Int J Health Care Finance Econ (2006) 6:103–117
DOI 10.1007/s10754-006-7107-7

Provider choice of quality and surplus

Nolan Miller · Karen Eggleston · Richard Zeckhauser

Abstract We study the quality choices of institutional health-care providers, such as hospitals, assuming that the utility function of the key organizational decision maker includes both quality of care and financial surplus. We are primarily concerned with how changes in outside claims—particularly proportional outside claims—on the provider's financial surplus affect his choice of quality. We use the term "rate of surplus retention" to refer to the fraction of surplus remaining after deducting all such claims. Using the Arrow-Pratt coefficient of relative risk aversion as a measure of curvature of the provider's utility-from-money function, we show that increasing the surplus retention rate increases (decreases) quality if the provider's coefficient of relative risk aversion is greater than (less than) 1.

Keywords Quality · Surplus · Risk aversion · Surplus retention · Provider maximization · Nonprofit · For-profit

JEL Classification I1

1. Introduction

Health care providers—such as physicians, hospitals and HMOs—must trade off the quality of care they deliver against financial returns. We focus on how the structure of their preferences affects this tradeoff. We follow an extensive literature on the goals of providers, which usually identifies quality and returns as being primary objectives, though sometimes one or the other

N. Miller (✉) · R. Zeckhauser
Kennedy School of Government, Harvard University, Cambridge, MA 02138
e-mail: nolan_miller@harvard.edu

R. Zeckhauser
e-mail: richard_zeckhauser@harvard.edu

K. Eggleston
Department of Economics, Tufts University, Medford, MA 02155
e-mail: karen.eggleston@tufts.edu

is identified as a constraint (Newhouse, 1970; Pauly, 1987; Sloan, 2000; Lakdawalla and Philipson, 1998; Philipson and Lakdawalla, 2001). Following this literature, we consider the incentives of providers whose objective functions are the sum of *quality* and *utility from financial surplus.*[1]

Many parties may lay claim to a health care provider's financial surplus: workers, suppliers, the government, and so on.[2] We investigate the consequences for quality of changing the *fraction* of financial surplus retained by the key decision maker. If, for example, the provider secures better terms from his workers, suppliers, or grantors, or (at a for-profit provider) receives a larger bonus incentive or is subjected to a lower tax rate, the provider's rate of financial surplus retention increases.

When the provider's rate of surplus retention increases, how will his choice of quality change in response? Naïve intuition suggests that he might choose to provide more quality. After all, the provider values quality, and allowing him to retain more financial surplus means that he has greater resources available to finance quality provision. However, such a change does more than increase the provider's income; it also increases the opportunity cost or "price" of quality (in terms of foregone income). Thus, increasing the percentage of surplus retained differs from increasing the absolute level of surplus, just as obtaining a larger slice of a given pie differs from enlarging the pie itself. We show that whether such a change leads the provider to increase or decrease quality depends on both the nature of the change and the shape of the provider's utility function.

Our central question has direct relevance for policy. How will a provider—a nursing home, the administrator of a behavioral health carve-out, etc.—change quality in response to an increase in percentage fee? Will a nursing home chain that sells a significant equity stake, thus retaining less of any given surplus, have an incentive to reduce the quality of the personnel he hires and facilities he maintains? Will reducing tax rates on providers who pay taxes spur investment in quality or merely increase their profits?

We find that an increase in surplus retention produces two countervailing effects. First, it increases the decision maker's surplus, holding quality constant. In response, the decision maker spends some of the additional surplus to increase quality. This effect is akin to consumer theory's income effect. Second, an increase in surplus retention increases the rate at which the decision maker's surplus must be sacrificed in order to increase quality. This leads the decision maker to decrease quality, as in consumer theory's substitution effect.

Which effect dominates depends on the degree of curvature of the decision maker's utility-for-money (UFM) function. We show that if the UFM function is sufficiently curved (as measured by the Arrow-Pratt coefficient of relative risk aversion), then increasing the rate of surplus retention leads to an increase in quality, as expected. However, if the UFM function has modest curvature (specifically, if the coefficient of relative risk aversion is less than one), then increasing the rate of surplus retention leads the provider to decrease quality.

