CESifo Working Papers
www.cesifo.org/wp

Hospital Competition with Soft Budgets

Kurt R. Brekke
Luigi Siciliani
Odd Rune Straume

CESifo Working Paper No. 4073
Category 1: Public Finance
January 2013

CESifo *Working Paper No. 4073*

Hospital Competition with Soft Budgets

Abstract

We study the incentives for hospitals to provide quality and expend cost-reducing effort when their budgets are soft, i.e., the payer may cover deficits or confiscate surpluses. The basic set up is a Hotelling model with two hospitals that differ in location and face demand uncertainty, where the hospitals run deficits (surpluses) in the high (low) demand state. Softer budgets reduce cost efficiency, while the effect on quality is ambiguous. For given cost efficiency, softer budgets increase quality since parts of the expenditures may be covered by the payer. However, softer budgets reduce cost-reducing effort and the profit margin, which in turn weakens quality incentives. We also find that profit confiscation reduces quality and cost-reducing effort. First best is achieved by a strict no-bailout and no-profit-confiscation policy when the regulated price is optimally set. However, for suboptimal prices a more lenient bailout policy can be welfare improving. When we allow for heterogeneity in costs and qualities, we also show that a softer budget can raise quality for high-cost patients (and therefore reduce ‘skimping’ on such patients).

JEL-Code: I110, I180, L130, L320.

Keywords: hospital competition, soft budgets, quality, cost efficiency.

Kurt R. Brekke
Department of Economics
Norwegian School of Economics
Helleveien 30
Norway – 5045 Bergen
kurt.brekke@nhh.no

Luigi Siciliani
Department of Economics and Centre for Health Economics / University of York
Heslington
UK – York YO10 5DD
ls24@york.ac.uk

Odd Rune Straume
Department of Economics / NIPE
University of Minho
Campus de Gualtar
Portugal – 4710-057 Braga
o.r.straume@eeg.uminho.pt

September 13, 2012

1 Introduction

Hospital deficits and government-sponsored bailouts are frequently observed in many countries. Recently, the English newspaper The Telegraph wrote (15 Sept. 2011):

> "Taxpayers will face a £5 billion bill to bail out failing NHS hospitals over the next few years unless radical action is taken."

This headline was based on a report written by a former health adviser (Prof. Paul Corrigan) to Tony Blair.[1] According to this report more than 40 UK hospitals are facing financial problems, and the National Health Service must deliver £20 billion of efficiency savings to balance the budgets. This situation is not specific to the UK. In Norway, for instance, hospitals have been running deficits over a long period. In 2002 the government introduced a reform to harden the hospitals' budget constraint by transforming the public hospitals into state-owned enterprises. However, Hagen and Kaarbøe (2006) show that the hospitals continued to report deficits and receive supplementary funding from the government after the reform. Hospital deficits and bailouts are also present in health care systems with a larger private sector as in the US. According to Shen and Eggleston (2009), the share of general acute US hospitals reporting negative income grew from 21 percent in 1995 to 29 percent in 2004.

In a recent paper, Kornai (2009) discusses what he calls the 'soft budget syndrome' in the hospital sector.[2] Here he points out that, despite being a widespread phenomenon, the research on soft budgets in the hospital sector is very limited.[3] Our paper aims at bridging this gap by studying the impact of soft budgets on the provision of hospital care. We ask two key questions: (i) What are the incentives for hospitals to be cost efficient and invest in quality when facing soft budgets? (ii) Is there a scope for soft budgets to be welfare improving in the hospital sector?

[1] The report "The hospital is dead, long live the hospital" was published by the UK think tank *Reform* in September 2011.

[2] Kornai is considered to be the father of the concept of soft budget constraints. For a general review of this literature, see Kornai, Maskin and Roland (2003) and Kornai (2001).

[3] There are a couple of (mainly) empirical papers on soft budgets in the hospital market, see e.g., Duggan (2000), Shen and Eggleston (2009) and Eggleston and Shen (2011). We will describe the related literature in more detail below.

To analyse these questions, we consider a model with two competing hospitals that differ in location (à la Hotelling) and face demand uncertainty. The hospitals invest in quality and expend effort on cost reductions before the state of demand is revealed.[4] We consider the Nash equilibrium where the hospitals earn negative profits in the high-demand state and positive profits in the low-demand state. There are two reasons why hospitals run deficits (surpluses) when demand is high (low).[5] First, prices are regulated, implying that hospitals cannot profit on high demand by increasing their prices. This is contrary to most other industries, where high demand enables firms to charge higher prices and earn more profits. Second, hospital production involves diseconomies of scale.[6] This means that the profit margin is likely to become negative in the high-demand state when prices are fixed and hospitals operate close to (or at) their capacity constraints.

We introduce soft budgets in the standard way by assuming that the government (sponsor) with some probability will bailout the hospital ex post if it runs a deficit. The bailout probability measures the softness of the budget. We also allow for the possibility that the surplus in the low-demand state is confiscated. Duggan (2000) argues that public hospitals enjoying soft budgets face the problem that their surpluses in good times may be expropriated by the government ('ratchet effect'). Within this framework, we study how soft budgets (bailouts) and profit confiscation influence the hospitals' incentives for cost efficiency and quality investments. In an extension we allow for patients' heterogeneity on costs and investigate the incentives towards 'skimping' (under-provision of quality to high-cost patients) and 'creaming' (over-provision of quality to low-cost patients).

[4] We therefore consider dimensions of quality and cost efficiency that are more long-term and cannot be easily adjusted according to demand fluctuations. This can be investments in machinery, medical equipment, training programs for medical staff, treatment protocols, management procedures, etc.

[5] There are also some indicative evidence that high demand is the 'bad state' in the hospital sector. For instance, Hagen and Kaarbøe (2006) find that hospital deficits in Norway tend to be higher when the hospitals treat more patients than expected (planned). They argue that this is the key source of hospital deficits.

[6] Aletras (1999) suggests that economies of scale appear to be fully exploited in acute hospitals with 100–200 beds. Larger hospitals with 300–600 beds display diseconomies of scale (see also Ferguson, Sheldon and Posnett, 1999, and Folland, Goodman and Stano, 2004, for literature surveys). In England, the average size of hospitals is above 600 beds (Hughes and McGuire, 2003; Siciliani, Stanciole and Jacobs, 2009). Diseconomies of scale are therefore likely to be present.

Our main findings are the following. First, we show that *softer budgets* weaken hospitals' incentives for *cost efficiency*. The reason is that a higher bailout probability reduces the expected deficit in the high-demand state, and weakens the incentives for hospitals to expend effort on cost reductions. This finding is in line with the general concern that bailouts give rise to moral hazard by the hospital management in running the hospital efficiently. Indeed, bad management is frequently mentioned as a main reason for hospital deficits in the policy debate. In the UK, for instance, the government has a policy of firing the managers of NHS trusts reporting significant deficits. Our finding suggests that soft budgets may be the source of bad management, and that a no-bailout policy might be more effective in reducing hospital deficits than firing hospital managers.

Second, we find, somewhat surprisingly, that *softer budgets* have an ambiguous effect on *quality* incentives. In the high-demand (low-demand) state, hospitals earn a negative (positive) profit margin, and have therefore an incentive to reduce (increase) quality to avoid (attract) patients. A higher bailout probability increases the hospitals' expected (ex ante) profit, and therefore has a direct, positive impact on the incentives for quality investments. However, softer budgets also reduce the incentives for cost efficiency, as pointed out above, and this implies a negative, indirect effect on quality incentives due to a lower profit margin. The net effect of a more lenient bailout policy on hospital quality is therefore ambiguous and depends on the relative strengths of the direct and the indirect effects.

Third, we show that *profit confiscation* is always bad for hospital quality and cost efficiency. Hospitals expend effort on cost reductions to obtain a higher profit margin. However, when parts of this profit margin might be confiscated by the government (sponsor), then the incentive for cost efficiency is obviously reduced. Hospitals' incentives to invest in quality is related to the expected profit. A higher probability of profit confiscation reduces the expected profit directly by the reduction of the surplus in the low-demand state, but also indirectly through the incentives for cost efficiency. Thus, contrary to the case of bailouts, the direct and indirect effects of profit confiscation pull in the same direction, resulting in weaker incentives for hospitals to invest in quality.

