WORKING PAPERS IN ECONOMICS

No. 07/08

ODDVAR M. KAARBØE AND LUIGI **SICILIANI**

Multitasking, Quality and Pay for **Performance**



Multitasking, Quality and Pay for Performance

Oddvar M. Kaarbøe* Luigi Siciliani^{†‡}

27 May 2008

Abstract

We present a model of optimal contracting between a purchaser and a provider of health services when quality has two dimensions. We assume that one dimension of quality is verifiable (dimension 1) and one dimension is not verifiable (dimension 2). We show that the power of the incentive scheme for the verifiable dimension depends critically on the extent to which quality 1 increases or decreases the provider's marginal disutility and the patients' marginal benefit from quality 2 (i.e. substitutability or complementarity). Our main result is that under some circumstances a high-powered incentive scheme can be optimal even when the two quality dimensions are substitutes.

Keywords: quality, altruism, pay for performance. JEL: D82; I11; I18; L51.

^{*}Department of Economics, University of Bergen, Hermann Fossgate 6, 5007 Bergen, Norway; and Health Economics Bergen (HEB). E-mail: oddvar.kaarboe@econ.uib.no; phone +47 5558 9216.

[†]Department of Economics and Related Studies, and Centre for Health Economics, University of York, Heslington, York, YO10 5DD, United Kingdom; and Centre for Economic Policy and Research (C.E.P.R.), 90-98 Goswell Street, London EC1V 7DB, UK. E-mail: ls24@york.ac.uk; phone: +44 1904 432300.

[‡]We would like to thank Eirik Christensen, Hugh Gravelle, Tor Iversen, Rosella Levaggi, Albert Ma, Matt Sutton, and other participants at the 9th European Health Economics Workshop held in Bergen in May 2008 for helpful comments and suggestions.

1 Introduction

Policymakers aim to design incentive schemes that encourage better performance in the health care sector. This is often referred to as *Paying for Performance*. For example, the Medicare Programme in the United States provides higher transfers to hospitals that perform well according to measurable quality indicators, such as rates of cervical cancer screening and hemoglobin testing for diabetic patients (Rosenthal et al., 2005). In the United Kingdom general practitioners who perform well on certain quality indicators, such as the measurement of blood pressure and cholesterol in patients with ischemic heart disease, can receive substantial financial rewards. These can amount to about 20% of a general-practitioner's budget (Doran et al., 2006). Rosenthal et al. (2004) provide 36 other examples of Pay-for-Performance programs in the United States. Similar initiatives are being discussed in Australia, Canada, New Zealand, the Netherlands and Spain (Gravelle, Sutton and Ma, 2008).

The Pay for Performance (P4P) programs outlined above define quality in such a way that it is verifiable. That is, the reimbursement contract between the payer and the provider must be written such that quality indicators can be observed and verified ex post by a third party (e.g. by court). However, a major issue in rewarding performance is that while some quality dimensions are verifiable through performances indicators, other dimensions of quality are not. For example, both communication about medical conditions, and hemoglobin testing affect the quality of care for diabetic patients. While the latter dimension can be verified by a third party, the former dimension is not. Another example can be found in the P4P-program Quality and Outcome Framework (QOF) introduced for UK general practices in 2004 (Gravelle, Sutton and Ma, 2008). While diagnosis factors like blood pressure, body mass index and smoking status are incentivised for diabetes patients (and hence are verifiable), alcohol consumption is not. It is well known from the contract literature that problems of non-verifiability and multi-tasking may impose severe difficulties in effective incentive design (Holmstrom and Milgrom, 1991; Baker 1992).

Recently, Eggleston (2005) has provided a model with two quality dimensions. She

shows that if one dimension of quality is verifiable, while one dimension of quality is not, then the introduction of a P4P-program may increase the verifiable quality dimension, which will increase patients' benefit, but may decrease the non-verifiable one, which will reduce patients' benefit. The overall welfare effect is therefore ambiguous. The purpose of this study is to investigate two related questions: 1) under what conditions is it desirable to introduce a Pay-for-Performance incentive scheme? 2) If the introduction is desirable, how strong should be the power of the optimal incentive scheme?

We show that the optimal incentive scheme depends critically on the extent to which quality 1 (the verifiable dimension) increases or decreases providers' marginal disutility and patients' marginal benefit of quality 2 (the non-verifiable one), i.e. the extent to which quality 1 and 2 are substitutes or complements. Quality dimensions can be substitutes when they are time consuming for the doctor or the provider, so that an increase in one dimension of quality tends to reduce the other dimension. Quality dimensions can be complements in the presence of scope economies or learning by doing: if induced to increase quality in one dimension, the provider becomes better at providing the other dimension as well. The Quality and Outcome Framework program in the UK illustrates other examples of cases where quality dimensions may be susbstitutes and complements. This program mainly focused on the care of people with ten targeted chronic conditions. In a survey with a random national sample of GPs in England conducted before the introduction of the P4P-contract, nearly one-third of the GPs thought that care for patients with acute conditions would deteriorate as a result of the increase in quality for chronic conditions (i.e. these quality dimensions are substitutes; Whalley et al., 2008). Sutton et al. (2008) estimate the possible spillovers from verifiable to non-verifiable quality dimensions in the QOF-contract by analyzing annual rates of recording of clinical effective factors (blood pressure, cholesterol, alcohol consumption, etc.) from 315 general practices over the period 2000-2006. They find that, following the introduction of the QOF, the recording of nonverifiable, clinically-effective factors for the targeted groups increased by 10.9 percentage

¹In a follow-up survey conducted after the introduction of QOF, GPs were less likely to believe that the contract had decreased quality of care for patients with acute disease (Whalley et al, 2008)

points (i.e. quality dimensions are complements).

We might intuitively expect that the incentive scheme will be low powered when quality dimensions are substitutes. This intuition holds true in some circumstances. However, we show that in some cases the optimal incentive scheme can instead be high-powered even when the two quality dimensions are substitutes. Moreover, in other cases it may well arise that the incentive scheme breaks down.

In more detail, we show that if the two quality dimensions are *substitutes*, three possible solutions arise. 1) The incentive scheme *breaks down*: it is not optimal to introduce Pay for Performance as the gain of welfare from an increase in quality dimension 1 is lower than the loss of welfare from a reduction in quality dimension 2. This result arises when the benefits from the quality dimension that is not verifiable are relatively more important.

- 2) The optimal incentive scheme is *low powered*: the price for the verifiable quality 1 is below the marginal benefit of quality 1. Both quality dimensions are positive. This result arises when the benefits from the quality dimension that is verifiable are relatively more important but need to be traded off with the reductions in the quality dimension that is not verifiable.
- 3) The optimal incentive scheme is *high powered*: the price for the verifiable quality 1 is equal to the marginal benefit of quality dimension 1 and the optimal quality in dimension 2 is zero. This result arises when the quality dimension that is not verifiable falls quickly to the minimum when the price is raised, while the benefits from the quality dimension that is verifiable are large. This is, to some extent, a surprising result, as we would intuitively expect the incentive scheme to have low power when quality dimensions are substitutes.

If the two quality dimensions are *complements*, the incentive scheme is always *high* powered. The price for the verifiable quality 1 is above the marginal benefit of quality dimension 1. Both quality dimensions are positive.

We also compare our solutions with what can be obtained if both dimensions of quality are verifiable and the optimal prices are implemented. Obviously, the second-best quality, when quality dimension 2 is not verifiable, is generally different from the first-best quality, when quality dimension 2 is also verifiable, however not necessarily lower. Second-best verifiable quality may be higher in second best if the two dimensions of quality are complements. This follows since providing incentives for the verifiable quality is an indirect way of incentivising the non-verifiable quality.

This study contributes to the literature on provider incentives in health care. Despite the increase in the use of performance indicators, most of the existing theoretical literature assumes that quality is not verifiable (for example Pope, 1989; Ma, 1994; Rogerson, 1994; Ellis, 1998; Ellis and McGuire, 1990; Chalkley and Malcomson, 1998a and 1998b; Mougeot and Naegelen, 2005). As quality indicators become increasingly available, quality becomes partially verifiable. Therefore, there is increasing scope for analysing incentive schemes within this imperfect environment. As far as the authors are aware, this is one of the first attempts to derive the optimal incentive scheme when such indicators are available within the healthcare sector.

The paper is organised as follows. Section 2 introduces the main assumptions of the model and derives the equilibrium price when only one dimension of quality is verifiable. Section 3 provides comparative statics with respect to the price. Sections 4 derives the first-best solution, when both dimensions of quality are verifiable, and compares it with the second-best solution derived in section 2. Section 5 discusses possible extensions of the model. Section 6 concludes.

2 The model

There are two active players, the sponsor (the payer or a purchaser of health services) and the provider (a hospital or a family doctor). The sponsor provides reimbursement to the provider, and the provider exerts effort on two quality tasks. In addition, fully insured patients, whose benefit is increasing in the quality provided on both tasks, seek treatment to the provider. The model is solved by backwards induction, starting with the provider's choice of quality levels.

2.1 The provider

There are two dimensions of quality, q_1 and q_2 . The disutility from exerting quality effort q_1 and q_2 is $\phi(q_1, q_2)$. The disutility is increasing in quality and strictly convex: $\phi_{q_i} > 0$, $\phi_{q_i q_i} > 0$, where $\phi_{q_i} := \partial \phi_i / \partial q_i$ and $\phi_{q_i q_i} := \partial^2 \phi_i / \partial q_i^2$ for i = 1, 2. If the two dimensions of quality are substitutes, then an increase in quality 1 increases the marginal disutility of quality 2 and $\phi_{q_2 q_1} > 0$. If they are complements, an increase in quality 1 reduces the marginal disutility of quality 2 and $\phi_{q_2 q_1} < 0$. We also assume $\phi_{q_1}(0, q_2) = \phi_{q_2}(q_1, 0) = 0$, and $\phi_{q_i q_i} > |\phi_{q_i q_j}|$.