A number of factors affect the curvature of the decision maker's utility function. Drawing on the connection between curvature of the utility function and choice under uncertainty, a general phenomenon emerges: factors that tend to increase curvature (i.e., risk aversion)—such as breakeven concerns, low levels of financial assets, high levels of background risk, and liquidity constraints—tend to be associated with financially difficult times. Thus, we might expect that financially secure organizations would respond to a decrease in outside claims on

[1]Choné and Ma (2005) discuss the "imperfect agency" model of physician behavior and provide a list of other works employing a similar objective function.

[2]Changes in hospital financial surplus often arouse controversy; see for example Rowland, 2006.

surplus by decreasing quality (i.e., by taking the windfall as profit), while more financially strapped organizations would respond to the same increase in surplus retention by increasing quality. After all, these latter organizations are the ones whose financial constraints prevent them from providing as much quality as they would choose to provide in an unconstrained world. This intuition proves apt.

Probably the most closely related work is McGuire and Pauly (1991). They focus on supplier-induced demand and over-provision under fee-for-service payment; we focus on the provider's incentive to stint on quality, a pervasive concern under prospective payment, capitation, or other forms of "supply-side cost sharing" (Ellis and McGuire, 1990; Newhouse, 2002). In the McGuire-Pauly model, a provider's concern for his patients is reflected in a disutility from inducing excessive volume. By contrast, our provider's concern for patients takes the form of boosting quality despite a loss in surplus from doing so. Like McGuire and Pauly, we use a representative agent's utility function, but we are concerned with an organizational decision maker allocating resources to quality, rather than an individual physician's labor-leisure trade-off.

In the remainder of this paper, we first present the model and analysis. We then briefly discuss the implications for organizational decision making and how other types of efficiency changes present more straightforward policy implications. A brief discussion concludes.

2. The model and analysis

Consider a provider organization with a central decision maker who values both quality provided to patients and retained surplus (profit). The provider is paid at least partially prospectively, i.e., independent of actual expenditures. Examples are a hospital under a prospective payment system or a capitated managed care plan.

Let q denote quality provided to patients, with the understanding that $q = 0$ represents the contractible minimum level of quality that must be provided to patients. Let $u(y)$ be the decision maker's utility for money, where y denotes the provider's financial surplus net of any claims placed by other parties, which we call **retained surplus**.[3] We assume that $u(y)$ is strictly increasing, strictly concave, and twice differentiable. The provider's utility function is additively separable in quality and retained surplus, and takes the form:

$$V(q, y) = q + u(y), \qquad (1)$$

where quality is normalized so that utility is linear in q.[4] Thus, $u'(y)$ gives a cardinal measure of the decision maker's willingness to pay for an additional unit of quality when retained surplus is y. To keep clear the distinction between overall utility, $V(q, y)$, and utility-for-money, $u(y)$, we will continue to refer to $u(y)$ as the decision maker's utility-for-money (UFM) function.

Let $s(q)$ denote the gross per-patient surplus (before deducting any outside claims on surplus) when quality is q. We assume that per-patient surplus is positive when minimal quality is provided, $s(0) > 0$. Although in practical situations the provider may be rewarded with higher surplus for quality above the minimum (i.e., $s'(q) > 0$ for some q), it is straight-forward to

[3]We use the terms financial surplus and retained surplus instead of gross and net profit in order to emphasize the fact that the providers we consider are not necessarily profit maximizers.

[4]As long as additive separability is maintained, the assumption that utility $V(q, y)$ is linear in quality is without loss of generality, since the quality variable can always be rescaled to restore linearity.

show that an optimizing decision maker will choose q such that $s'(q) < 0$, and so, without loss of generality, we assume that $s'(q) < 0$ for all q. We also assume that surplus is concave in quality, $s''(q) < 0$.

We will be particularly interested in proportional claims on surplus. Our independent variable is $b \in (0, 1)$, the fraction of gross surplus that accrues to the provider. Thus, fraction $1 - b$ of the surplus is captured by other parties. The cleanest example is taxation (in the case of for-profit providers), but employee profit sharing, either formal or informal (e.g., expectations of increased wages when times are flush), can also be interpreted as such a claim. Retained surplus is therefore given by $y = bs(q)$ when quality is q.

The provider's maximization problem (*PMP*) is to choose quality and retained surplus in order to maximize utility, and is written as:

$$\begin{aligned} &\max_{q \geq 0, y \geq 0} \quad V(q, y) \qquad (PMP) \\ &s.t. \qquad y \leq bs(q). \end{aligned}$$

Since $V(q, y)$ is strictly increasing in q and y, the constraint in (*PMP*) will bind at the optimal solution. Further, since $V(q, y)$ is strictly quasiconcave in quality and the constraint set is strictly convex, (*PMP*) has a unique solution, which we denote (q^*, y^*).