Fourth, we analyse whether soft budgets can be welfare improving or not in the hospital sector. First, we show that, if the government can set socially optimal prices, then first-best is achieved by a strict no-bailout and zero-profit confiscation policy. However, in practice, the pricing schemes used for hospitals are not maximising social welfare, but are often cost-based.[7] We show that for an exogenous price, a strictly positive bailout policy may improve social welfare if (i) the hospitals' profit margin is sufficiently low and (ii) the disutility of cost-containment effort is sufficiently high. In this case, quality is suboptimal, and a higher bailout probability has a weak negative effect on cost efficiency and a stronger positive effect on quality. We also show that a more lenient bailout policy can be an optimal policy substitute to higher treatment prices.

Finally, when we allow for patient heterogeneity with respect to treatment costs, we show that a softer budget can increase quality for high-cost patients and therefore reduce 'skimping'. Whether it increases or decreases the quality for low-cost patients depends on the profit margin on these patients in the high-demand state: if the margin is positive (negative) then a softer budget will reduce (increase) quality for low-cost patients. If the cost difference between the two patient types is sufficiently high, we show that softer budgets may reduce not only 'skimping' but also 'creaming' (overprovision of quality to low-cost patients). If this is the case, welfare unambiguously increases as a result of softer budgets (if prices are fixed and equal for both patient types).

The literature on soft budgets in hospital markets is scarce, but there has been some recent, mainly empirical, papers on this issue. Shen and Eggleston (2009) study the effect of soft budgets on access and quality in hospital care both theoretically and empirically. The theoretical part builds on an incomplete contract model by Eggleston (2008) on soft budget constraints, and focuses on the interaction between a government purchaser and a hospital manager.[8] The hospital manager can develop and implement innovations in

[7] For instance, the pricing scheme based on Diagnosis Related Groups (DRGs), which is frequently used in practice, involves average cost pricing.

[8] Eggleston (2008) offers an extension to the seminal work by Hart, Shleifer and Vishny (1997) on the scope of government. She shows that government managers' soft incentives arise from their lack of control rights, because of softer budget constraints.

quality and cost efficiency. If the management fails to implement the innovations, the hospital faces a probability of bailout or closure. They show that softer budgets leads to less cost control, whereas the effect on quality is ambiguous. The latter result relies on the assumption that cost savings damages quality. In absence of this quality-damaging effect, softer budgets have a negative impact on quality.

While our paper addresses the same issues and arrives at similar results as Shen and Eggleston (2009), the two studies differ substantially in their modelling approaches and the mechanisms that drive the results. We start out with a standard model of hospital competition and introduce soft budgets into this framework, whereas Shen and Eggleston (2009) makes use of property right theory, and applies it to the hospital market. In their paper softer budgets have a negative impact on quality unless the quality-damaging effect of cost efficiency is sufficiently strong. In our paper, in contrast, softer budgets have a *positive* impact on quality unless the indirect effect of lower cost efficiency incentives is sufficiently strong. We therefore believe the two papers complement each other.

Shen and Eggleston (2009) measures soft budgets by the inverse of the probability of hospital closures. They first estimate this probability by using US data from hospital closures for 1990-2000, and subsequently use this to estimate the impact on hospital quality (AMI mortality rates) and cost efficiency (safety-net services). Their empirical results show that softer budgets are associated with lower cost efficiency and higher quality. This study is followed up by Eggleston and Shen (2011) that use the same framework, but focuses on the interaction between hospital ownership and soft budgets. In this paper, they find that for-profit hospitals are more cost efficient (more likely to drop safety-net services) and offers lower quality (higher AMI mortality rates) than non-profit and public hospitals, when controlling for difference in the softness in the budget constraints.

Duggan (2000) is another empirical paper on hospital ownership and soft budget constraints using US hospital data. This paper uses an exogenous change in the hospital financing that was intended to improve the medical care for the poor to test different theories of organisational behaviour, and find that the difference between the three types of hospitals is caused by the soft budget constraint of public hospitals. In particular, he finds that the increase in funding was matched by a one-by-one dollar reduction in the subsidy from the local authorities that owned the hospitals. This profit confiscation implied that public hospitals did not respond, whereas for-profit and non-profit did respond, to the financial incentives provided by the change in the hospital financing targeted towards the poor.

Our paper also relates to the literature on hospital competition. This includes work by, for instance, Gravelle (1999), Lyon (1999), Beitia (2003), Brekke, Nuscheler and Straume (2006, 2007), Karlsson (2007) and Brekke, Siciliani and Straume (2011).[9] To our knowledge, the present paper is the first attempt to analyse competition between hospitals in the presence of soft budgets. We contribute to this literature in two ways. First, we show that the incentives for quality competition crucially depend on the softness of the hospitals' budget constraints. As described above, a higher probability of profit confiscation has a negative impact on quality competition incentives, whereas the effect of a higher bailout probability is ambiguous. Second, we add cost efficiency to the basic model, and show that this also has an impact on the quality competition incentives.[10] A higher probability of bailout or profit confiscation both weakens the incentives for cost-containment effort, which in turn reduces the incentives for hospitals to invest in quality to compete for patients.

The rest of the paper is organised as follows. In Section 2 we present the model and derive the Nash equilibrium when the provider chooses quality and cost-reducing effort. In this section we analyse the impact of bailout and profit confiscation on hospitals' quality and cost-reducing effort. In Section 3 we analyse the welfare implications of soft budgets considering both first-best and second-best policies. Finally, in Section 4 we present some concluding remarks.

[9] See also Brekke, Siciliani and Straume (2008) who analyse the effect of hospital competition on waiting times, which are frequently interpreted as a quality dimension.

[10] Brekke, Siciliani and Straume (2012) consider quality competition where providers also choose cost-containment effort. However, this paper differs as there is no demand uncertainty and providers are profit constrained (e.g., non-profit) and altruistic.

2 Model

Consider a market for hospital treatments where consumers (patients) are uniformly located on the line segment $S = [0, 1]$. Each patient demands one unit of treatment and there are two hospitals in the market that can provide adequate treatments. Hospital 1 is located at the left endpoint of the line, while Hospital 2 is located at the right endpoint. The utility of a patient located at $x \in S$ and seeking treatment at Hospital i is given by

$$U(x) = \begin{cases} v + q_1 - tx & if \quad i = 1 \\ v + q_2 - t(1 - x) & if \quad i = 2 \end{cases}, \tag{1}$$

where q_i is the quality provided by Hospital i and t is a transportation cost parameter measuring the importance of travelling distance relative to quality differences. The patient who is indifferent between the two hospitals is located at

$$\widehat{x} = \frac{1}{2} + \frac{q_1 - q_2}{2t}. \tag{2}$$

We assume that the two hospitals face uncertainty about the total number of patients seeking treatment. More specifically, we assume that the distribution of patients is known and given, but the density can take one of two values. In *State L*, which occurs with probability μ, demand is low with density equal to 1, while in *State H*, demand is high with density equal to $n > 1$. Thus, the demands for treatment at the two hospitals are given as follows.

Hospital 1:

$$D_1 = \begin{cases} \widehat{x} & \text{in } State\ L \\ n\widehat{x} & \text{in } State\ H \end{cases}. \tag{3}$$

Hospital 2:

$$D_2 = \begin{cases} 1 - \widehat{x} & \text{in } State\ L \\ n(1 - \widehat{x}) & \text{in } State\ H \end{cases}. \tag{4}$$

The profit of Hospital i in *State* j is given by

$$\pi_i^j = pD_i^j - \frac{c_i}{2}\left(D_i^j\right)^2 - \frac{k}{2}q_i^2, \tag{5}$$

where p is a fixed (regulated) price per treatment and c_i and k are cost parameters related to output and quality investment, respectively. While quality investment costs are equal for both hospitals, we assume that each hospital can reduce its treatment costs by making an ex ante investment in cost containment. More specifically, we assume that $c_i := \sigma - e_i$, where e_i is the amount of cost containment effort invested by Hospital i. This investment imposes a non-monetary disutility $\frac{w}{2}e_i^2$ on the hospital management/staff.[11]

We consider a simultaneous-play game where each hospital chooses its level of quality (q_i) and cost containment effort (e_i). These two decisions are made before demand is realised and are therefore not state dependent. We will look for a symmetric Nash equilibrium where the hospitals have a positive profit in *State L* and a negative profit in *State H*. We also assume that any positive profits will be confiscated by the regulator with a probability θ, while a deficit will be covered with probability β. In other words, β is the probability that a hospital running a deficit will be bailed out and is thus a measure of the degree of budget softness.[12]

2.1 The Nash equilibrium

We assume that each hospital chooses quality and cost containment to maximise expected profits minus the disutility of cost-containment effort. The expected payoff of Hospital i is given by

$$\Pi_i = \mu\left(1-\theta\right)\pi_i^L + \left(1-\mu\right)\left(1-\beta\right)\pi_i^H - \frac{w}{2}e_i^2, \tag{6}$$

[11] The objective function does not explicitly take into account any potential (personal) costs to hospital managers if they run deficits, such as the possibility of being fired. However, if these costs take the form of a fixed penalty (e.g., lower salary), which is the most likely effect, then our analysis and results will not be affected.