Patients' benefit from receiving quality q_1 and q_2 is $B(q_1, q_2)$ with $B_{q_i} > 0$, and $B_{q_i q_i} \le 0$, i = 1, 2. Patients' benefit increases with quality and is concave. If $B_{q_1 q_2} = 0$ then the two dimensions of quality are independent. If $B_{q_1 q_2} < 0$ then an increase in quality 1 decreases the marginal benefit of quality 2, and the two dimensions of quality are substitutes. If $B_{q_1 q_2} > 0$ then an increase in quality 1 increases the marginal benefit of quality 2, and the two dimensions of quality are complements. We will consider all these three cases, although the presence of complementarity in quality seems more plausible. To simplify the exposition and without loss of generality, we assume that the third-order derivatives on patients' benefit and provider's disutility are zero. We also assume $-B_{q_i q_i} > |B_{q_i q_i}|$.

The incentive scheme is based only on the verifiable dimension of quality q_1 . That is, we assume that no contract on q_2 can be enforced: it is prohibitively costly to specify this outcome ex ante in such a way that it can be verified by a court ex post. Therefore, the payment (price or bonus) can be based only on q_1 and not q_2 . The payment for each unit of verifiable quality q_1 is $p \geq 0$. Below we will refer to p as the price. The provider also receives a lump-sum payment $T \geq 0.2$

The provider is semi altruistic (see Ellis and McGuire, 1986; Chalkley and Malcomson, 1998a; Eggleston, 2005; Jack, 2005). Altruism is captured by the parameter $\alpha \geq 0$.

²For realism we consider linear contracts. This is without loss of generality as the optimal solution derived in section 3 can also be implemented by a non-linear contract or a contract which specifies a fixed transfer conditional on a certain amount of verifiable quality.

Provider's utility from providing quality q_1 and q_2 to a representative patient is

$$U = T + pq_1 + \alpha B(q_1, q_2) - \phi(q_1, q_2) \tag{1}$$

subject to $q_1 \ge 0$, $q_2 \ge 0$. Suppose that both quality dimensions are positive in equilibrium $(q_1 > 0, q_2 > 0)$. Then the optimal levels of quality provided by the provider are given by the following First Order Conditions (FOCs):

$$p + \alpha B_{q_1}(q_1, q_2) = \phi_{q_1}(q_1, q_2) \tag{2}$$

$$\alpha B_{q_2}(q_1, q_2) = \phi_{q_2}(q_1, q_2) \tag{3}$$

The optimal quality for dimension 1 is determined such that the marginal benefit from the price plus the altruistic component are equal to the marginal disutility of providing quality. The optimal quality for dimension 2 is determined such that the marginal benefit from the altruistic component is equal to the marginal disutility.

In the Appendix we show that the Second Order Conditions (SOCs) are satisfied and $U_{q_1q_1}U_{q_2q_2}-U_{q_2q_1}^2>0 \text{ so that}$

$$\frac{\partial q_1}{\partial p} = \frac{-\alpha B_{q_2q_2} + \phi_{q_2q_2}}{U_{q_1q_1}U_{q_2q_2} - U_{q_2q_1}^2} > 0, \quad \frac{\partial q_2}{\partial p} = \frac{\alpha B_{q_1q_2} - \phi_{q_1q_2}}{U_{q_1q_1}U_{q_2q_2} - U_{q_2q_1}^2} \ge 0.$$
(4)

A higher price always increases quality dimension 1 ($\partial q_1/\partial p > 0$). This follows since a higher price increases the provider's marginal benefit of providing the verifiable quality. He therefore responds by increasing q_1 .

The effect of an increase in price on the non-verifiable quality q_2 depends on whether the two quality dimensions are substitutes, independent or complements in patients' benefits and provider's disutility.

Definition 1 The two quality dimensions are substitutes in patients' benefit and in provider's disutility if $\phi_{q_1q_2} > 0$ and $B_{q_1q_2} < 0$, or if $(\alpha B_{q_1q_2} - \phi_{q_1q_2}) < 0$. The two quality dimensions are independent if the patient's benefit and provider's disutility function is separable

in the two quality dimensions $(B_{q_1q_2} = \phi_{q_1q_2} = 0)$, or if $(\alpha B_{q_1q_2} - \phi_{q_1q_2}) = 0$. The two quality dimensions are complements in patients' benefit and in provider's disutility when $\phi_{q_1q_2} < 0$ and $B_{q_1q_2} > 0$, or if $(\alpha B_{q_1q_2} - \phi_{q_1q_2}) > 0$.

From equation (4) it follows that an increase in price decreases quality dimension 2 when the two quality dimensions are substitutes in patients' benefit and in provider's disutility. A higher price increases quality 2 if the two quality dimensions are complements in patient's benefit and in provider's disutility. In this case introducing a positive price is clearly welfare improving for the patients (compared to no price) although there is still an issue of how to set the optimal price. If patients' benefit and provider's disutility function is separable in the two quality dimensions then a higher price has no effect on quality 2 but still increases quality 1. Note that it is the overall effect $(\alpha B_{q_1q_2} - \phi_{q_1q_2})$ which determines the relationship between the quality dimensions: if the two quality dimensions are complements in patients' benefit but are substitutes in provider's disutility function, then the overall effect will depend on the relative strength of the two effects. As a special case, it can happen that the two effects cancel each other out and $(\alpha B_{q_1q_2} - \phi_{q_1q_2}) = 0$.

Finally, if the constraint $q_2 \geq 0$ is binding with strict equality (which arises when $\alpha B_{q_2} - \phi_{q_2} < 0$), then the FOC for quality 1 is:

$$p + \alpha B_{q_1}(q_1, q_2 = 0) = \phi_{q_1}(q_1, q_2 = 0)$$
 (5)

and $\partial q_1/\partial p = 1/(-U_{q_1q_1}) > 0$. Notice that for any positive level of the price p, the non-negativity constraint for quality 1 is never binding in equilibrium as quality 1 is always positive $(q_1 > 0)$. This is not the case for quality 2 which can reach zero for sufficiently high price p, when the two quality dimensions are substitutes.

2.2 The purchaser

The purchaser maximises the difference between patients' benefit and the transfers to the provider $B(q_1, q_2) - T - pq_1$ subject to the participation constraint: $U \ge 0$ or $T + pq_1 \ge 0$

 $\phi(q_1, q_2) - \alpha B(q_1, q_2)$. Since this is binding with equality, the problem becomes:

$$\max_{p>0} W = (1+\alpha)B(q_1(p), q_2(p)) - \phi(q_1(p), q_2(p))$$
(6)

subject to:

$$p + \alpha B_{q_1}(q_1, q_2) - \phi_{q_1}(q_1, q_2) \le 0, \quad q_1 \ge 0, \tag{7}$$

$$\alpha B_{q_2}(q_1, q_2) - \phi_{q_2}(q_1, q_2) \le 0, \quad q_2 \ge 0,$$
 (8)

where the inequalities in the incentive–compatibility constraints hold with complementary slackness. The question is: will a strictly positive price increase the purchaser's utility? The trade-off is that a higher price increases quality in dimension 1 and therefore welfare, but might also reduce quality in dimension 2, which reduces welfare.

The First Order Condition with respect to price, if an interior solution exists (i.e. $q_2 \ge 0$ is not binding with strict equality), is:⁴

$$\frac{dW(q_1(p), q_2(p))}{dp} = [(1+\alpha)B_{q_1} - \phi_{q_1}](\partial q_1/\partial p) + [(1+\alpha)B_{q_2} - \phi_{q_2}](\partial q_2/\partial p) = 0 \quad (9)$$

Using the provider's FOCs $(\alpha B_{q_1} - \phi_{q_1} = -p)$, the optimal price is given by

$$p^* = B_{q_1} + B_{q_2} \frac{\partial q_2/\partial p}{\partial q_1/\partial p} \tag{10}$$

The optimal price is set equal to the marginal benefit of quality 1 adjusted for the the ratio of the responsiveness of the two quality dimensions to price times the marginal benefit of quality 2. From this it follows that the optimal price will be below, equal or above

³We could assume instead that the purchaser maximises a utilitarian welfare function. Define λ as the opportunity cost of public funds. Then a utilitarian welfare function is given by $B - (1 + \lambda) (T + pq_1) + U$, which after substituting for U = 0 and setting $T + pq_1 = \phi - \alpha B$, provides $B(1 + \alpha + \lambda \alpha) - (1 + \lambda) \phi$. This formulation is similar to Boadway, Marchand and Sato (2004). Chalkley and Malcomson (1998a) argue that this formulation leads to double counting of the altruistic component, and that the altruistic component into the welfare function should be excluded. If this approach is followed instead, then the welfare function is: $B(1 + \lambda \alpha) - (1 + \lambda) \phi$. These alternative formulations would not qualitative impact on our main results.

⁴In the Appendix we show that the SOC is satisfied and the problem is well behaved.

the marginal benefit of quality 1 depending on whether the two quality dimensions are substitutes, independent or complements in patients' benefits and provider's disutility. If the two dimensions are substitutes then the optimal price is below the marginal benefit of quality 1: $p^* < B_{q_1}(q_1(p^*), q_2(p^*))$. If a higher price has no effect on quality 2 (i.e the two dimensions are independent), then the price is equal to the marginal benefit of quality 1: $p^* = B_{q_1}(q_1(p^*), q_2)$. Finally, if the two dimensions are complements, then the price is set above the marginal benefit of quality 1: $p^* > B_{q_1}(q_1(p^*), q_2(p^*))$.

If the optimal price is above or equal to the marginal benefit of quality 1, $p^* \geq B_{q_1}$, we call the incentives high-powered. Similarly, if $p^* < B_{q_1}$, then the incentives are low-powered.

