the constrained bargaining is illustrated. When the hospital's bargaining power is low (Figure 7-3), the social welfare effects are relative small of the regulation. The interesting case is therefore the scenario with hospital domination of the bargaining, in which the regulation acts as to countervail the rent exploitation by the hospitals towards the users (insurers).

Figure 7-3 Constrained bargaining ( $g = 0.3$ )Figure 7-4 Constrained bargaining ( $g = 0.5$ ), cf. Figure 4.3.Figure 7-5 Constrained bargaining ( $g = 0.8$ ).

The impact of the demand function, or more specifically the slope  $\beta$  of the benefit function is illustrated in Figure 7-6 for the case of  $a = 0.5$ ,  $\alpha = 2$ ,  $w = 0.5$ ,  $\gamma = 0.5$ . The unconstrained production level  $y^{**}$  is monotonously increasing with decreasing concavity, which confirms the intuition of demand curve exploitation. In Figure 7-7 and Figure 7-8 the welfare loss, measured as the difference in welfare  $W(y) = b(y) - C(y)$  is illustrated as a function of its slope  $b$ . As the bargaining space is in the form of an ellipsoid, the welfare loss is not monotone when the unconstrained optimum is close to the ends of the interval.

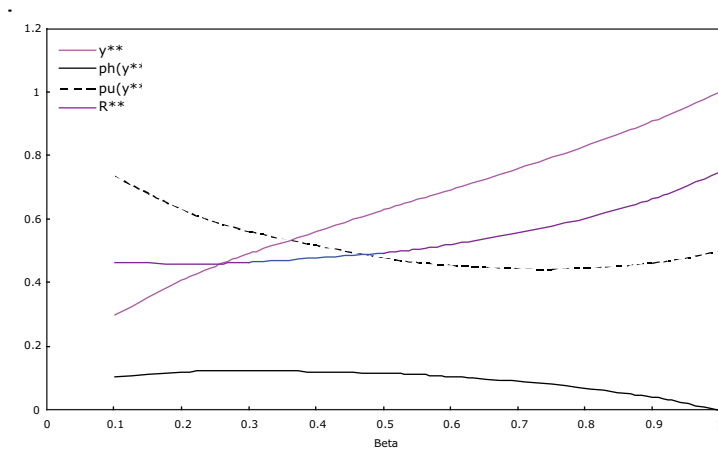


Figure 7-6 Impact of concavity of the benefit function.

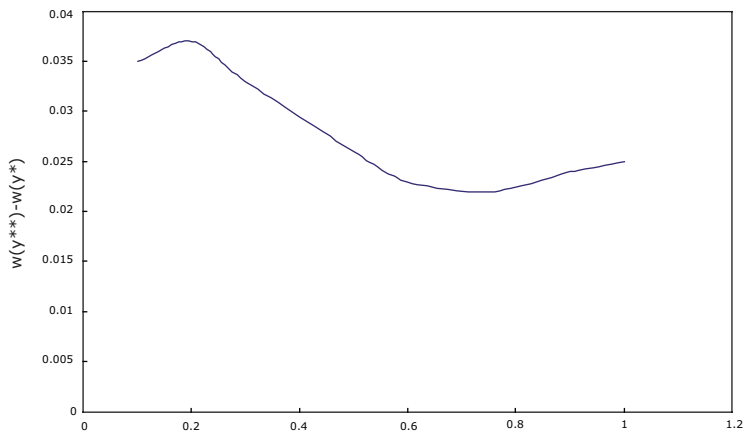


Figure 7-7 Welfare loss as a function of slope  $b$  of benefit function.

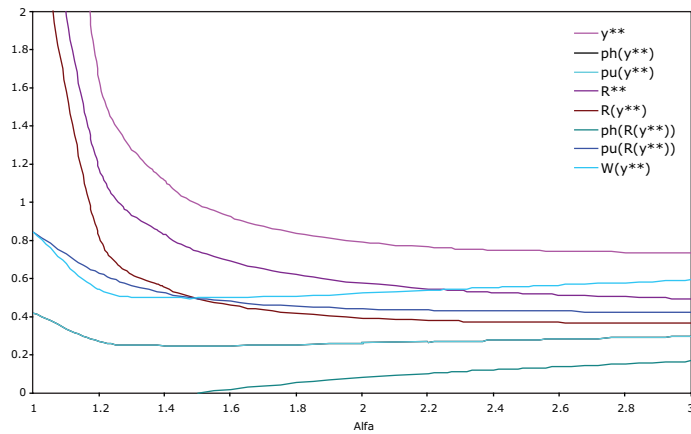


Figure 7-8 Impact on unconstrained and constrained bargaining by the slope  $\alpha$  of the cost function.

Table 7-2 Sensitivity analysis of concavity of benefit function, welfare losses and curvature.