[12] Alternatively, we can intepret β as the share of the deficit that will be covered by the regulator if *State H* occurs, while θ is the share of the profits that will be confiscated in *State L*.

where we assume that $\pi_i^L > 0$ and $\pi_i^H < 0$. Simultaneously maximising (6) with respect to q_i and e_i for $i = 1, 2$, the candidate symmetric Nash equilibrium, denoted by q^* and e^*, has quality and cost containment given by

$$q^* = \frac{\mu(1-\theta)\left(p - \frac{c^*}{2}\right) + (1-\mu)(1-\beta)\, n\left(p - \frac{nc^*}{2}\right)}{2kt\,(1-\beta+\mu(\beta-\theta))} \tag{7}$$

and[13]

$$e^* = \frac{\mu(1-\theta) + n^2(1-\mu)(1-\beta)}{8w}, \tag{8}$$

where $c^* := \sigma - e^*$.

Proposition 1 *(Equilibrium existence). If*

$$\mu > \frac{n(1-\beta)}{n(1-\beta)+1-\theta}$$

and

$$w > (n+1)\,\frac{\mu(1-\theta)+n^2(1-\mu)(1-\beta)}{8(\sigma(1+n)-4p)},$$

there always exists a $\sigma \in (\underline{\sigma}, \overline{\sigma})$ and a $t \in (\underline{t}, \overline{t})$ that ensure the existence of a symmetric Nash equilibrium with interior solutions for q^* and e^*.

Proof. Equilibrium existence requires $\pi^L > 0$, $\pi^H < 0$, $q^* > 0$ and $c^* > 0$. It is straightforward to show that $\pi^L > \pi^H$ if $\sigma > \underline{\sigma} := \frac{4p}{n+1} + e^*$ while $q^* > 0$ if $\sigma < \overline{\sigma} := 2p\frac{\mu(1-\theta)+n(1-\beta)(1-\mu)}{\mu(1-\theta)+n^2(1-\beta)(1-\mu)} + e^*$. Notice that neither of these conditions depend on t. Using the candidate equilibrium value of q^*, we have that $\lim_{t\to\infty} \pi^L = \frac{1}{8}(4p + e^* - \sigma) > 0$ if $\sigma < \widehat{\sigma} := 4p + e^*$. Since $\frac{\partial \pi^j}{\partial t} > 0$ and $\lim_{t\to 0} \pi^j = -\infty$ for $j = L, H$, there exist $\underline{t}$ and $\overline{t}$, where $\underline{t} < \overline{t}$, and such that $\pi^H < 0 < \pi^L$ (for $\sigma > \underline{\sigma}$) if $t \in (\underline{t}, \overline{t})$, where $\overline{t}$ may or may not be infinitely large. Furthermore, since $\widehat{\sigma} - \overline{\sigma} = 2p\frac{n(1-\mu)(1-\beta)(2n-1)+\mu(1-\theta)}{\mu(1-\theta)+n^2(1-\beta)(1-\mu)} > 0$, the restriction $\sigma < \overline{\sigma}$ always ensures that $\sigma < \widehat{\sigma}$. It remains to check if $\overline{\sigma} > \underline{\sigma}$. It is straightforward to

[13] An interior solution in cost-containment effort, i.e., $c^* := \sigma - e^* \geq 0$, requires

$$w \geq \frac{\mu(1-\theta)+n^2(1-\mu)(1-\beta)}{8\sigma}.$$

confirm that $\overline{\sigma} - \underline{\sigma} = 2p\,(n-1)\,\frac{\mu(1-\theta)-(1-\mu)n(1-\beta)}{(n+1)(\mu(1-\theta)+n^2(1-\beta)(1-\mu))} > 0$ if $\mu > \frac{n(1-\beta)}{n(1-\beta)+1-\theta}$. Finally, when imposing the constraint $\sigma > \underline{\sigma}$, an interior solution in cost containment effort requires $w > (n+1)\,\frac{\mu(1-\theta)+n^2(1-\mu)(1-\beta)}{8(\sigma(1+n)-4p)}$. ■

Notice that there are two reasons why hospitals may run a deficit in equilibrium. First, they cannot increase the price when demand is high and, second, they cannot turn down patients that demand treatment. The first reason also explains why *State H* is the financially bad state for the hospitals, in contrast to most other markets where producers are able to capitalise on high demand by increasing prices.

Using the equilibrium conditions derived in the proof of Proposition 1, we can also confirm that the two states differ, *in equilibrium*, with respect to the sign of the profit margin, given by $\frac{\partial \pi_i}{\partial D_i} = p - c_i D_i$.[14] Since $\underline{\sigma} < \sigma < \overline{\sigma}$ implies $\frac{c^*}{2} < p < \frac{nc^*}{2}$, it follows that both profits *and* the profit margin are negative in *State H*, while in *State L* the hospitals have positive profits and a positive profit margin in equilibrium. As we will show below, this particular property of the equilibrium, which relies on the assumption of convex treatment costs, is important for explaining the mechanisms behind some of the main results or our analysis.

2.2 Quality and cost efficiency

Although we are foremostly interested in analysing the effects of soft budgets on hospital quality and cost efficiency, we can also use the model to analyse other institutional and regulatory features of the hospital market and how these interact with budget softness. More specifically, there are four main parameters of interest in our model: the degree of budget softness (β), the probability of profit confiscation (θ), the inverse measure of competition intensity (t) and the regulated price (p). In the remainder of this section we will explore how marginal changes along each of these dimensions affect equilibrium quality and cost containment, and also how expected hospital deficits are affected.

[14] Notice that we define the profit margin in terms of output (rather than quality).

2.2.1 Budget softness

The effect of softer budgets on equilibrium quality is given by the sum of a direct and an indirect effect:

$$\frac{dq^*}{d\beta} = \frac{\partial q^*}{\partial \beta} + \frac{\partial q^*}{\partial e}\frac{\partial e^*}{\partial \beta}. \tag{9}$$

When making quality choices in the face of uncertainty, each hospital chooses optimally to invest in quality up to the point where the expected marginal revenue is equal to the marginal cost of quality. The marginal revenue of quality investments is the increase in demand (due to higher quality) times the profit gain of treating these extra patients. In equilibrium, this profit gain (i.e., the profit margin) is positive in *State L* and negative in *State H*, which means that *State H* contributes *negatively* to the expected marginal revenue of quality investments. For a given level of cost containment, a softer budget reduces the expected deficit in *State H*, which implies that the profit margin becomes less negative in this state. This means that the expected revenue of quality investments increases, which consequently strengthens each hospital's incentive for investing in quality. This is the direct effect of softer budgets on quality provision, and it is analytically given by

$$\frac{\partial q^*}{\partial \beta} = \frac{\mu(1-\theta)(1-\mu)(n-1)((n+1)c^* - 2p)}{4kt(1-\beta+\mu(\beta-\theta))^2} > 0. \tag{10}$$

However, there is also an indirect effect of budget softness on quality investments through cost-containment effort. The hospitals can reduce the deficit in *State H* by investing in cost containment ex ante. However, the incentives for doing so are weaker the more likely it is that the government will cover a deficit. Thus, a softer budget will lead to less cost efficiency in equilibrium:

$$\frac{\partial e^*}{\partial \beta} = -\frac{(1-\mu)n^2}{8w} < 0. \tag{11}$$

Less cost efficiency implies in turn that the profit margin is lower in both states. All else equal, this reduces the incentive to invest in quality to attract patients, since these are less profitable to treat for a less cost-efficient hospital:

$$\frac{\partial q^*}{\partial e} = \frac{\mu(1-\theta) + n^2(1-\beta)(1-\mu)}{4kt(1-\beta+\mu(\beta-\theta))} > 0 \tag{12}$$

The indirect effect of softer budgets on equilibrium quality is therefore given by

$$\frac{\partial q^*}{\partial e}\frac{\partial e^*}{\partial \beta} = -\frac{\big(\mu(1-\theta) + n^2(1-\beta)(1-\mu)\big)(1-\mu)n^2}{32wkt(1-\beta+\mu(\beta-\theta))} < 0. \tag{13}$$

Thus, the overall effect of budget softness on equilibrium quality is the sum of a positive (Eq. (10)) and a negative (Eq. (13)) effect. Comparing (10) and (13), we see that the total effect is positive if w is sufficiently large (which implies that the indirect effect is sufficiently small). In other words, softer budgets lead to higher quality provision in equilibrium if the scope for improved cost efficiency is sufficiently limited. Otherwise, it is relatively straightforward to construct numerical examples where the sum of (10) and (13) is either positive or negative, confirming the general ambiguity of the total effect.