Notice that if the two dimensions are substitutes, then there is always a level of price $p = \overline{p}$ such that the level of quality 2 hits zero. In other words, since quality 2 is decreasing in price, there has to be a price high enough to bring quality 2 to zero (the minimum level of enforceable quality). Analytically, if $\partial q_2/\partial p < 0$ then $\exists p = \overline{p}$ such that $q_2 = 0 \ \forall p \geq \overline{p}$ and

$$\frac{dW(q_1(p), q_2(p) = 0)}{dp} \bigg|_{p > \overline{p}} = \left[(1 + \alpha) B_{q_1} - \phi_{q_1} \right] (\partial q_1 / \partial p) \ge 0.$$
 (11)

The point is that when quality 2 is zero and price is above \overline{p} , then a marginal increase in price can be welfare improving (reducing) if the marginal benefit from quality 1 is larger (smaller) than the marginal disutility.

Define p^{sb} as the price under the second-best solution. We define this price second best because one dimension of quality is not verifiable. In section 4 we derive the optimal price under the first best, i.e. when the two dimensions of quality are verifiable.

The following Propositions 1-4 identify the conditions under which the power of the incentive scheme is respectively zero, positive but low, and high. We first investigate the case when it is optimal for the purchaser to set the price equal to zero. In these situations, even if it is possible for the purchaser to write contracts on some dimensions of quality, she prefers not to. Intuitively, this case arises when quality dimension 2 is relatively more important for the sponsor, and when a positive price shifts the provider's choice of quality production towards the first task. The following Proposition 1 provides a sufficient

condition for having no incentive scheme, i.e. for setting $p^{sb}=0$.

Proposition 1 Suppose that: (i) at p = 0, $B_{q_1} \left(\phi_{q_2 q_2} - \alpha B_{q_2 q_2} \right) < B_{q_2} \left(\phi_{q_1 q_2} - \alpha B_{q_1 q_2} \right)$ where $\left(\phi_{q_1 q_2} - \alpha B_{q_1 q_2} \right) > 0$; (ii) $dW(q_1(p), q_2(p) = 0) / dp|_{p = \overline{p}} < 0$. Then, dW(p = 0) / dp < 0 and $p^{sb} = 0$. The incentive scheme breaks down.

Proof. Appendix.

Condition i) in Proposition 1 suggests that quality 2 is relatively more important than quality 1. However, notice that for this condition to hold it is not enough having the marginal benefit of quality 1 smaller than the marginal benefit of quality 2 at p=0. The degree of substitutability between the two quality dimensions $(\phi_{q_1q_2} - \alpha B_{q_1q_2})$ also has to be sufficiently high. If this condition is not met, then the results in Proposition 2 below apply. Condition ii) guarantees that at $p=\overline{p}$ (the price required to bring quality 2 to zero) the marginal gain from quality 1 is below its marginal disutility, and an additional increase in price would reduce welfare. Figure 1.a illustrates the solution.⁵ The solid line provides the welfare when both quality dimensions are allowed to vary, and the dotted line the welfare when quality 2 is set to zero. The two lines cross at $p=\overline{p}$. It is straightforward to verify that the optimal price p^* is in this case equal to zero. The incentive scheme breaks down. The purchaser is better off without the incentive scheme.

Condition (ii) in Proposition 2 is sufficient. As Figure 1.b shows, even if there are gains in quality 1 that can be obtained from an increase in price after quality 2 has hit zero (i.e. even if $dW/dp|_{p=\bar{p}} > 0$), it may still be optimal to have no incentive scheme and set $p^{sb} = 0$. This arises if the welfare when the two dimensions of quality are strictly positive is higher than the welfare when the verifiable quality 1 is set at the level where marginal benefit equals marginal disutility while the non-verifiable quality 2 is set equal to zero (the minimum enforceable level).

[Figure 1.a and 1.b]

⁵ In Figures 1-3 the disutility and benefit functions are assumed to be quadratic: $B(q_1,q_2) = a_1q_1 - (b_1/2)q_1^2 + a_2q_2 - (b_2/2)q_2^2$ and $\phi(q_1,q_2) = (\phi_1/2)q_1^2 + (\phi_2/2)q_2^2 + mq_1q_2$ (see section 4 for details). In Figure 1.a and 1.b we assume $a_1 = a_2 = b_1 = b_2 = \phi_1 = \phi_2 = 1$. For Figure 1.a we set $\alpha = 0.5$ and m = 0.5. For Figure 1.b we set $\alpha = 0.25$ and m = 0.5.

The following Proposition 2 establishes conditions that ensure $p^{sb} = p^*$, where p^* is the price derived in Eq.(10) when q_2 is strictly positive (i.e. the constraint $q_2 \ge 0$ is not binding).

Proposition 2 Suppose that: (i) quality 1 and 2 are substitutes; (ii) $B_{q_1}(p=0) \ge B_{q_2}(p=0)$; (iii) $dW(q_1(p), q_2(p)=0)/dp|_{p=\overline{p}} < 0$. Then, dW(p=0)/dp > 0 and the optimal price is below the marginal benefit of quality 1:

$$p^{sb} = p^* = B_{q_1} - B_{q_2} \left(\phi_{q_1 q_2} - \alpha B_{q_1 q_2} \right) / \left(-\alpha B_{q_2 q_2} + \phi_{q_2 q_2} \right) < B_{q_1}.$$
 (12)

The incentive scheme is low powered.

Proposition 2 suggests that when conditions (i-iii) are met, then the optimal incentive scheme is low-powered and the price is below the marginal benefit of quality dimension 1. Condition ii) requires that if no incentive scheme is introduced (i.e. p = 0), then the marginal benefit from quality dimension 1 is higher than the marginal benefit from quality dimension 2 (quality 1 matters more to the patients than quality 2). Condition iii) says that at $p = \overline{p}$, the price required to bring quality 2 to zero, the marginal gain from quality 1 is below its marginal disutility. In other words, at $p = \overline{p}$ an additional increase in price and in quality 1 would reduce welfare.

This case resembles situations where quality dimension 1 is relatively more important for the purchaser but has a negative impact on the other quality dimension. In these cases, the price is positive but *low-powered* since a too high price will crowd-out valuable quality on the non-verifiable task.

Figure 2.a shows the solution. Again the solid line shows the welfare when both quality dimensions are allowed to vary. The dotted line shows the welfare when quality 2 is set to zero. The two lines cross at $p = \overline{p}$, where quality 2 is equal to zero in both welfare functions. Note that for $p > \overline{p}$ quality 2 is negative along the solid line while it is positive when $p < \overline{p}$. Therefore, the relevant welfare is the solid line for $p < \overline{p}$ and the dotted line

for $p \geq \overline{p}$. Since welfare is decreasing for $p \geq \overline{p}$ along the dotted line, it is straightforward that the welfare-maximising price is $p^{sb} = p^*$.

Conditions i-iii) in Proposition 2 are sufficient but not necessary for $p^{sb} = p^*$. Instead of the sufficient conditions i-ii), a necessary but not sufficient condition for Proposition 1 to hold is the weaker condition: $B_{q_1}\left(-\alpha B_{q_2q_2} + \phi_{q_2q_2}\right) > B_{q_2}\left(\phi_{q_1q_2} - \alpha B_{q_1q_2}\right)$. This condition can hold even when the marginal benefit from quality 1 is smaller than the marginal benefit from quality 2 as long as the degree of substitutability between the two quality dimensions is sufficiently small. The condition guarantees that the introduction of a price is welfare improving, dW/dp(p=0)>0, and that $p^*>0$. However, it does not establish whether p^* generates the maximum welfare. Whether it does it depends on how the welfare function varies with p for $p>\overline{p}$. If $dW/dp|_{p=\overline{p}}<0$ (the marginal welfare from an increase in quality 1 is negative when quality 2 hits zero), then p^* is certainly a maximum, as Figure 2.a has shown. However, this is again a sufficient condition. Figure 2.b shows the case where $dW/dp|_{p=\overline{p}}>0$ (the marginal welfare from an increase in quality 1 is positive when quality 2 hits zero) but $p^{sb}=p^*$ is still a maximum. In Figure 2.b, the price \widetilde{p} denotes the price such that $\widetilde{p}=:\frac{dW(q_1(p),q_2(p)=0)}{dp}\Big|_{p>\overline{p}}=0.6$

If the optimal price $p^{sb} = p^*$ is positive, then the FOC can be rewritten as:

$$[(1+\alpha)B_{q_1} - \phi_{q_1}](\partial q_1/\partial p) = [(1+\alpha)B_{q_2} - \phi_{q_2}](-\partial q_2/\partial p)$$
(13)

The optimal price is such that the marginal welfare gain from an increase in quality dimension 1 is equal to the marginal welfare loss from a reduction in quality dimension 2.

The optimality condition for the price $p^{sb}=p^*$ can also be written in terms of elasticities:

$$\epsilon_{q_1}^W \epsilon_p^{q_1} = \epsilon_{q_2}^W (-\epsilon_p^{q_2}) \tag{14}$$

In Figure 2.a and 2.b we assume $a_1=b_1=\phi_1=\phi_2=1,\ a_2=2,\ b_2=0$ and m=0.5. For Figure 2.a we set $\alpha=0.2$, while $\alpha=0.1$ for Figure 2.b.

where $\epsilon_{q_i}^W = \partial W/\partial q_i(q_i/W)$ denotes the elasticity of welfare with respect to quality dimension i and $\epsilon_p^{q_i} = \partial q_i/\partial p(p/q_i)$ the elasticity of quality i with respect to price. The optimal price is such that the product of the elasticity of welfare with respect to quality and the elasticity of quality with respect to price are equal for each quality dimension.⁷

The following two propositions establish conditions for having *high-powered* incentive schemes. Proposition 3 shows that the optimal incentive scheme can be high powered even though the two quality dimensions are substitutes.