Problem (*PMP*) resembles the standard optimization problem from consumer theory. However, in our case the budget set is nonlinear. The slope of the budget set, $bs'(q)$, embodies an increasing "price" of quality in terms of foregone surplus, given the assumption that $s''(q) < 0$.

Figure 1 illustrates the provider's optimal quality choice for a given rate of surplus retention, b.[5] Curve $y = bs(q)$, the constraint from (*PMP*), can be thought of as the production possibility frontier for transforming quality into retained surplus or vice-versa. The dashed curves represent the provider's utility isoquants. Solving (*PMP*) leads the provider to choose point (q^*, y^*), where the utility isoquant is tangent to the constraint, $y = bs(q)$.

Since the constraint in (*PMP*) will bind at any optimal solution, the provider's maximization problem is equivalent to an unconstrained problem in which the provider chooses q to maximize:

$$v(q) = q + u(bs(q)). \qquad (2)$$

Keeping in mind the non-negativity constraint on q,differentiating (2) yields the first-order condition for the provider's optimal quality choice:

$$bs'(q^*)u'(bs(q^*)) \leq -1. \qquad (3)$$

If the provider chooses to provide more than minimum quality, i.e., $q^* > 0$, then (3) must be satisfied with equality. Since $v(q)$ is concave in quality, the following condition is sufficient for an interior maximum:

$$bs'(0)u'(bs(0)) > -1. \qquad (4)$$

[5] As in Newhouse (1970), the provider chooses the level of the two arguments in the utility function to reach the highest possible indifference curve, facing a nonlinear production possibility frontier or budget set. Whereas in the Newhouse model the provider maximizes utility from quality and quantity subject to a break-even constraint, here the provider maximizes utility from quality and (per-patient) surplus, for a given level of surplus retention b (and is otherwise unconstrained).

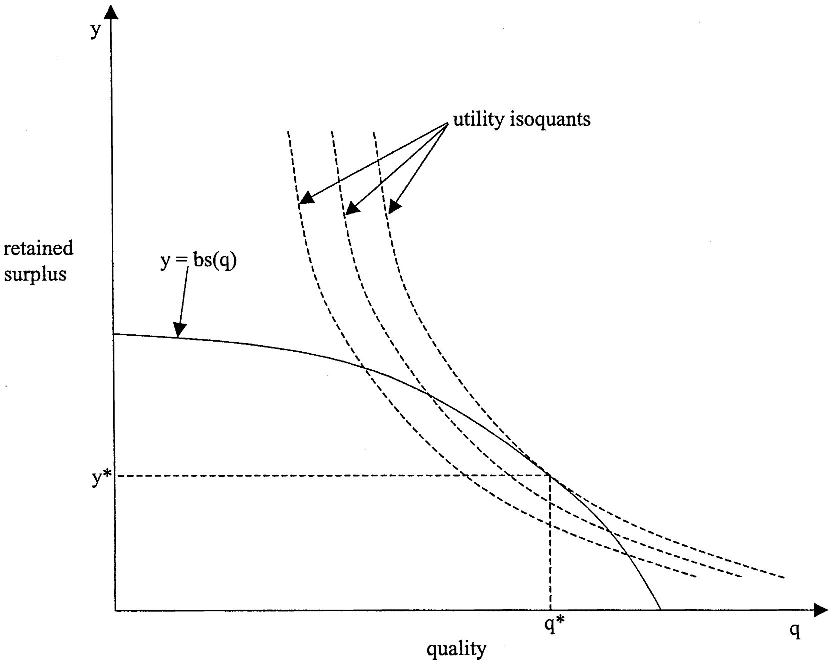

Fig. 1 The provider's problem

For the moment, we assume that (4) is satisfied and consider the properties of an interior solution. We return to the case where (4) is not satisfied later.

When (4) is satisfied, the provider's optimality condition, (3), can be rewritten as:

$$u'(bs(q^*)) = -\frac{1}{bs'(q^*)}. \qquad (5)$$

The left-hand side of (5) represents the marginal utility of increasing retained surplus. The right-hand side represents the magnitude of the loss in utility from the decrease in quality that accompanies the increase in retained surplus. To see this, let $h(y)$ be the quality provided when retained surplus is y, i.e., $h(bs(q)) = q$. Since utility v is linear in quality q, $h(y)$ is measured in units of utility. Differentiating, $\frac{dh}{dy}bs'(q) = 1$, or $\frac{dh}{dy} = \frac{1}{bs'(q)}$ when $q = h(y)$. Thus, the provider chooses a level of quality to equate the marginal utility of additional surplus with the marginal disutility of the quality reduction it entails.