Proposition 2 *Softer budgets lead to less cost efficiency in equilibrium, while the effect on equilibrium quality provision is ambiguous. There is a positive relationship between budget softness and quality provision if the disutility of cost-containment effort is sufficiently high.*

2.2.2 Profit confiscation

As for the case of budget softness, the effect of profit confiscation on equilibrium quality is given by the sum of a direct and an indirect effect:

$$\frac{dq^*}{d\theta} = \frac{\partial q^*}{\partial \theta} + \frac{\partial q^*}{\partial e}\frac{\partial e^*}{\partial \theta}. \tag{14}$$

For a given level of cost containment, a higher probability of profit confiscation reduces the profit margin in *State L* and therefore reduces the marginal revenue of quality investments, implying that the hospitals have weaker incentives for quality provision. This is the direct effect, which is analytically given by

$$\frac{\partial q^*}{\partial \theta} = -\frac{\mu (1-\beta)(1-\mu)(n-1)((n+1)c^* - 2p)}{4kt(1-\beta+\mu(\beta-\theta))^2} < 0. \tag{15}$$

In addition, a higher probability of profit confiscation will also reduce each hospital's incentive to invest in cost containment, since the profit gain from more cost-efficient production is more likely to be confiscated by the regulator:

$$\frac{\partial e^*}{\partial \theta} = -\frac{\mu}{8w} < 0. \tag{16}$$

A lower level of cost efficiency will lead to a further reduction in the hospitals' incentives for quality provision (cf. (12)). Thus, the direct and the indirect effect both go in the same direction.

Proposition 3 *A higher probability of profit confiscation leads to less cost efficiency and lower quality provision in equilibrium.*

2.2.3 Competition intensity and treatment prices

Increased competition (interpreted as a reduction in t) or a higher treatment price (p) do not affect the hospitals' incentives for reducing their costs, as these incentives are only determined by the expected profit gain from lower treatment costs for a given level of (expected) demand. Thus, the equilibrium level of cost efficiency does not depend on the competition intensity or the treatment price, and changes in either of these dimensions will only have a direct effect on equilibrium quality provision, given as follows:

$$\frac{\partial q^*}{\partial t} = \frac{-\mu (1-\theta)\left(p - \frac{c^*}{2}\right) - n(1-\mu)(1-\beta)\left(p - \frac{nc^*}{2}\right)}{2kt^2(1-\beta+\mu(\beta-\theta))} < 0, \tag{17}$$

$$\frac{\partial q^*}{\partial p} = \frac{n(1-\mu)(1-\beta) + \mu(1-\theta)}{2kt(1-\beta+\mu(\beta-\theta))} > 0. \tag{18}$$

The effect of increased competition on quality is composed of two opposite sub-effects, as represented by the two terms (with opposite signs) in the numerator of (17). A reduction in t increases demand responsiveness to quality, which stimulates quality incentives if the profit margin is positive but discourages quality incentives if the profit margin is negative. Although the profit margin is negative in equilibrium in *State H*, the expected profit margin is nevertheless positive, implying that the first term in the numerator of (17) is larger (in absolute value) than the second term.[15] Thus, in line with the existing theoretical literature on competition between profit-maximising hospitals facing fixed prices, we find an unambiguously positive relationship between competition intensity and equilibrium quality.

The positive effect of a price increase on quality provision is also in line with the existing literature, and the intuition is very straightforward. A higher price increases the profit gain of attracting more patients (i.e., the profit margin) in *both* states and the hospitals will consequently respond by competing harder (on quality) for patients.

Proposition 4 *Increased competition between the hospitals or a higher treatment price has no effect on cost efficiency, but leads to higher quality provision in equilibrium.*

2.3 Expected hospital deficit

What determines the expected hospital deficit (given by $(1-\mu)\,\pi^H$) in equilibrium? Since total demand is constant in each of the two states, the expected deficit depends (negatively) on quality investments and (positively) on cost efficiency. Based on the analysis in the previous section, we can say something about the effect of different policies on the expected size of hospital deficits.

Softer budgets affect the expected deficit through two different channels. A higher bailout probability increases the deficit through less cost efficiency, while the effect via quality costs is *a priori* ambiguous. If $\partial q^*/\partial \beta > 0$, a more lenient bailout policy will increase expected hospital deficits due to higher treatment *and* quality costs. From Proposition 2 we know that this result applies for a sufficiently large value of w.

A higher probability of *profit confiscation* will have two counteracting effects on the

[15] The unambiguous sign of (17) is confirmed by applying the equilibrium condition $\underline{\sigma} < \sigma < \overline{\sigma}$.

expected deficit. First, it will increase treatment costs through lower cost-containment effort. On the other hand, it will also reduce quality costs through weakened incentives for quality provision. The overall effect depends on the relative strength of these two counteracting effects. For example, if k is very high, so that equilibrium quality is not very responsive to changing market conditions, a higher probability of profit confiscation will increase the expected hospital deficit through less cost containment. If this is the case, the policy implication would be that, in order to reduce the problem of hospital deficits, the regulator should let the hospitals keep their profits more often in periods where they run a surplus. On the other hand, this result would be overturned if the parameter w is sufficiently high, which implies that it is hard to influence hospitals' cost efficiency. In this case, profit confiscation could be an effective policy instrument to reduce hospital deficits by curtailing their incentives for quality investments.

Stronger *competition* or a higher *treatment price* affect hospital deficits through changes in quality costs, while cost efficiency remains unchanged. The effect of stronger competition (lower t) is unambiguous, since competition strengthens quality incentives and therefore leads to higher expected deficits. An increase in the treatment price, on the other hand, has two counteracting effects on hospital deficits. As for the case of lower travelling costs, a price increase stimulates quality competition and leads to higher quality costs, and therefore higher expected hospital deficit, in equilibrium. On the other hand, a price increase has a direct positive effect on revenues (in both states), which counteracts the effects of higher quality costs. The total effect is given by

$$\frac{\partial \pi^H}{\partial p} = \frac{n}{2} - \frac{\left(n\left(1-\mu\right)\left(1-\beta\right)+\mu\left(1-\theta\right)\right)\left(\mu\left(1-\theta\right)\left(p-\frac{c^*}{2}\right)+n\left(1-\mu\right)\left(1-\beta\right)\left(p-\frac{nc^*}{2}\right)\right)}{4kt^2\left(1-\beta+\mu\left(\beta-\theta\right)\right)^2}, \tag{19}$$

where the first term is the direct (positive) revenue effect while the second term is the (negative) effect due to stronger quality competition. We see that the first term dominates the second if either t or k is sufficiently high, implying that equilibrium quality responds relatively little to price changes.