Proposition 3 Suppose that: (i) quality 1 and 2 are substitutes;

(ii)
$$dW(q_1(p), q_2(p) = 0)/dp|_{p=\overline{p}} > 0$$
; (iii) $W(\widetilde{p}) > W(p^*)$ or $W(\widetilde{p}) > W(p = 0)$. Then, $p^{sb} = \widetilde{p} = B_{q_1}$. The incentive scheme is high powered.

Condition ii) ensures that, once quality 2 hits zero, welfare increases with price for $p > \overline{p}$ up to price $p = \widetilde{p}$. In words, when quality 2 is driven to zero, a marginal increase in price p is such that the marginal benefit from quality 1 is bigger than its marginal disutility. This might be the case when the level of altruism is sufficiently low, so that quality 2 quickly drops to zero when price increases. Condition iii) guarantees that $p = \widetilde{p}$ is the global maximum. The purchaser is better off when quality 2 is equal to zero, quality 1 is high and the price is equal to the marginal benefit of quality 1, compared to a scenario where both quality dimensions are positive but low, and the price is set below the marginal benefit of quality 1.

Figures 3.a-3.c show three possible scenarios. In Figure 3.a we have dW(p=0)/dp < 0. In this case an increase in the price reduces welfare for low p because it reduces a lot quality 2. However, reached price $p = \overline{p}$, the level of quality 2 is zero and therefore given assumption ii) welfare increases after that up to price $p = \widetilde{p}$. Condition (iii) guarantees that $p = \widetilde{p}$ is the global maximum.

(ii) welfare increases after that up to price $p = \tilde{p}$. Condition (iii) guarantees that $p = \tilde{p}$ is the global maximum. In Figure 3.a and 3.b it is always the case that $p^* < \bar{p} < \tilde{p}$. Figure 3.c provides an example where $\bar{p} < p^* < \tilde{p}$.

Finally, our next proposition provides the optimal incentive scheme when the two quality dimensions are complements in the provider's disutility or patients' benefit function.

Proposition 4 Suppose that the two quality dimensions are complements. Then,

$$p^{sb} = p^* = B_{q_1} + B_{q_2} \left(-\phi_{q_1q_2} + \alpha B_{q_1q_2} \right) / \left(-\alpha B_{q_2q_2} + \phi_{q_2q_2} \right) > B_{q_1}. \tag{15}$$

The incentive scheme is high powered.

Since the two quality dimensions are complements, the introduction of a positive price increases not only quality 1 but also quality 2. Therefore, in this case the introduction of the price is always welfare improving. However, there is still an issue on how high the price should be set.

Proposition 4 suggests that the optimal price is set above the marginal benefit of quality dimension 1. In this case we do not need to worry about the constraint $q_2 \geq 0$, because it is always satisfied with strict inequality. This is because $dq_2/dp > 0$.

The optimal price is such that

$$[(1+\alpha)B_{q_2} - \phi_{q_2}](\partial q_2/\partial p) = -[(1+\alpha)B_{q_1} - \phi_{q_1}](\partial q_1/\partial p).$$
 (16)

In equilibrium the marginal welfare gain from an increase in quality 2 is equal to the marginal welfare loss from an increase in quality 1. In this case, the price is set at such

⁸In Figure 3.a -3.c we assume $a_1 = b_1 = \phi_1 = \phi_2 = 1$, and m = 0.5. For Figure 3.a we set $a_2 = 1.2$, $b_2 = 0$, $\alpha = 0.05$, for Figure 3.b $a_2 = b_2 = 1$, $\alpha = 0.2$, and in Figure 3.c $a_2 = b_2 = 1$, and $\alpha = 0.1$.

a high level that the marginal benefit from quality 1 is lower than its marginal disutility $((1+\alpha)B_{q_1} < \phi_{q_1})$. It is also the case that the marginal benefit of quality 2 is above its marginal disutility $((1+\alpha)B_{q_2} > \phi_{q_2})$.

Recall from Definition 1 that the two quality dimensions might be complements even when the two quality dimensions are substitutes in the provider's disutility function (if quality dimensions are complements in patients' benefit and the degree of complementarity is sufficiently strong). The empirical results from Sutton et al. (2008) indicate that the complementarity effects between verifiable and non-verifiable quality may be strong. We thus believe that the relevance of this proposition should not be underestimated.

3 Comparative statics

In this section we provide some comparative-statics results in the case where benefit and disutility functions are quadratic. The point we want to make is that even in this simple case the interaction between the quality dimensions in patients' benefits and provider's disutility make the comparative-statics results quite complex.

Suppose that $B(q_1, q_2) = a_1q_1 - (b_1/2)q_1^2 + a_2q_2 - (b_2/2)q_2^2$ and $\phi(q_1, q_2) = (\phi_1/2)q_1^2 + (\phi_2/2)q_2^2 + mq_1q_2$, so there are no interaction effects in benefits. By solving the provider's First Order Conditions (equation (2) and (3)) for the quality levels we obtain:

$$q_{1}^{*}(p) = \frac{(p + \alpha a_{1}) (\phi_{2} + \alpha b_{2}) - m\alpha a_{2}}{\phi_{1}\phi_{2} + \alpha b_{1}\phi_{2} + \alpha b_{2}\phi_{1} + \alpha^{2}b_{1}b_{2} - m^{2}}$$

$$q_{2}^{*}(p) = \frac{\alpha a_{2} (\phi_{1} + \alpha b_{1}) - m (p + \alpha a_{1})}{\phi_{1}\phi_{2} + \alpha b_{1}\phi_{2} + \alpha b_{2}\phi_{1} - m^{2} + \alpha^{2}b_{1}b_{2}}$$
(17)

Then, the optimal price is (follows from equation (10)):

$$p^{sb} = p^* = B_{q_1} - B_{q_2} \frac{m}{\alpha b_2 + \phi_2}. (18)$$

First, we consider the case when the marginal benefit is constant. In this case the optimal price is decreasing (increasing) in $\phi_{q_1q_2} = m > 0$ (< 0). That is, it is decreasing

(increasing) in m when the two quality dimensions are substitutes (complements) in the provider's disutility function. Then we show that if marginal benefits are decreasing, the optimal price can be increasing in $\phi_{q_1q_2} = m$ even in the case where the quality dimensions are substitutes in the provider's disutility function, i.e. when m > 0.

3.1 Constant marginal benefit

Suppose that the marginal benefit is constant $(b_1 = b_2 = 0)$. Then the optimal price is

$$p^{sb} = p^* = a_1 - m\frac{a_2}{\phi_2} \tag{19}$$

The optimal price p^{sb} is increasing in the marginal benefit of quality 1 and decreasing in the marginal benefit of quality 2. The price is decreasing in m, as expected: the more the two quality dimensions are substitutes, the smaller is the price $(\partial p/\partial m < 0)$.

The higher the marginal disutility of dimension 2 the higher is the price $(\partial p/\partial \phi_2 = a_2 m/\phi_2^2 > 0)$. This is somewhat counter-intuitive, but follows from the fact that an increase in the marginal disutility of quality 2 reduces quality 2 and thus the marginal disutility of quality 1, which ultimately increases the level of quality 1.

Finally, notice that price does not vary with altruism (since marginal benefits are constant) nor with the marginal disutility of quality 1 (as price is equal to marginal benefit). Note that the above results also hold for small b_1 and b_2 (i.e. for $b_1 \to 0$ and $b_2 \to 0$).

Substituting the optimal price into the FOCs of the provider, we obtain

$$q_1^* = \frac{(1+\alpha)(a_1\phi_2 - ma_2)}{\phi_1\phi_2 - m^2}; \ q_2^* = \frac{\alpha(a_2\phi_1 - ma_1) - \frac{m}{\phi_2}(a_1\phi_2 - ma_2)}{\phi_1\phi_2 - m^2}.$$
 (20)

The following corollary establishes how the optimal levels of quality vary with the different parameters.

Corollary 1 Suppose $p^* > 0$. (a) $\partial q_i^* / \partial m \ge 0$ with i = 1, 2; (b) $\partial q_i^* / \partial a_i > 0$ with i = 1, 2; (c) $\partial q_i^* / \partial a_j < 0$ with j = 1, 2 and $i \ne j$; (d) $\partial q_i^* / \partial \phi_i < 0$ with i = 1, 2; (e)

 $\partial q_i^*/\partial \phi_j > 0$ with j = 1, 2 and $i \neq j$; (f) $\partial q_i^*/\partial \alpha > 0$ with i = 1, 2;.

Proof. Appendix.

To understand these results notice that the effect of each parameter on quality reflects the sum of the direct effect on quality plus the indirect effect through the price (see equation (17)). In most cases the indirect effect reinforces the direct effect and the results are unambiguous. For example, each quality level is decreasing in the marginal benefit of the other quality dimension $(\partial q_i^*/\partial a_j < 0, i = j = 1, 2 \text{ and } i \neq j)$. An increase in a_1 decreases q_2^* for two reasons: for a given price, a higher marginal benefit of quality 1 decreases quality 2 but also implies a higher price which also decreases quality 2 (case c).

In case (a) however, the direct and indirect effect counteract.⁹ A higher m implies a more negative spillover effect of a high level of q_1 on the marginal disutility of providing quality 2. This effect tends to reduce quality 1. However, a higher m also implies a tendency to reduce q_2 : this effect tends to increase q_1 . Which effect dominates depends on the relative size of the marginal disutility of producing q_1 and q_2 , and the relative marginal benefits (a_1 and a_2). If the relative benefits favour quality dimension 1 (a_1 large relative to a_2 , and large relative to marginal disutility) then q_1 tends to increase with m, while q_2 tends to decrease with m. A similar result occurs if the marginal disutility of providing q_1 is relatively small to the marginal disutility of providing q_2 (and the difference in marginal benefits is small).

We now comment briefly on the other cases.