a	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
$\alpha$	2	2	2	2	2	2	2	2
w	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
$\gamma$	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
$\beta$	0.1	0.2	0.3	0.5	0.6	0.75	0.9	1.0
$\gamma^{**}$	0.298	0.409	0.493	0.63	0.694	0.794	0.909	1.000
$\pi_h(\gamma^{**})$	0.105	0.121	0.125	0.117	0.106	0.082	0.041	0.000
$\pi_u(\gamma^{**})$	0.737	0.632	0.562	0.479	0.456	0.444	0.463	0.500
$R^{**}$	0.465	0.46	0.465	0.496	0.522	0.579	0.665	0.750
$R(\gamma^{**})$	0.149	0.204	0.246	0.315	0.347	0.397	0.454	0.500
$W(\gamma^{**})$	0.842	0.753	0.687	0.595	0.562	0.526	0.505	0.500
$\gamma^*$	0.131	0.223	0.304	0.449	0.516	0.61	0.703	0.768
$\pi_h(\gamma^*)$	0.057	0.087	0.106	0.124	0.125	0.119	0.104	0.089
$\pi_u(\gamma^*)$	0.751	0.629	0.598	0.449	0.414	0.385	0.377	0.384
$R^*$	0.149	0.204	0.246	0.315	0.347	0.397	0.454	0.500
$R(\gamma^*)$	0.065	0.111	0.152	0.225	0.258	0.305	0.352	0.384
$W(\gamma^*)$	0.807	0.716	0.654	0.569	0.539	0.504	0.481	0.475
$W(\gamma^{**}) - W(\gamma^*)$	0.035	0.037	0.033	0.026	0.023	0.022	0.024	0.025
$\gamma^{**} - \gamma^*$	0.167	0.186	0.189	0.181	0.178	0.184	0.206	0.232

The numerical illustrations show that the regulation may lead to some distortions in production level. The distortion in physical terms will decline as the elasticity of demand gets smaller. The dead-weight loss, i.e. the loss of social welfare, may increase as illustrated in the example. This is a consequence of the fact that as the elasticity of demand diminishes, any physical distortion also becomes more costly in terms of social welfare.

### Analytical example

Consider a given hospital. We now assume that the users' benefits are linear in treatments, costs are quadratic, and the bargaining powers are equal



$$b_h(y_h)=y_h \quad c_h(y_h)=(y_h)^2 \quad \gamma_h=\gamma_u=0.5$$

It is clear that unregulated bargaining will lead to the socially optimal outcome  $y_h=0.5$  and a settled payment somewhere between 0.25 and 0.5. Also, as long as the regulated  $p_h$  is 1 or above, the regulation will not matter. With equal bargaining power, the parties would settle for a payment (defined one way or the other in a contract) of  $3/8$  for  $y_h=0.5$ . This is equivalent of a unit price of  $3/4$ . Hence, if the regulation price is  $p_h \geq 0.75$  and that parties are equally powerful, the regulation will not matter.

If  $p_h < 0.75$ , however, the regulation will matter. In this case, the bargaining would solve

$$\max_{y_h} (b_h(y_h) - R_h)(R_h - c_h(y_h))$$

$$\max_{y_h} (y_h - p_h y_h)(p_h y_h - y_h^2)$$

To understand the implications of this, assume for example that the regulator uses a unit price equal to average costs at the optimal production level. In this case, we see that the bargaining leads to  $y_h=1/3$ , i.e. a fall in production below first best production level.

This illustrates that even though the production initially is optimal, the regulation even though it uses average costs (possibly of other and similar units) will cause it to be too low in this next period. In this case, there regulation does not support the optimal production level.

In terms of convergence, we see that the production level actually diverges from the optimal level. This is especially clear when we continue the regulation. In the next period, the production level of  $1/3$  would make us infer that average costs are  $1/3$ , and using this in the third period we get even lower production, namely  $y_h=2/9$ .

As another example, we may imagine that the hospital has all the bargaining power. It would then in an unregulated past set price equal to 1 and produce 0.5. The average costs at this production level (for other hospitals that we believe are similar) would be 0.5 and in the next period the hospital would choose output level  $1/4$ . The average costs is now  $1/4$  and in the third period, the hospital would produce  $1/8$  etc.

The numerical illustration and the first analytical example clearly show the main characteristics of the regime: the shift of the bargaining power

to the users and the limitation of the information rent. Whereas a mere reduction (cap) of the revenues for the hospital might violate the individual rationality constraint, the adjusted constrained optimization reestablishes part of the information rent for the hospital, yet at a lower level than in a pure bilateral bargaining. However, as any regulatory constraint, this provision of incentives is a second-best solution that comes at a social cost. The distortion is proportional to the elasticity of the demand and the convexity of the cost function. From this we can infer that

- The regulation may not support the optimal production level in all cases when it is active
- The regulation may in fact make the production level diverge more and more from the optimal production level
- The divergence is faster the more bargaining power the hospital has. In other words, the usage of average costs and the tendency to react to marginal costs leads to distorted production. This distortion is present as long as the hospital has some bargaining power – but it is less dramatic the more power the insurance companies have. Hence, real bargaining dampens the regulatory distortions – but it will not eliminate it.

To sum up, we have demonstrated that some distortions in production level are possible. The distortions in production level demonstrated in the numerical and analytical examples presume, however, a demand that is somewhat price sensitive. In the Dutch context, and given the universal insurance and the fact that the general practitioners serve as gate-keepers independently of the insurance companies, it is likely that the demand for the vast majority of DBCs is entirely price-inelastic. In that case, the distortions will tend to vanish. In that case, the only role of the regulation is to give the users more bargaining power so that they can keep a larger share of the social gains.



The Dutch Healthcare Authority (NZA) is the regulator of health care markets in the Netherlands. The NZa is established at October 1, 2006 and is located in Utrecht.

The NZa promotes, monitors and safeguards the working of health care markets. The protection of consumer interests is an important mission for the NZa. The NZa aims at short term and long term efficiency, market transparency, freedom of choice for consumers, access and the quality of care. Ultimately, NZa aims to secure the best value for money for consumers.

The Research Paper Series present scientific research on health care markets and addresses an international forum. The Research Paper Series offers NZa staff and invited authors an opportunity to disseminate their research findings intended to generate discussion and critical comments. The goal is to enhance the knowledge and expertise on the regulation of health care markets.

This paper reflects the personal views of its authors, which are not necessarily those of their employers. This paper is not in any way binding the board of the NZa.