3 Social welfare

Considering social welfare as a sum of aggregate consumer utility net of treatment costs, quality costs and disutility of cost containment effort, the state-dependent welfare expressions are given by[16]

$$\begin{aligned} W^L &= \int_0^{\widehat{x}} (v + q_1 - ts)\, ds + \int_{\widehat{x}}^1 (v + q_2 - t(1-s))\, ds \\ &\quad -\frac{c_1}{2}\left(D_1^L\right)^2 - \frac{c_2}{2}\left(D_2^L\right)^2 - \frac{k}{2}q_1^2 - \frac{k}{2}q_2^2 - \frac{w}{2}e_1^2 - \frac{w}{2}e_2^2 \end{aligned} \tag{20}$$

and

$$\begin{aligned} W^H &= n\left(\int_0^{\widehat{x}} (v + q_1 - ts)\, ds + \int_{\widehat{x}}^1 (v + q_2 - t(1-s))\, ds\right) \\ &\quad -\frac{c_1}{2}\left(D_1^H\right)^2 - \frac{c_2}{2}\left(D_2^H\right)^2 - \frac{k}{2}q_1^2 - \frac{k}{2}q_2^2 - \frac{w}{2}e_1^2 - \frac{w}{2}e_2^2. \end{aligned} \tag{21}$$

Expected social welfare is therefore

$$W = \mu W^L + (1-\mu)\, W^H. \tag{22}$$

3.1 The first-best solution

It is straightforward to calculate the first-best optimal levels of quality and cost containment:

$$q^{fb} = \frac{n(1-\mu) + \mu}{2k} \tag{23}$$

[16] Social welfare can be interpreted as the difference between the sum of (gross) aggregate consumer utility and all costs and disutilities or, alternatively, as the sum of aggregate consumer utility *net* of payments to the provider and the utility of the provider (i.e., profit minus disutilities). Both specifications lead to the same expression for social welfare. Under the latter specification, the state-dependent welfare expressions for state L is given by

$$\begin{aligned} W^L &= \int_0^{\widehat{x}} (v + q_1 - ts)\, ds + \int_{\widehat{x}}^1 (v + q_2 - t(1-s))\, ds - (p\widehat{x} + p(1-\widehat{x})) \\ &\quad +(1-\theta)\left(\pi_1^L + \pi_2^L\right) - \frac{w}{2}e_1^2 - \frac{w}{2}e_2^2 + \theta\left(\pi_1^L + \pi_2^L\right), \end{aligned}$$

and similarly for W^H.

and

$$e^{fb} = \frac{n^2(1-\mu)+\mu}{8w}. \tag{24}$$

The regulator needs two instruments to implement the first-best solution. Comparing (24) with (8) we see that the optimal level of cost containment is implemented by setting $\beta = \theta = 0$. In other words, positive profits should never be confiscated and hospitals running deficits should never be bailed out. If the regulator is not able to commit to a strict no-bailout policy, the equilibrium level of cost-containment will always be less than socially optimal for all $\theta \in (0,1)$. In fact, with $\beta > 0$, the socially optimal level of cost containment can only be implemented by setting $\theta = -\left(\frac{1-\mu}{\mu}\right) n^2\beta < 0$, which implies giving a financial reward to hospitals that run a financial surplus. The socially optimal level of hospital quality, on the other hand, can be implemented by an appropriate choice of p, which does not affect cost efficiency (since (8) does not depend on p).

Proposition 5 *The first-best solution is implemented with no profit confiscation and a strict no-bailout policy, along with an appropriately chosen level of the treatment price.*

3.2 Second-best policies

Although the first-best solution is in principle implementable, there are several reasons why second-best policies may be more relevant in practice. First, in countries using DRG pricing, the treatment price is set according to a cost-based rule which is unlikely to coincide with the first-best optimal level when hospitals' incentives for quality provision and cost efficiency are taken into account. Second, it may be difficult, for political and other reasons, for the government to commit to a strict no-bailout policy. In this subsection we will consider each of these cases, where either p or β is exogenously given at a suboptimal level. First, we consider the case of an exogenous price p (which is not at the first-best level) and analyse if and when a deviation from a strict no-bailout policy may improve welfare. Second, we consider the case of an exogenous (and strictly positive) value of β and analyse how the optimal price depends on the bailout probability; in other words, whether higher treatment prices and a more lenient bailout policy are substitutes or complements from an optimal (second-best) policy perspective. Towards the end of this section we also examine how social costs of public funds affect the scope for welfare-improving bailout policies.

3.2.1 Welfare-improving bailout policies

Suppose that p is exogenously given while β can be thought of as a policy parameter, where a more lenient bailout policy corresponds to a higher value of β. On general form, the welfare effect of a higher bailout probability is given by

$$\frac{dW}{d\beta} = \frac{\partial W}{\partial q}\frac{\partial q^*}{\partial \beta} + \frac{\partial W}{\partial e}\frac{\partial e^*}{\partial \beta}. \tag{25}$$

Since the first-best level of cost containment effort is implemented for $\theta = \beta = 0$, the second term in (25) is negative for any $\theta \geq 0$.[17] Thus, for softer budgets to be welfare improving, the distortion of the hospitals' quality incentives must be reduced (implying that the first term in (25) is positive) to an extent that is sufficient to outweigh the welfare loss from lower cost efficiency. A complete characterisation of the optimal bailout policy is not feasible, due to the complexity of the mathematical expressions involved. Instead, we will characterise the scope for welfare-improving deviations from a strict no-bailout policy. This amounts to evaluate the sign of (25) at $\beta = 0$. If this is positive, we know that the optimal bailout policy involves a strictly positive bailout probability. When performing this analysis we also set $\theta = 0$, which implies that there are no welfare distortions with respect to cost efficiency from a marginal increase in β from zero (i.e., the second term in (25) vanishes), and therefore maximises the scope for a strictly positive bailout probability to be welfare-optimal. For $\theta = 0$, a marginal increase in β (evaluated at $\beta = 0$) leads to higher expected welfare either if there is a positive relationship between β and q^* and quality is underprovided ($q^* < q^{fb}$) or, vice versa, if quality is overprovided

[17] Formally, the second term in (25) is given by

$$\frac{\partial W}{\partial e^*}\frac{\partial e^*}{\partial \beta} = -\frac{n^2(1-\mu)\left(\mu\theta + (1-\mu)\beta n^2\right)}{32w}.$$

and a higher bailout probability reduces equilibrium quality. We will here consider only the former case and characterise the parameter space where a more lenient bailout policy can improve welfare by stimulating quality provision.

In the case under consideration, a more lenient bailout policy is welfare improving if $\left.\frac{\partial q^*}{\partial \beta}\right|_{\theta=\beta=0} > 0$ and $\left.\frac{\partial W}{\partial q}\right|_{q=q^*;\theta=\beta=0} > 0$. Setting $\theta = \beta = 0$ in (10) and (13), and summing these two terms, we derive

$$\left.\frac{\partial q^*}{\partial \beta}\right|_{\theta=\beta=0} = \frac{(1-\mu)}{4kt}\left(\mu(n-1)\left((n+1)c^* - 2p\right) - n^2 e^*\right). \tag{26}$$

As we already know from Proposition 2, this expression is positive if w is sufficiently high (which implies that e^* is correspondingly low). In addition, we see that the scope for softer budgets to increase quality is higher if p is lower relative to σ, or if μ is higher.[18]

The welfare effect of a higher quality provision, evaluated at the equilibrium quality level, and for $\theta = \beta = 0$, is given by

$$\left.\frac{\partial W}{\partial q}\right|_{q=q^*;\theta=\beta=0} = n(1-\mu) + \mu + \frac{1}{2t}\left(\Theta - \frac{\left(\mu + n^2(1-\mu)\right)^2}{8w}\right), \tag{27}$$

where $\Theta := n(\sigma n - 2p) - \mu(n-1)(\sigma(1+n) - 2p)$. Notice that Θ is monotonically decreasing in p and increasing in σ. Thus, the scope for higher quality to increase welfare is larger when p is small relative to σ. This is intuitive, as the hospitals have weaker incentives to invest in quality when the profit margin is lower. The same is true if the degree of competition is weaker (i.e., if t is higher). The scope for quality to be underprovided in equilibrium is also larger if w is higher. The reason is that a higher w reduces the chosen amount of cost-containment effort, which in turn reduces the profit margin and therefore incentives for quality provision.

Comparing the conditions for $\left.\frac{\partial q^*}{\partial \beta}\right|_{\theta=\beta=0} > 0$ and $\left.\frac{\partial W}{\partial q}\right|_{q=q^*;\theta=\beta=0} > 0$, we see that a

[18] Notice that

$$\frac{\partial\left(\mu(n-1)\left((n+1)c^* - 2p\right) - n^2 e^*\right)}{\partial \mu} = (n-1)\left(\frac{\mu(n-1)(n+1)^2}{4w} + \sigma(1+n) - 2p\right) > 0.$$

lower profit margin (i.e., lower p or higher σ) and a higher disutility of cost containment (i.e., higher w) increase the scope for both expressions to be positive, while less intense competition (i.e., higher t) increases the scope for quality to be underprovided ($\frac{\partial W}{\partial q} > 0$) while not affecting the scope for a positive relationship between budget softness and quality. However, it remains to be explicitly shown that the parameter space where the optimal bailout policy has a strictly positive bailout probability is non-empty. This can be easily confirmed by setting the price at the lowest permissible level ($p = \frac{\mu+n^2(1-\mu)}{2(\mu+n(1-\mu))}c^*$) while setting w at the highest permissible level ($w \to \infty$, which implies $e^* = 0$). The welfare effect of deviating from a strict no-bailout policy is then given by

$$\lim_{w\to\infty} \frac{dW}{d\beta}\bigg|_{\theta=\beta=0;p=\frac{\mu+n^2(1-\mu)}{2(\mu+n(1-\mu))}c^*} = \frac{(n-1)\,n\mu\,(1-\mu)\,c}{4kt} > 0, \tag{28}$$

which confirms that the welfare-optimal bailout policy in this case always has a strictly positive bailout probability.