- (b) Each quality level is increasing in its marginal benefits $(\partial q_i^*/\partial a_i > 0, i = 1, 2)$: a higher marginal benefit from quality implies a higher price but also a stronger altruistic component for the provider. Both effects induce higher quality in equilibrium.
- (d) Each quality level is also decreasing in its own marginal disutility, i.e. $\partial q_i^*/\partial \phi_i < 0$, i = 1, 2. A higher marginal disutility for quality 1, ϕ_1 , decreases quality 1 (direct effect) and has no effect on price (follows from equation (19)). A higher marginal disutility for quality 2, ϕ_2 , decreases quality 2 (direct effect). In this case there is also an indirect effect

⁹The effects of an increase in m is symmetric so we only consider the effect on q_1 .

via the price; an increase in ϕ_2 increases the price which further decreases quality 2.

- (e) Each quality level is increasing in the marginal disutility of the other quality dimension. A higher marginal disutility for quality 2, ϕ_2 , reduces the optimal quality 2 and therefore reduces the marginal disutility of quality 1 which increases quality 1; moreover it implies a higher price which also increases quality 1. Similarly, a higher marginal disutility of quality 1, ϕ_1 , reduces the optimal quality 1 and therefore reduces the marginal disutility of quality 2 which increases quality 2 (there is no effect through the price).
- (f) Higher altruism increases the marginal benefit of quality and therefore increases quality (direct effect) and has no effect on the price (due to the assumption of constant marginal benefit).

3.2 Decreasing marginal benefit

We now consider the case with decreasing marginal benefit. The point we want to make is that the effects of an increase in the degree of substitutability in the disutility function, $\phi_{q_1q_2}=m$, are quite complicated and most often indeterminate when the marginal benefit is decreasing. For example, it might indeed be the case that the price is actually increasing in m. This happens when marginal benefits decrease sufficiently fast.

The optimal price is now given by:

$$p^{sb} = a_1 - b_1 q_1(p^{sb}) - \left[a_2 - b_2 q_2(p^{sb})\right] \frac{m}{\alpha b_2 + \phi_2}$$
(21)

Consider the effect of an increase in m > 0. Totally differentiating we obtain:

$$\left[1 + b_1 \frac{\partial q_1(p^{sb})}{\partial p} - \frac{b_2 m}{\alpha b_2 + \phi_2} \frac{\partial q_2(p^{sb})}{\partial p}\right] dp^{sb}
+ \left[\frac{a_2 - b_2 q_2(p^{sb})}{\alpha b_2 + \phi_2} + b_1 \frac{\partial q_1}{\partial m} - \frac{b_2 m}{\alpha b_2 + \phi_2} \frac{\partial q_2}{\partial m}\right] dm = 0$$
(22)

Now, since $\partial q_1(p^{sb})/\partial p > 0$ and $\partial q_2(p^{sb})/\partial p < 0$, then

$$sign\frac{dp^{sb}}{dm} \iff sign\left[-\frac{a_2 - b_2 q_2(p^{sb})}{\alpha b_2 + \phi_2} - b_1 \frac{\partial q_1}{\partial m} + b_2 \frac{m}{\alpha b_2 + \phi_2} \frac{\partial q_2}{\partial m}\right]. \tag{23}$$

If $b_1 = b_2 = 0$ we obtain as a special case the previous result, so that $dp^{sb}/dm < 0$.

To show that the optimal price can increase in m>0 when the marginal benefit decreases sufficiently fast, let $\alpha=0.8$, $a_1=a_2=b_1=b_2=2$, and let $\phi_1=\phi_2=1$. Figure 4 shows that $dp^{sb}/dm>0$ for some values of m>0.10

To see why this result may arise, consider the case where the marginal benefit of quality 2 is constant, and the marginal benefit of quality 1 is decreasing (i.e., $b_2 = 0$, and $b_1 > 0$). Moreover, suppose that $\partial q_1/\partial m < 0$, i.e. higher substitutability reduces the verifiable quality (recall that in general $\partial q_1/\partial m$ is indeterminate even when marginal benefit is constant). Then, since the marginal benefit of quality 1 is decreasing, the purchaser reacts to the lower verifiable quality by adjusting upwards the price. If this effect is sufficiently strong then $dp^{sb}/dm > 0$.

[Figure 4]

4 Comparison with First Best

In this section we first define the first best solution and then compare the results obtained in Propositions 1-4, which we refer to as the second-best solution, with the first-best solution. Our main result below is that if the marginal benefit of quality is constant, then the price under the first-best solution will be higher than in the second best if the two quality dimensions are substitutes, as we might intuitively expect. Furthermore, if patients' benefit is symmetric in both quality dimensions, then both aggregate quality and patients' benefit are higher in first best.

4.1 First best

We define the first-best solution a setting where the purchaser can observe both quality dimensions and maximize over the quality levels directly. This is equivalent to set two different prices p_1 and p_2 for q_1 and q_2 respectively.

 $^{^{10}}$ With these parameter values $q_1 > 0, q_2 > 0$, and the SOC is fullfiled.

The purchaser maximises the difference between patients' benefit and the transfers to the provider $B(q_1, q_2) - T - p_1q_1 - p_2q_2$ subject to the participation constraint: $U \ge 0$ or $T + p_1q_1 + p_2q_2 \ge \phi(q_1, q_2) - \alpha B(q_1, q_2)$. Since this is binding with equality, the problem becomes:

$$\max_{p_1 \ge 0, p_2 \ge 0} W = (1 + \alpha)B(q_1(p_1), q_2(p_2)) - \phi(q_1(p_1), q_2(p_2))$$
(24)

subject to the incentive-compatibility (IC) constraints (which follow from the provider's First Order Conditions):

$$p_1 + \alpha B_{q_1}(q_1, q_2) - \phi_{q_1}(q_1, q_2) \le 0, \quad q_1 \ge 0$$
 (25)

$$p_2 + \alpha B_{q_2}(q_1, q_2) - \phi_{q_2}(q_1, q_2) \le 0, \quad q_2 \ge 0.$$
 (26)

The First Order Conditions with respect to price are:

$$\frac{dW(q_1(p_1), q_2(p_2))}{dp_1} = \left[(1 + \alpha)B_{q_1} - \phi_{q_1} \right] (\partial q_1/\partial p_1) = 0$$
(27)

$$\frac{dW(q_1(p_1), q_2(p_2))}{dp_2} = \left[(1+\alpha)B_{q_2} - \phi_{q_2} \right] (\partial q_2/\partial p_2) = 0$$
 (28)

Using the FOCs for the provider (the ICs) we obtain

$$p_i^{fb} = B_{q_i} \left(q_1^{fb}, q_2^{fb} \right), \qquad i = 1, 2$$
 (29)

$$q_i^{fb}$$
: $(1+\alpha)B_{q_i} = \phi_{q_i}$, $i = 1, 2$. (30)

Hence, the price of each quality dimension is set equal to the marginal benefit this dimension generates. Furthermore, the optimal level of quality is such that the marginal benefit of quality (weighted for the altruistic component) is equal to the marginal disutility.

4.2 Comparison of first best and second best

We start by comparing prices of quality dimension 1. However since both the marginal benefit and the marginal disutility of quality 1 depends on the level of quality 2 we are

not able to compare prices and quality levels without making further assumptions. To compare prices we assume that marginal benefit of quality dimension 1 is constant, and that marginal benefit of quality 1 is independent of quality 2. The following corollary compares solutions.

Corollary 2 Suppose the marginal benefit is constant, $B_{q_1q_1} = B_{q_1q_2} = 0$. i) If the conditions in Propositions 1-2 hold then $p_1^{fb} > p^{sb}$. ii) If the conditions in Proposition 3 hold then $p_1^{fb} = p^{sb}$. iii) If the conditions in Proposition 4 holds, then $p_1^{fb} < p^{sb}$.

The Corollary shows that the second-best price coincides with first-best price only when quality 2 is zero in the second best. However, the real allocations, i.e. the choice of quality differs in first- and second best also in this case. This follows since $q_2^{fb} > q_2^{sb} = 0$, which implies that welfare will differ in first- and second-best solution. We now turn to comparing the level of quality under the two settings.

Comparing the levels of quality is not straightforward. The problem is that even if we can rank the prices for quality 1, the marginal disutility of quality 1 depends on the level of quality 2. To compare the levels of quality we impose the following restrictions. First, we assume that marginal benefit is constant and that the benefit function is symmetric, $B_{q_1q_1} = B_{q_2q_2} = B_{q_1q_2} = 0$, so that $B_{q_1} = B_{q_2} = B$. Second, let the disutility function be symmetric, $\phi_{q_1q_1} = \phi_{q_2q_2}$. (For simplicity) let the third-order derivatives on disutility be zero so $\phi_{q_iq_i} = \phi > \phi_{q_iq_j} = m > 0$, where the inequality follows from the second-order conditions of the provider's maximization problem.¹¹

Under these assumptions it follows from equation (29) and (30) that prices and quality for both tasks are identical in the first best. Furthermore, it follows from the provider's first-order conditions (equation (2) and (3)) that $\phi_{q_1^{sb}}(q_1^{sb}, q_2^{sb}) > \phi_{q_2^{sb}}(q_1^{sb}, q_2^{sb})$ if and only if $p^{sb} > 0$. The following lemma compares quality levels in second best.

 $^{^{11} \}text{Note that the Hessian is } |H| \equiv \left| \begin{array}{cc} -\phi_{q_1q_1} & -\phi_{q_1q_2} \\ -\phi_{q_1q_2} & -\phi_{q_2q_2} \end{array} \right| = \phi_{q_1q_1}\phi_{q_2q_2} - \left(\phi_{q_1q_2}\right)^2 > 0 \text{ from SOCs. Hence } \phi_{q_iq_i} > \phi_{q_iq_j}.$

Lemma 1 Let i) $B_{q_1q_2} = B_{q_1q_1} = B_{q_2q_2} = 0$, ii) $B_{q_1} = B_{q_2}$, and iii) $\phi_{q_1q_1} = \phi_{q_2q_2}$. If $p^{sb} > 0$, then $\phi_{q_1^{sb}}(q_1^{sb}, q_2^{sb}) > \phi_{q_1^{sb}}(q_1^{sb}, q_2^{sb}) \Rightarrow q_1^{sb} > q_2^{sb}$.

Hence, when the disutility function and benefit function are symmetric in the two quality dimensions, quality for the verifiable dimension is higher than for non-verifiable dimension if the second best price is strictly positive.

Since the marginal benefit is constant and the benefit function is symmetric, marginal benefit is equal for both quality dimensions both in first and second best. Hence we are able to compare the marginal disutility in first best and second best. From the first order conditions (equations (2), (3) and (30)) we have

$$B = \frac{\phi_{q^{fb}}(q^{fb}, q^{fb})}{(1+\alpha)} = \frac{\phi_{q_1^{sb}}(q_1^{sb}, q_2^{sb})}{(1+\alpha-\Phi)} = \frac{\phi_{q_2^{sb}}(q_1^{sb}, q_2^{sb})}{\alpha}.$$
 (31)

where $\Phi = \phi_{q_1q_2}/\phi_{q_1q_1} > 0$. We thus obtain the following ranking of marginal disutility (the last inequality follows from Lemma 1)

$$\phi_{q^{fb}}(q^{fb}, q^{fb}) > \phi_{q_1^{sb}}(q_1^{sb}, q_2^{sb}) > \phi_{q_2^{sb}}(q_1^{sb}, q_2^{sb}).$$
 (32)

Obviously, the conditions given in (32) holds for $q^{fb} > q_1^{sb} > q_2^{sb}$. The following proposition gives an upper boundary for q_2^{sb} in the case where $q^{fb} < q_1^{sb}$.

Proposition 5 Suppose the conditions in Lemma 1 hold and $q_1^{sb} > q_2^{sb}$. If $q^{fb} < q_1^{sb}$ then $q_1^{sb} < q_1^{fb} - \frac{\phi}{m} \left(q_1^{sb} - q_1^{fb} \right) < q_1^{fb}$, $q_1^{sb} + q_1^{sb} < 2q^{fb}$, and iii) $B(q_1^{fb}, q_2^{fb}) > B\left(q_1^{sb}, q_2^{sb} \right)$.

Hence, if quality 1 is higher in the second best than in the first best, $q^{fb} < q_1^{sb}$, then quality 2 is lower in the second best compared to the first best, $q_2^{sb} < q^{fb}$. Furthermore, aggregate quality and patients' benefit is lower in the second best compared with the first best.

5 Discussion

We now comment upon two of the assumptions we have imposed. The first one is that provider costs are measured in disutility, and not in monetary costs; the second is that both verifiable and non-verifiable quality are one-dimentional.

In the analysis presented in sections 2-4 we have assumed that providers costs are non-monetary, i.e. that increasing quality in each dimension raises the disutility of the provider. The main hints from our analysis also hold if we reinterpret costs as monetary. However, if costs are monetary, a limited-liability constraint needs to be added to the problem of the purchaser in addition to the participation constraint. Limited liability implies that the provider cannot make losses. To make the point more clearly, suppose that instead of the disutility function $\phi(q_1, q_2)$, we assume that all costs are monetary and are captured by the function $C(q_1, q_2)$, with $C_{q_i} > 0$, $C_{q_i q_i} > 0$, $C_{q_i q_j} > 0$ if quality dimensions are substitutes, and $C_{q_iq_j} < 0$ if quality dimensions are complements. Since the limited-liability constraint is binding, while the participation constraint is not for any positive level of altruism, the purchaser maximises $W = B(q_1(p), q_2(p)) - C(q_1(p), q_2(p))$ subject to the usual incentive-compatibility constraints. If both quality dimensions are non-negative, the optimal price is now given by: $p^* = (B_{q_1} + B_{q_2} \frac{\partial q_2/\partial p}{\partial q_1/\partial p})(1-\alpha)$. The optimal price now depends also directly on the degree of altruism, with higher altruism implying a lower price. However, the qualitative features of the optimal price remain unchanged.

In this paper we have assumed that both the verifiable and non-verifiable quality is one-dimensional. This will typically not be the case. For example the contract for UK general practices introduced in 2004 rewarded practices according to performance on 146 indicators. Clearly, also non-verifiable quality will typically be multidimensional. One way to capture this in the model is to assume that $\mathbf{q}_1 = (q_{11}, q_{21}, ..., q_{n1})$ is an n-dimensional vector of quality, $\mathbf{p}\mathbf{q}_1 = \omega_1q_{11} + \omega_2q_{21} + ... + \omega_nq_n$ is the scalar product, and ω_i , i = 1, ...n is the weight given to quality dimension i. Similarly, $\mathbf{q}_2 = (q_{12}, q_{22}, ..., q_{m2})$ is

¹²The weights correspond to the value of the points given in the UK-GP contract (Doran et al, 2006).

an *m*-dimensional vector of non-verifiable quality. By solving the provider's maximisation problem, and using the implicit-function theorem we obtain the comparative statics results that determine if the vectors of quality dimensions are complements or substitutes. What matters is not how each element of the verifiable quality vector relates to one other element of the non-verifiable vector, but the *overall* interaction effect.

6 Conclusions

Purchasers make increased use of pay-for-performance incentive schemes in the attempt of fostering quality in the health care sector. However, inevitably some dimensions of quality remain not verifiable. Existing incentive schemes have been criticised on the ground that paying for quality will increase quality in the dimensions that are verifiable but will reduce quality for the dimensions that are not verifiable (Roland, 2004; Whalley et al., 2005). This criticism then raises the question whether such incentive schemes should be introduced, and if introduced how powered should be the incentive schemes.

We show that in some cases *low powered* incentive schemes are optimal. Introducing the scheme is useful in increasing welfare when the quality that is verifiable is relatively important. However, this needs to be traded-off with the reductions in the quality dimension that is not verifiable. In other cases it is optimal not to introduce an incentive scheme. This is likely to be the case when the quality dimensions that are not verifiable are relatively more important.

Finally, there are some cases where the optimal incentive scheme is *high powered*. This arises in two circumstances. First, if the quality dimension that is not verifiable falls to zero quickly with price (due for example to low altruism), the benefit from increasing the quality in the dimensions that are verifiable can be quite large. Second, if the different quality dimensions face some complementarity, then providers become better at providing also the dimensions of quality that are not verifiable, when induced to increase the quality dimensions that are verifiable.

The incentive schemes that have been recently implemented in the US and the UK,

suggest that these schemes are reasonably high powered. Our model suggest that this policy is optimal when one of these conditions hold: i) quality dimensions are independent so that an increase in one does not come at the cost of the other; ii) there are some elements of substitutability and complementarity: for example it could be that they are substitutes on costs but they are complements on patients' benefits (i.e. the marginal benefit from the non-verifiable quality increases when the verifiable one is increased); iii) the non-verifiable quality is perceived by the purchaser to be so low, that the marginal gains from an increase in the verifiable quality overcome the losses from a reduction in the non-verifiable one.

Epstein (2006) argues that policy changes might lead to unexpected consequences, such as higher payments to physicians and increased budget deficits. In the model outlined in this paper providers' utility is the same (since the purchaser adjust the lump-sum payments so that the provider's participation constraint holds with equality). However, the total payments will change when paying for performance is introduced. This follows since the provided levels of quality are (weakly) higher under a regime of paying for performance. Note that this will be an intended effect. In some cases the introduction of paying for performance programs are not matched with a decrease in lump-sum payments, and provider payments might increase unintentionally. The introduction of the paying for performance program for UK family practitioners in 2004 might be such an example since it increased the gross income of the family practitioners by about 20% (Doran et al., 2006). We believe that a generous performance scheme will reduce resistance against P4P-programs, and hence facilitate their introduction. In the long run we would however expect payments to be adjusted to counteract the unintended consequences observed at the time paying-for-performance schemes are introduced.

7 References

Atlanta Information Services, 2003, Case studies on physician compensation in medical groups, Washington DC.

Baker, G., 1992, Incentive contracts and performance measurement, *Journal of Political Economy*, 100(3), 598–614

Boadway, R., Marchand M., Sato M., 2004, An optimal contract approach to hospital financing, *Journal of Health Economics*, 23(1), 85-110.

Chalkley, M., Malcomson J.M., 1998a, Contracting for health services when patient demand does not reflect quality, *Journal of Health Economics*, 17, 1-19.

Chalkley, M., Malcomson J.M., 1998b, Contracting for health services with unmonitored quality, *The Economic Journal*, 108, 1093-1110.

Doran, T., Fullwood C., Gravelle H., Reeves D., Kontopantelis E., Hiroeh U., Roland M., 2006, Pay for performance programs in family practices in the United Kingdom, *New England Journal of Medicine*, 355(4), 375-384.

Ellis, R.P., McGuire T.G., 1990, Optimal payment systems for health services, Journal of Health Economics, 9 (4), 375–396.

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

Eggleston, K., 2005, Multitasking and mixed systems for provider payment, *Journal* of *Health Economics*, 24(1), 211-223.

Gravelle, H., Sutton M., Ma A., 2008, Doctor Behaviour Under a Pay for Performance Contract: Further Evidence from the Quality and Outcomes Framework, CHE Research Paper 34, Centre for Health Economics, University of York.

Holmstrom, B., Milgrom P., 1991, Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design, *Journal of Law, Economics and Organization*, 7, 24–52.

Jack, W., 2005, Purchasing health care services from providers with unknown altruism, Journal of Health Economics, 24(1), 73-93.

Ma, C.A., 1994, Health care payment systems: Cost and quality incentives, *Journal of Economics and Management Strategy*, 3(1), 93-112.

Mougeot, M., Naegelen F., 2005, Hospital price regulation and expenditure cap policy,

Journal of Health Economics, 24(1), 55–72.

Pope, G., 1989, Hospital nonprice competition and Medicare reimbursement policy, *Journal of Health Economics*, 8 (2), 147–172.