Proposition 6 *When p is exogenously given, the bailout policy that maximises expected welfare has a strictly positive bailout probability if the hospitals' profit margin is sufficiently low and the disutility of cost-containment effort is sufficiently high.*

3.2.2 Optimal treatment prices with bailouts

Now suppose instead that the regulator sets the treatment price that maximises expected welfare, but is not able to commit to a strict no-bailout policy. The optimal price p^* which implements first-best quality is such that

$$q^*(\beta, \theta, p^*) = q^{fb}, \tag{29}$$

or, more explicitly,

$$p^* = \frac{c^*\left(\mu\,(1-\theta) + (1-\mu)\,(1-\beta)\,n^2\right) + 2t\,(1-\beta+\mu\,(\beta-\theta))\,(n\,(1-\mu)+\mu)}{2\,(\mu\,(1-\theta) + n\,(1-\mu)\,(1-\beta))}. \tag{30}$$

As expected, the optimal price increases with treatment costs and decreases with the degree of competition (inversely measured by t). It also decreases with cost-containment effort since higher effort implies a higher profit margin, which increases incentives for quality provision.

What is the effect on the optimal price of the supposed inability of the regulator to commit to a strict no-bailout policy? By totally differentiating (29) we obtain

$$\frac{dp^*}{d\beta} = -\frac{\partial q^*/\partial \beta}{\partial q^*/\partial p}. \tag{31}$$

Since equilibrium quality is strictly increasing in the price level, the sign of $dp^*/\partial\beta$ depends solely on the sign of $\partial q^*/\partial\beta$. Compared to a strict no-bailout policy, soft budget constraints imply a lower optimal price whenever they also encourage higher quality investments. This will be the case if the disutility of cost-containment effort is sufficiently high (cf. Proposition 2), implying that a more lenient bailout policy is a substitute to more high-powered incentives through the choice of a higher treatment price. Notice however that if the regulator can set prices, he also has an incentive to implement a hard budget constraint, since soft budgets reduce cost-containment effort while quality can be maintained at the first-best level by adequately changing the price.

Proposition 7 *If the disutility of cost-containment effort is sufficiently high, a more lenient bailout policy is an optimal policy substitute to higher treatment prices.*

3.2.3 Social costs of public funds

We finally discuss how our main result from the welfare analysis – that soft budgets may increase welfare when prices are fixed – is affected when the government has a strictly positive cost from raising money through distortionary taxation (for example due to distortions in the labour market). We assume that the distortionary cost of raising taxes is equal to λ per unit of taxes raised.

The state-dependent welfare expressions are now given by

$$\begin{aligned} W^L &= \int_0^{\widehat{x}} (v + q_1 - ts)\, ds + \int_{\widehat{x}}^1 (v + q_2 - t(1-s))\, ds - (1+\lambda)(p\widehat{x} + p(1-\widehat{x})) \\ &\quad + (1-\theta)\left(\pi_1^L + \pi_2^L\right) - \frac{w}{2}e_1^2 - \frac{w}{2}e_2^2 + \theta(1+\lambda)\left(\pi_1^L + \pi_2^L\right) \end{aligned} \tag{32}$$

and

$$\begin{aligned} W^H &= n\left(\int_0^{\widehat{x}} (v + q_1 - ts)\, ds + \int_{\widehat{x}}^1 (v + q_2 - t(1-s))\, ds\right) - (1+\lambda)\, n\, (p\widehat{x} + p(1-\widehat{x})) \\ &\quad + (1-\beta)\left(\pi_1^H + \pi_2^H\right) - \frac{w}{2}e_1^2 - \frac{w}{2}e_2^2 + \beta(1+\lambda)\left(\pi_1^H + \pi_2^H\right), \end{aligned} \tag{33}$$

where the transfer to the providers, as well as the profits confiscated in the low state and the deficits covered in the high state, are now multiplied by $(1+\lambda)$. These welfare expressions can be rewritten as

$$\begin{aligned} W^L &= \int_0^{\widehat{x}} (v + q_1 - ts)\, ds + \int_{\widehat{x}}^1 (v + q_2 - t(1-s))\, ds \\ &\quad - (1+\lambda\theta)\left[\frac{c_1}{2}\left(D_1^L\right)^2 + \frac{c_2}{2}\left(D_2^L\right)^2 + \frac{k}{2}q_1^2 + \frac{k}{2}q_2^2\right] \\ &\quad - \frac{w}{2}e_1^2 - \frac{w}{2}e_2^2 - \lambda(1-\theta)\, p \end{aligned} \tag{34}$$

and

$$\begin{aligned} W^H &= n\left(\int_0^{\widehat{x}} (v + q_1 - ts)\, ds + \int_{\widehat{x}}^1 (v + q_2 - t(1-s))\, ds\right) \\ &\quad - (1+\lambda\beta)\left[\frac{c_1}{2}\left(D_1^H\right)^2 + \frac{c_2}{2}\left(D_2^H\right)^2 + \frac{k}{2}q_1^2 + \frac{k}{2}q_2^2\right] \\ &\quad - \frac{w}{2}e_1^2 - \frac{w}{2}e_2^2 - n\lambda(1-\beta)\, p. \end{aligned} \tag{35}$$

Expected social welfare is again $W = \mu W^L + (1-\mu) W^H$, with

$$\frac{dW}{d\beta} = \frac{\partial W}{\partial q}\frac{\partial q^*}{\partial \beta} + \frac{\partial W}{\partial e}\frac{\partial e^*}{\partial \beta} + 2(1-\mu)\lambda\pi^H. \tag{36}$$

Compared to the analysis with $\lambda = 0$, a softer budget constraint now has an additional cost.[19] With $\lambda = 0$ a higher coverage of a deficit implies a transfer from the consumer to the provider leaving welfare unaffected. With a positive opportunity cost of raising taxes it instead implies a reduction in welfare since the loss for the consumer is now larger than the gain for the provider. This welfare loss (together with the loss from lower effort) has to be traded off against the potential gains from higher quality. The case for a softer budget is therefore reduced.

Finally, it is also worth noticing that the presence of a positive opportunity cost of public funds introduces a potential welfare gain from more profit confiscation, since

$$\frac{dW}{d\theta} = \frac{\partial W}{\partial q}\frac{\partial q^*}{\partial \theta} + \frac{\partial W}{\partial e}\frac{\partial e^*}{\partial \theta} + 2\mu\lambda\pi^L. \tag{37}$$

More profit confiscation in states with positive profits implies that the government will have to raise less revenues through distortionary taxation.[20] With $\lambda = 0$ it is never optimal to confiscate profits since it reduces quality and effort and is therefore welfare decreasing. In the presence of distortionary taxation the benefits from recovering revenues need to be traded off against the welfare losses from lower quality provision and cost efficiency.

4 Extension: Creaming and Skimping

In the hospital sector, it is often argued that prospective payment systems of the DRG type can give distorted incentives in the provision of quality to patients with different severity and costs. Since the DRG price is fixed, the provider may have an incentive to provide low quality to patients with high costs and severity in an attempt to discourage them from seeking treatment at that hospital and to provide high quality to patients with low costs and severity in an attempt to attract them. These practices are known as, respectively, 'skimping' and 'creaming' (Ellis, 1998).

In this section we extend the analysis with treatment cost heterogeneity to investigate the effect of soft budgets: i) on the incentives to cream and skimp respectively low- and

[19] Recall that $\pi^H < 0$ in equilibrium.
[20] Recall that $\pi^L > 0$ in equilibrium.

high-cost patients, and ii) on welfare. We assume that there are two types of patients that differ in the marginal cost of treatment $d = \{c, s\}$ with $c < s$ (where c stands for 'creaming' and s for 'skimping'). Like Ellis (1998), we assume that hospitals, but not the purchaser, can observe the different types of patients and therefore provide different levels of quality (observable to patients), while receiving the same price p for both types of patients.