Rogerson, W.P., 1994, Choice of treatment intensities by a non-profit hospital under prospective pricing. *Journal of Economics and Management Strategy*, 3 (1), 7–51.

Roland, M., 2004, Linking physician pay to quality of care: a major experiment in the UK, New England Journal of Medicine, 351, 1488-54.

Rosenthal, M.B., Fernandopulle R., Song H.R., and Landon B., 2004, Paying for quality: Providers' incentives for quality improvement, *Health Affairs*, 23, 127-141.

Rosenthal, M.B., Frank R.G., Li A., Epstein A.M., 2005, Early experience with payfor-performance. From concept to practice, *Journal of the American Medical Association*, 294, 1788-1793.

Sutton, M., Elder R., Guthrie B., Watt G., 2008, Record rewards: the effect on risk factor monitoring of the new financial incentives for UK general practices, mimeo.

Whalley, D., Gravelle H., Sibbald B., 2008, Effect of the new contract on GPs' working lives and perceptions of quality of care, *British Journal of the General Practice*, January, 8-14.

8 Appendix

In this appendix we provide details regarding some of the calculations in this paper.

The Second Order Conditions (SOCs) of the provider's problem are:

$$U_{q_1q_1} = \alpha B_{q_1q_1} - \phi_{q_1q_1} < 0, \quad U_{q_2q_2} = \alpha B_{q_2q_2} - \phi_{q_2q_2} < 0$$

$$U_{q_1q_1}U_{q_2q_2} - U_{q_2q_1}^2 = \left(\alpha B_{q_1q_1} - \phi_{q_1q_1}\right) \left(\alpha B_{q_2q_2} - \phi_{q_2q_2}\right) - \left(\alpha B_{q_1q_2} - \phi_{q_1q_2}\right)^2 > 0$$

A sufficient but not necessary condition for $U_{q_1q_1}U_{q_2q_2} - U_{q_2q_1}^2 > 0$ to be satisfied is $\phi_{q_1q_1}\phi_{q_2q_2} > \phi_{q_1q_2}^2$ and $B_{q_1q_1}B_{q_2q_2} > B_{q_1q_2}^2$.

To find the effects $\partial q_i/\partial p$, i=1,2 we use Cramer's rule. Define

$$F^{1}(q_{1}, q_{2}; p, \alpha) \equiv p + \alpha B_{q_{1}}(q_{1}, q_{2}) - \phi_{q_{1}}(q_{1}, q_{2}) = 0$$
$$F^{2}(q_{1}, q_{2}; p, \alpha) \equiv \alpha B_{q_{2}}(q_{1}, q_{2}) - \phi_{q_{2}}(q_{1}, q_{2}) = 0.$$

The Jacobian determinant is

$$|J| \equiv \begin{vmatrix} \alpha B_{q_1q_1} - \phi_{q_1q_1} & \alpha B_{q_1q_2} - \phi_{q_1q_2} \\ \alpha B_{q_1q_2} - \phi_{q_1q_2} & \alpha B_{q_2q_2} - \phi_{q_2q_2} \end{vmatrix} = U_{q_1q_1} U_{q_2q_2} - U_{q_2q_1}^2 > 0,$$

where the last inequality follows from the assumption that the SOC is satisfied.

Since $-\partial F^1/\partial p = -1$, and $-\partial F^2/\partial p = 0$ we obtain

$$\frac{\partial q_1}{\partial p} = \frac{\begin{vmatrix} -1 & \alpha B_{q_1q_2} - \phi_{q_1q_2} \\ 0 & \alpha B_{q_2q_2} - \phi_{q_2q_2} \end{vmatrix}}{|J|} = \frac{-\alpha B_{q_2q_2} + \phi_{q_2q_2}}{U_{q_1q_1}U_{q_2q_2} - U_{q_2q_1}^2} > 0,$$

$$\frac{\partial q_2}{\partial p} = \frac{\begin{vmatrix} \alpha B_{q_1q_2} - \phi_{q_1q_1} & -1 \\ \alpha B_{q_1q_2} - \phi_{q_1q_2} & 0 \\ |J| & = \frac{\alpha B_{q_1q_2} - \phi_{q_1q_2}}{U_{q_1q_1}U_{q_2q_2} - U_{q_2q_1}^2}.$$

The SOC of the sponsor's problem is satisfied and the problem is well behaved since

$$\frac{d^2W}{dp^2} = \left[(1+\alpha)B_{q_1q_1} - \phi_{q_1q_1} \right] (\partial q_1/\partial p)^2 + \left[(1+\alpha)B_{q_2q_2} - \phi_{q_2q_2} \right] (\partial q_2/\partial p)^2 < 0.$$

Proof of Proposition 1. $\frac{dW(q_1(p),q_2(p))}{dp}\Big|_{p<\overline{p}} = \left[(1+\alpha)B_{q_1} - \phi_{q_1}\right](\partial q_1/\partial p) + \left[(1+\alpha)B_{q_2} - \phi_{q_2}\right](\partial q_2/\partial p).$ Using the FOCs $p + \alpha B_{q_1} = \phi_{q_1}$; $\alpha B_{q_2}(q_1,q_2) = \phi_{q_2}(q_1,q_2)$, we have

$$\begin{split} & \frac{dW(q_1(p), q_2(p))}{dp} \bigg|_{p < \overline{p}} \\ &= \left. (B_{q_1} - p) \left(\partial q_1 / \partial p \right) + B_{q_2} \left(\partial q_2 / \partial p \right) \right. \\ &= \left. \frac{\left(B_{q_1} - p \right) \left(\phi_{q_2 q_2} - \alpha B_{q_2 q_2} \right) + B_{q_2} \left(\alpha B_{q_1 q_2} - \phi_{q_1 q_2} \right)}{U_{q_1 q_1} U_{q_2 q_2} - U_{q_2 q_1}^2} \\ &= \frac{B_{q_1} \left(-\alpha B_{q_2 q_2} + \phi_{q_2 q_2} \right) + B_{q_2} \left(-\phi_{q_1 q_2} + \alpha B_{q_1 q_2} \right) - p \left(-\alpha B_{q_2 q_2} + \phi_{q_2 q_2} \right)}{U_{q_1 q_1} U_{q_2 q_2} - U_{q_2 q_1}^2} \end{split}$$

At p=0 the condition is $\frac{dW(q_1(p),q_2(p))}{dp}\Big|_{p<\overline{p}} = \frac{B_{q_1}(\phi_{q_2q_2}-\alpha B_{q_2q_2})+B_{q_2}(\alpha B_{q_1q_2}-\phi_{q_1q_2})}{U_{q_1q_1}U_{q_2q_2}-U_{q_2q_1}^2}$, which is negative when the necessary condition (i) in Proposition 1 is satisfied.

Proof of Proposition 2.

$$=\frac{\frac{dW(q_{1}(p), q_{2}(p))}{dp}\Big|_{p<\overline{p}}}{U_{q_{1}q_{1}}(-\alpha B_{q_{2}q_{2}} + \phi_{q_{2}q_{2}}) + B_{q_{2}}(-\phi_{q_{1}q_{2}} + B_{q_{1}q_{2}}) - p\left(-\alpha B_{q_{2}q_{2}} + \phi_{q_{2}q_{2}}\right)}{U_{q_{1}q_{1}}U_{q_{2}q_{2}} - U_{q_{2}q_{1}}^{2}}$$

Therefore, at p = 0 the condition is positive when (i) and (ii) in Proposition 2 are satisfied.

Proof of Proposition 3.
$$\frac{dW(q_1(p),q_2(p)=0)}{dp}\Big|_{p\geq \overline{p}} = \left[(1+\alpha)B_{q_1} - \phi_{q_1}\right](\partial q_1/\partial p). \text{ Using the FOC } p + \alpha B_{q_1} = \phi_{q_1}, \text{ then } \left.\frac{dW(q_1(p),q_2(p)=0)}{dp}\right|_{p\geq \overline{p}} = (B_{q_1}-p)\left(\partial q_1/\partial p\right) = 0, \text{ which implies } \widetilde{p} = B_{q_1}.$$

Proof of Proposition 4. In this case the quality dimension 2 is always strictly positive, $q_2 > 0$ and therefore $p = \tilde{p}$ cannot be in equilibrium. The solution is given by p^* .

Proof of Corollary 1.

To prove statement (a), note that

$$\frac{\partial q_1^*}{\partial m} = -\frac{\phi_2 (1+\alpha) \left[(a_2 \phi_1 - ma_1) - \frac{m}{\phi_2} (a_1 \phi_2 - ma_2) \right]}{(\phi_1 \phi_2 - m^2)^2},
\frac{\partial q_2^*}{\partial m} = (1+\alpha) \frac{m (a_2 \phi_1 - ma_1) - \phi_1 (a_1 \phi_2 - ma_2)}{(\phi_1 \phi_2 - m^2)^2}.$$

Hence

$$\operatorname{sign}\left(\frac{\partial q_1^*}{\partial m}\right) = \operatorname{sign}\left[\frac{m}{\phi_2}(a_1\phi_2 - ma_2) - (a_2\phi_1 - ma_1)\right]$$

$$= \operatorname{sign}\left[2ma_1 - a_2\left(\frac{m^2}{\phi_2} + \phi_1\right)\right],$$

$$\operatorname{sign}\left(\frac{\partial q_2^*}{\partial m}\right) = \operatorname{sign}\left[m\left(a_2\phi_1 - ma_1\right) - \phi_1\left(a_1\phi_2 - ma_2\right)\right].$$

$$= \operatorname{sign}\left[2ma_2 - a_1\left(\frac{m^2}{\phi_1} + \phi_2\right)\right].$$

Hence, $\frac{\partial q_i^*}{\partial m} \geq 0$ depending on the parameter values.