The main structure of the model is analogous to Section 2. A patient of type d who is indifferent between the two hospitals is located at $\widehat{x}^d = \frac{1}{2} + \frac{q_1^d - q_2^d}{2t}$. We assume that the two segments are of equal size. The demands for treatment at the two hospitals are given as follows:

$$D_1^d = \left\{ \widehat{x}^d \text{ in } \textit{State L; } n\widehat{x}^d \text{in } \textit{State H} \right\} \tag{38}$$

and

$$D_2^d = \left\{ (1 - \widehat{x}^d) \text{in } \textit{State L; } n(1 - \widehat{x}^d) \text{in } \textit{State H} \right\}. \tag{39}$$

To keep the analysis simple we omit cost-containment effort from the analysis.

The profit of Hospital i in *State* j treating patients' segment d is given by

$$\pi_i^{jd} = pD_i^{jd} - \frac{d}{2}\left(D_i^{jd}\right)^2 - \frac{k}{2}\left(q_i^d\right)^2 \tag{40}$$

The expected payoff of Hospital i is given by

$$\Pi_i = \mu (1 - \theta) \left(\pi_i^{Lc} + \pi_i^{Ls}\right) + (1 - \mu) (1 - \beta) \left(\pi_i^{Hc} + \pi_i^{Hs}\right). \tag{41}$$

As in the main model, we look for an equilibrium where the hospitals run a deficit when demand is high; i.e., $\pi_i^{Lc} + \pi_i^{Ls} > 0$ and $\pi_i^{Hc} + \pi_i^{Hs} < 0$.

Maximising (40) with respect to q_i^c and q_i^s, for $i = 1, 2$, the candidate symmetric Nash equilibrium has qualities given by

$$q^{*d} = \frac{\mu (1 - \theta) \left(p - \frac{d}{2}\right) + (1 - \mu) (1 - \beta) n \left(p - \frac{nd}{2}\right)}{2kt (1 - \beta + \mu (\beta - \theta))}, \ d = \{c, s\}, \tag{42}$$

which is analogous to Equation (7). q^{*d} is a decreasing function of the cost parameter d

and the quality provided to the low-cost segment (creaming) is higher than the quality for the high-cost segment (skimping):

$$q^{*c} > q^{*s}. \tag{43}$$

Proposition 8 *(Equilibrium existence). If $n > \frac{4s}{c+s} - 1$, there always exists a $\mu \in (\widetilde{\mu}, 1]$, a $p \in \left(\underline{p}, \overline{p}\right)$ and a $t \in \left(\underline{t}, \overline{t}\right)$ that ensure the existence of a symmetric Nash equilibrium with interior solutions for q^{*d}.*

Proof. Equilibrium existence requires $\pi_i^{Lc} + \pi_i^{Ls} > 0$, $\pi_i^{Hc} + \pi_i^{Hs} < 0$ and $q^{*s} > 0$. It is straightforward to show that $q^{*s} > 0$ if $p > \underline{p} := \frac{s}{2} \frac{\mu(1-\theta)+(1-\mu)(1-\beta)n^2}{\mu(1-\theta)+(1-\mu)(1-\beta)n}$, which is also a sufficient condition for $q^{*c} > 0$, since $q^{*c} > q^{*s}$. Given that $\pi^{Hc} + \pi^{Hs} = pn - \frac{c+s}{8}n^2 - \frac{k}{2}\left((q^s)^2 + (q^c)^2\right)$ and $\pi^{Lc} + \pi^{Ls} = p - \frac{c+s}{8} - \frac{k}{2}\left((q^c)^2 + (q^s)^2\right)$ it follows that $\pi^{Lc} + \pi^{Ls} > \pi^{Hc} + \pi^{Hs}$ when $p < \overline{p} := \frac{(c+s)(1+n)}{8}$. Using the candidate equilibrium value of q^{*d}, we have that $\lim_{t\to\infty}\left(\pi^{Lc} + \pi^{Ls}\right) = p - \frac{c+s}{8} > 0$ if $p > \frac{c+s}{8} < \underline{p}$. Since $\frac{\partial\left(\pi^{Lc}+\pi^{Ls}\right)}{\partial t} > 0$, $\frac{\partial\left(\pi^{Hc}+\pi^{Hs}\right)}{\partial t} > 0$ and $\lim_{t\to 0}\left(\pi^{jc} + \pi^{js}\right) = -\infty$, there exists a $\underline{t}$ and $\overline{t}$, with $\underline{t} < \overline{t}$, such that $\pi^{Lc} + \pi^{Ls} > 0 > \pi^{Hc} + \pi^{Hs}$ if $t \in \left(\underline{t}, \overline{t}\right)$ and $p \in \left(\underline{p}, \overline{p}\right)$. It remains to check that $\overline{p} > \underline{p}$. This holds if $(c+s)(n+1) > 4sA(\mu)$ where $A(\mu) := \frac{\mu(1-\theta)+(1-\mu)(1-\beta)n^2}{\mu(1-\theta)+(1-\mu)(1-\beta)n}$. Notice that: (i) for $\mu = 0$ we have $A(0) = n$ and the inequality is never satisfied (since $s > c$ and $n > 1$); (ii) $\frac{\partial A(\mu)}{\partial \mu} = -\frac{n(1-\theta)(1-\beta)(n-1)}{(n(1-\beta)(1-\mu)+\mu(1-\theta))^2} < 0$; and, (iii) for $\mu = 1$ we have $A(1) = 1$ and the inequality is satisfied if n is sufficiently large: $n > \frac{4s}{c+s} - 1$. Therefore, when the last condition is satisfied, there exists a $\widetilde{\mu} < 1$ such that $\underline{p} < \overline{p}$ for $\mu \in (\widetilde{\mu}, 1]$. ■

The effect of a softer budget on qualities provided to the two patient types is given as follows:

Proposition 9 *(i) Softer budgets always increase the quality provided to high-cost patients. (ii) Softer budgets increase the quality also for low-cost patients if the cost difference $(s - c)$ is sufficiently small. (iii) Softer budgets may reduce the quality provided to low-cost patients if there is a sufficiently large cost difference between the two patient types.*

Proof. From (42), the effect of softer budgets on equilibrium quality is given by

$$\frac{\partial q^{*d}}{\partial \beta} = -\frac{\mu(1-\theta)(1-\mu)n\left(p-\frac{nd}{2}\right)}{2kt\left(1-\beta+\mu(\beta-\theta)\right)^2},\ d=\{c,s\}. \tag{44}$$

The sign of $\frac{\partial q^{*s}}{\partial \beta}$ depends on the sign of $\left(p-\frac{ns}{2}\right)$, which is negative for any $\underline{p} < p < \overline{p}$, since $\overline{p}-\frac{ns}{2} < 0$. Thus, $\frac{\partial q^{*s}}{\partial \beta} > 0$. Similarly, the sign of $\frac{\partial q^{*c}}{\partial \beta}$ depends on the sign of $\left(p-\frac{nc}{2}\right)$. Using the definition of $\overline{p}$, it is straightforward to show that

$$\overline{p}-\frac{nc}{2} < (>)\, 0 \quad if \quad \frac{s}{c} < (>)\, \frac{3n-1}{n+1}.$$

Since $\frac{3n-1}{n+1} > 1$ for $n > 1$, $\overline{p}-\frac{nc}{2} < 0$, and thus $p-\frac{nc}{2} < 0$ for any $p \in \left(\underline{p}, \overline{p}\right)$, if the cost ratio $\frac{s}{c}$ is sufficiently small (i.e., sufficiently close to 1). In this case, $\frac{\partial q^{*c}}{\partial \beta} > 0$. Otherwise, if $\frac{s}{c}$ is sufficiently large, $p-\frac{nc}{2} > 0$ for some $p \in \left(\underline{p}, \overline{p}\right)$, which implies $\frac{\partial q^{*c}}{\partial \beta} < 0$. ■

The intuition for these results is the following. Suppose that the low-cost types are profitable on the margin even in the high-demand state. In this case, each hospital can (ex ante) try to reduce the deficit in the high-demand state by making the following quality adjustments: i) reduce quality for high-cost patients in order to attract fewer of these since they are unprofitable on the margin; ii) increase quality for low-cost patients in order to attract more of these since they are profitable on the margin. Softer budgets weaken the hospitals' incentives to reduce the deficit in the high-demand state. In other words, softer budgets reduce each hospital's incentive to change the patient mix in the high-demand state towards more profitable patients. Consequently, the hospitals will increase the quality offered to high-cost patients and reduce the quality offered to low-cost patients.