(b)
$$\frac{\partial q_1^*}{\partial a_1} = \frac{(1+\alpha)\phi_2}{\phi_1\phi_2 - m^2} > 0, \quad \frac{\partial q_2^*}{\partial a_2} = \frac{\alpha\phi_1 + m^2/\phi_2}{\phi_1\phi_2 - m^2} > 0$$
(c) $\frac{\partial q_1^*}{\partial a_2} = -\frac{(1+\alpha)m}{\phi_1\phi_2 - m^2} < 0, \quad \frac{\partial q_2^*}{\partial a_1} = -\frac{m(1+\alpha)}{\phi_1\phi_2 - m^2} < 0$
(d) $\frac{\partial q_1^*}{\partial \phi_1} = -\frac{(1+\alpha)(a_1\phi_2 - ma_2)\phi_2}{(\phi_1\phi_2 - m^2)^2} < 0$

$$\frac{\partial q_2^*}{\partial \phi_2} = -\frac{m^2a_2}{\phi_2^2(\phi_1\phi_2 - m^2)} - \phi_1 \frac{\left(\alpha(a_2\phi_1 - ma_1) - \frac{m}{\phi_2}(a_1\phi_2 - ma_2)\right)}{(\phi_1\phi_2 - m^2)^2}$$

$$= -\frac{1}{(\phi_1\phi_2 - m^2)} \left(\frac{m^2a_2}{\phi_2^2} + \phi_1q_2^*\right) < 0.$$
(e) $\frac{\partial q_1^*}{\partial \phi_2} = \frac{m(1+\alpha)(a_2\phi_1 - ma_1)}{(\phi_1\phi_2 - m^2)^2} > 0$ since $(a_2\phi_1 - ma_1) > 0$ if $q_2^* > 0$.
$$\frac{\partial q_2^*}{\partial \phi_1} = (1+\alpha)m\frac{a_1\phi_2 - ma_2}{(\phi_1\phi_2 - m^2)^2} > 0$$
(f) $\frac{\partial q_1^*}{\partial \alpha} = \frac{a_1\phi_2 - ma_2}{\phi_1\phi_2 - m^2} > 0, \quad \frac{\partial q_2^*}{\partial \alpha} = \frac{a_2\phi_1 - ma_1}{\phi_1\phi_2 - m^2} > 0$

Proof of Corollary 2. Obviously, when the price in second best is zero (Proposition 1), the first best price is higher. To prove the first statement i), recall that $p_1^{fb} = B_{q_1}$, $\phi_{q_1q_2} > 0$, and $p^{sb} = B_{q_1} + B_{q_2} \left(-\phi_{q_1q_2}\right) / \left(-\alpha B_{q_2q_2} + \phi_{q_2q_2}\right)$. Note that since $B_{q_1q_1} = 0$, then B_{q_1} is a constant. Therefore, since $B_{q_2} \left(-\phi_{q_1q_2}\right) / \left(-\alpha B_{q_2q_2} + \phi_{q_2q_2}\right) < 0$ then $p_1^{fb} > p^{sb}$. The second statement ii) follows since $p^{sb} = \tilde{p} = B_{q_1}$ when the conditions in Proposition 3 holds. The proof of the last statement iii) follows along similar lines as the first statement but now $\phi_{q_1q_2} < 0$.

Proof of Lemma 1. Suppose $q_1 = q_2 = q$. Then $\phi_{q_1} = \phi_{q_2} = K$ (by symmetry of the cost function). Let $\Delta > 0$. Then (starting in a symmetric situation)

$$K_1 \equiv \phi_{q_1}(q + \Delta, q) \approx K + \phi \Delta > K + m\Delta \approx \phi_{q_2}(q + \Delta, q) \equiv K_2.$$

Now, let $q_1 = q + \Delta > q_2$ (starting in an asymmetric situation).

$$\phi_{q_1}(q_1 + \Delta, q_2) \approx K_1 + \phi \Delta > K_2 + m\Delta \approx \phi_{q_2}(q_1 + \Delta, q_2).$$

Since Δ can be chosen arbitrarily small the approximation should hold. Furthermore, symmetry ensures that $\phi_{q_1} \leq \phi_{q_2} \iff q_2 \geq q_1$.

Proof of Proposition 5. We have

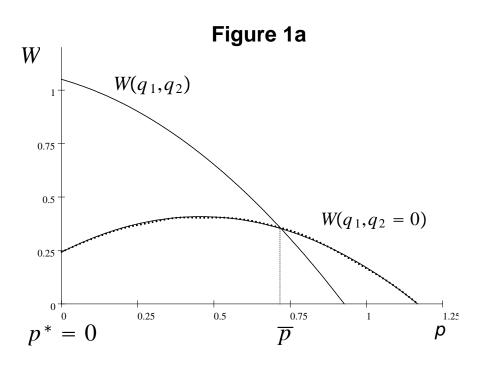
$$\begin{split} \phi_{q^{fb}}(q^{fb},q^{fb}) &= (\phi+m)\,q^{fb} > \phi q_1^{sb} + m q_2^{sb} = \phi_{q_1^{sb}}(q_1^{sb},q_2^{sb}) \\ &\frac{\phi}{m}q^{fb} + q^{fb} > \frac{\phi}{m}q_1^{sb} + q_2^{sb} \\ q^{fb} - \frac{\phi}{m}\left(q_1^{sb} - q^{fb}\right) > q_2^{sb}. \end{split}$$

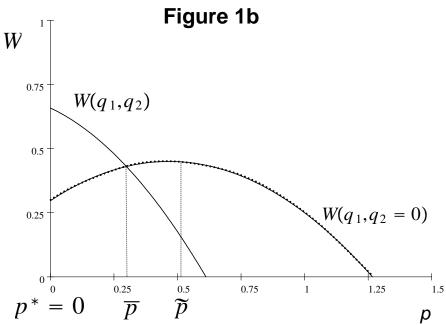
Hence, if $q^{fb} < q_1^{sb}$, then quality along the second dimension cannot be too high. The second statement follows since $q^{fb} - q_1^{sb} < 0$ and $0 < m < \phi$. Hence

$$-m\left(q^{fb}-q_2^{sb}\right)<\phi\left(q^{fb}-q_1^{sb}\right)<0$$

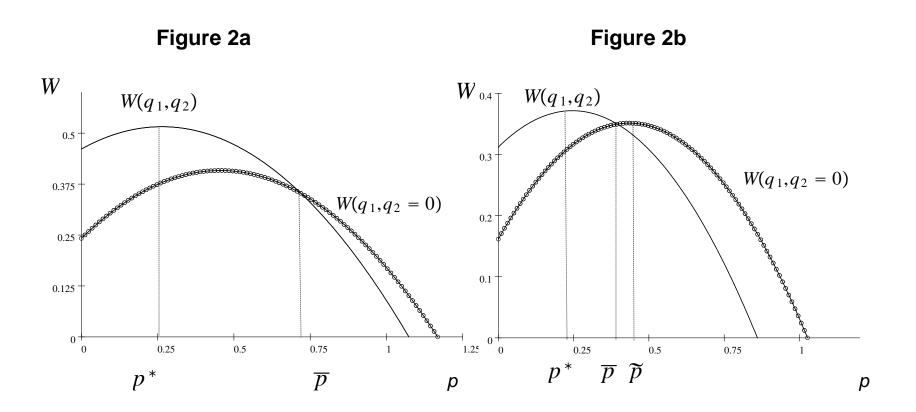
This last equation is fulfilled if and only if $-\left(q^{fb}-q_2^{sb}\right)<\left(q^{fb}-q_1^{sb}\right)$, or $q_1^{sb}+q_2^{sb}<2q^{fb}$.

Case 1. Incentive scheme breaks down



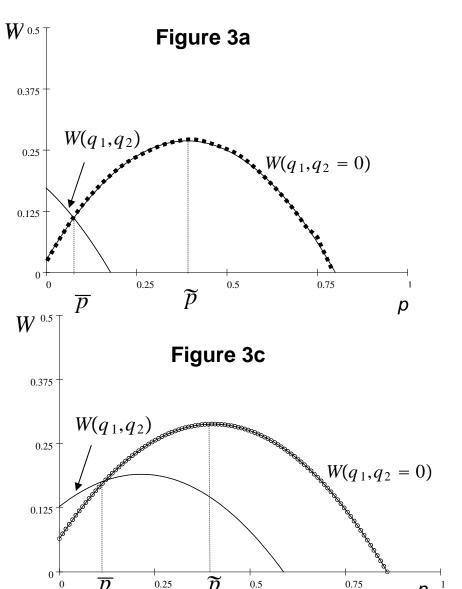


Case 2. Low-powered incentive scheme



Case 3: High-powered incentive scheme

36



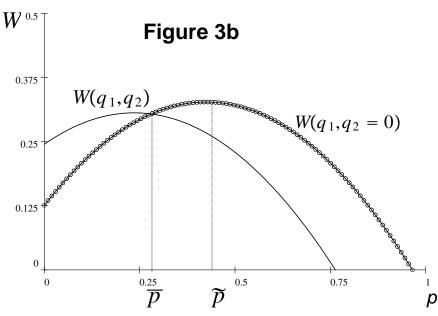
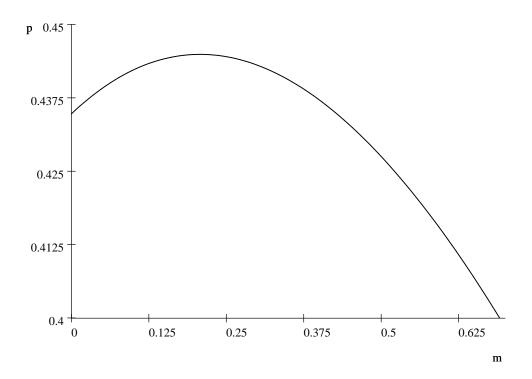


Figure 4. Optimal price may increase in m



Department of Economics University of Bergen Fosswinckels gate 6 N-5007 Bergen, Norway Phone: +47 55 58 92 00

Telefax: +47 55 58 92 10 http://www.svf.uib.no/econ