The above described scenario applies only for the case of a sufficiently high cost difference between the two patient types. On the other hand, if the cost difference is sufficiently low, both patient types will be unprofitable on the margin in the high demand state and softer budgets lead to higher quality provision for all patients.

What are the potential welfare implications of patient heterogeneity with respect to treatment costs? The first-best quality is given by $q^{fb} = \frac{n(1-\mu)+\mu}{2k}$, which is independent of the treatment cost parameter. For a given p, three scenarios may arise, depending on the price and treatment cost parameters:[21] (i) $q^{*s} < q^{*c} < q^{fb}$; (ii) $q^{*s} < q^{fb} < q^{*c}$; (iii) $q^{fb} < q^{*s} < q^{*c}$.

Under the first scenario, both qualities are lower compared to the first best. If a softer budget implies an increase in both qualities, then it will also imply an increase in welfare since both qualities move closer to the first-best level. If instead a softer budget increases quality for high-cost patients and reduces quality for low-cost patients, then the welfare effect is a priori indeterminate. Notice however that since the quality for high-cost patients is lower than for low-cost patients, a marginal increase in quality generates a higher welfare gain for high-cost patients than for low-cost patients. Therefore, a sufficient condition for a welfare gain is that $\left|\frac{\partial q^{*s}}{\partial \beta}\right| > \left|\frac{\partial q^{*c}}{\partial \beta}\right|$, which arises when the negative profit margin on high-cost patients is higher (in absolute value) than the positive profit margin on low-cost patients.

Under the second scenario, quality is still sub-optimal for high-cost patients but it is over-provided for low-cost patients. This is the scenario that corresponds to the case of 'creaming' and 'skimping', as defined by Ellis (1998); i.e., quality is underprovided for high-cost patients and overprovided for low-cost patients, relative to the social optimum. In this scenario, softer budgets always reduce skimping due to higher quality for high-cost patients. If the cost difference between the two patient types is sufficiently high, softer budgets may also reduce creaming due to lower quality for low-cost patients. In that case, softer budgets unambiguously increase welfare. On the other hand, if the cost difference is sufficiently low, softer budgets will reduce skimping but increase creaming, implying that the welfare effect of softer budgets is indeterminate.

Finally, in the third scenario, both qualities are overprovided. Welfare reduces when a softer budget increases qualities. A welfare increase may arise only under the special case when a softer budget reduces quality for low-cost patients and the associated increase in welfare is higher than the welfare loss due to even more overprovision of quality for high-cost patients.

[21] Notice that $q^{*s}\left(\underline{p}\right) = 0$ and that qualities are increasing in price and decreasing in treatment cost.

5 Concluding remarks

In this paper we have analysed the widespread phenomenon of soft budgets in the hospital sector. Two questions have been asked: (i) What are the incentives for hospitals to be cost efficient and invest in quality when facing soft budgets? (ii) Is there a scope for soft budgets to be welfare improving in the hospital sector? We have set up a model where hospitals face demand uncertainty, and invest in quality and cost-containment effort before the state of demand is revealed. Since prices are fixed (regulated) and hospitals face diseconomies of scale, they are – contrary to most other providers of goods and services – likely to run a deficit (surplus) when demand is high (low). Within this framework we have analysed the effects of bailouts when hospitals run deficits and profit confiscation when hospitals run surpluses.

Our paper provides the following answers. First, profit confiscation and bailout are always harmful for cost-containment effort by the hospitals, as these measures reduce the expected profit from expending such effort. Our finding is consistent with the general claim that bailouts are likely to induce moral hazard in hospital management. Second, profit confiscation is always harmful for quality investments by the hospitals, but the effect of bailout is ambiguous. For a given profit margin, a higher bailout probability increases quality incentives. However, since a higher bailout probability reduces the cost-containment effort, the profit margin becomes smaller, and this weakens the incentives for hospitals to invest in quality. Finally, we show that there might be scope for a bailout policy to be welfare improving. This is the case if (i) prices are not set at the socially optimal level, (ii) the profit margin is sufficiently low, and (iii) the disutility of cost-containment effort is sufficiently high. If socially optimal pricing is feasible, then a strict no-bailout and no-profit-confiscation policy would implement the first-best solution. There may also be scope for welfare-improving bailout policies in the presence of patients' heterogeneity on treatment costs since a softer budget increases quality for high-cost patients and therefore reduces 'skimping'. If the patient heterogeneity is sufficiently large, softer budgets may also reduce quality for low-cost patients and may therefore also lead to less 'creaming'.

Our results are based on several assumptions. First, we assume that hospitals cannot adjust quality and cost-containment effort according to demand fluctuations. If this was the case, then hospitals can avoid running deficits, for instance, by setting quality to a minimum level. We therefore require that these variables involve decisions that are more long-term, such as investments in machinery, medical equipment, training programs, management programs and protocols.

Second, we assume that hospitals are profit maximisers. This assumption is fairly standard and made for analytical convenience, but we are aware that it may not fully capture the objectives of public or not-for-profit hospitals. As some empirical studies indicate, soft budgets might be more pronounced for these ownership structures. However, it is beyond the scope of the current paper to address the issue of ownership structures and soft budgets in the hospital sector, so we leave this issue for future research.

References

[1] Aletras, V.H., 1999. A comparison of hospital scale effects in short-run and long-run cost functions. Health Economics, 8, 521–530.

[2] Beitia, A., 2003. Hospital quality choice and market structure in a regulated duopoly. Journal of Health Economics, 22, 1011–1036.

[3] Brekke, K.R., Nuscheler, R., Straume, O.R., 2006. Quality and location choices under price regulation. Journal of Economics & Management Strategy, 15, 207–227.

[4] Brekke, K.R., Nuscheler, R., Straume, O.R., 2007. Gatekeeping in health care. Journal of Health Economics, 26, 149–170.

[5] Brekke, K.R., Siciliani, L., Straume, O.R., 2008. Competition and waiting times in hospital markets. Journal of Public Economics, 92, 1607–1628.

[6] Brekke, K.R., Siciliani, L., Straume, O.R., 2011. Hospital competition and quality with regulated prices. Scandinavian Journal of Economics, 113, 444–469.

[7] Brekke, K.R., Siciliani, L., Straume, O.R., 2012. Quality competition with profit constraints. Journal of Economic Behavior & Organization, forthcoming.

[8] Duggan, M.G., 2000. Hospital ownership and public medical spending. Quarterly Journal of Economics, 115, 1343–1373.

[9] Eggleston, K., 2008. Soft budget constraints and the property rights theory of ownership. Economics Letters, 100, 425–427.

[10] Eggleston, K., Shen, Y.-C., 2011. Soft budget constraints and ownership: Empirical evidence from US hospitals. Economics Letters, 110, 7–11.

[11] Ellis, R., 1998, Creaming, skimping and dumping: provider competition on the intensive and extensive margins. Journal of Health Economics, 17, 537–555.

[12] Ferguson, B., Sheldon, T., Posnett, J., (eds.), 1999. Concentration and Choice in Health Care. Royal Society of Medicine, London.

[13] Folland, S., Goodman, A.C., Stano, M., 2004. The Economics of Health and Health Care, Prentice Hall, Upper Saddle River, NJ.

[14] Hagen, T.P., Kaarbøe, O.M., 2006. The Norwegian hospital reform of 2002: Central government takes over ownership of public hospitals. Health Policy, 76, 320–333.

[15] Hart, O., Shleifer, S., Vishny, R.W., 1997. The proper scope of government: theory and an application to prisons. Quarterly Journal of Economics, 112, 1127–1161.

[16] Hughes, D., McGuire, A., 2003. Stochastic demand, production responses, and hospital costs. Journal of Health Economics, 22, 999–1010.

[17] Kornai, J., 2009. The soft budget constraint syndrome in the hospital sector. International Journal of Health Care Finance and Economics, 9, 117–135.

[18] Kornai, J., 2001. Hardening the budget constraint: The experience of the post-socialist countries. European Economic Review, 45, 1573–1599.

[19] Kornai, J., Maskin, E., Roland, G., 2003. Understanding the soft budget constraint. Journal of Economic Literature, 41, 1095–1136.

[20] Shen, Y.-C., Eggleston, K., 2009. The effect of soft budget constraints on access and quality in hospital care. International Journal of Health Care Finance and Economics, 9, 211–232.

[21] Siciliani, L., Stanciole, A., Jacobs, R., 2009. Do waiting times reduce hospital costs? Journal of Health Economics, 28, 771–780.