Conditions (3) and (5) highlight the countervailing roles played by quality in the provider's objective function. When the provider optimizes, increasing quality increases the provider's utility but decreases the provider's retained surplus. At the optimum, the provider balances the marginal utility contributions of these two effects.

2.1. The effect of a change in the surplus retention rate

We now consider the effect of a change in the surplus retention rate, b, on the provider's quality choice. Such changes might arise from changes in the tax code, regulatory interventions (e.g., a cap on retained surplus), formation of a successful union, new executive compensation arrangements, or other factors.

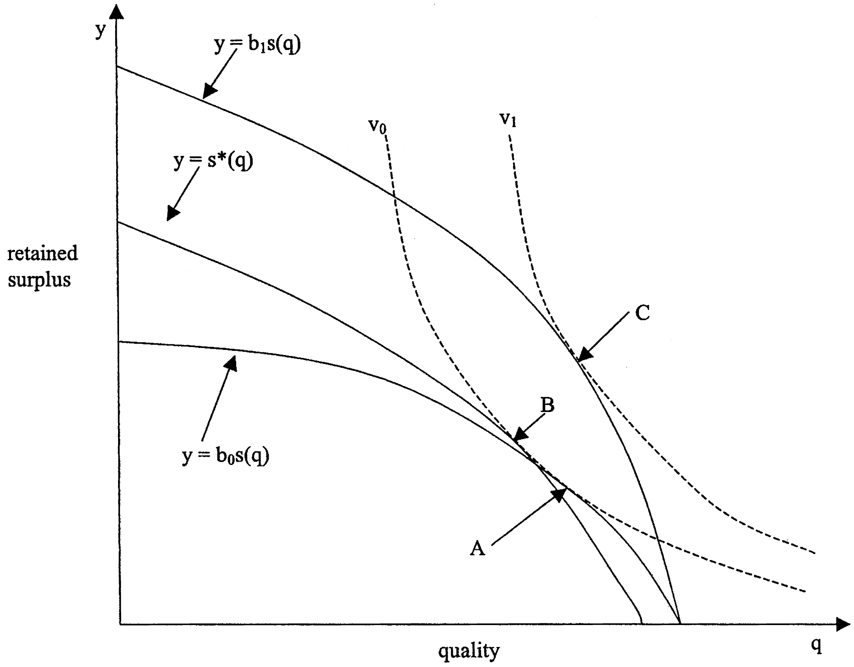

Fig. 2 The substitution and income effects

Figure 2 illustrates how an increase in b from b_0 to b_1 affects the provider's quality choice. The dashed curves labeled v_0 and v_1 represent the provider's utility isoquants when the surplus retention rates are b_0 and b_1, respectively.[6] Initially, the provider retains fraction b_0 of his surplus and chooses point A, where the utility isoquant is tangent to the constraint $y = b_0 s(q)$, as in Fig. 1 above. When the surplus retention rate increases from b_0 to b_1, the provider chooses point C, where his utility isoquant is tangent to the new constraint, $y = b_1 s(q)$.

The impact of an increase in b on the provider's quality choice can be decomposed into two effects, analogous to the income and substitution effects of standard consumer theory. An increase in b decreases the relative price of retained surplus (measured in terms of forgone quality). Holding fixed purchasing power, this gives the provider an incentive to choose higher surplus and lower quality. This is the substitution effect. At the same time, increasing b expands the budget set outward; for any q, it is now possible to retain more surplus. This "income effect" gives the provider an incentive to increase both surplus and quality (argument provided below).[7] The net effect is that increasing b will always induce the provider to choose more surplus, but the overall effect on his quality choice is ambiguous.

[6]Since utility is quasilinear in quality, each utility isoquant is a "horizontal translation" of any other. Therefore, holding retained surplus constant, the slopes of the utility isoquants do not change as quality increases.

[7]The analogy to the income effect of neoclassical consumer theory is not exact: the budget set here is nonlinear and the increase in b induces an upward shift in the constraint (rather than a radial expansion). Nevertheless, our income effect is the natural extension of the neoclassical concept to our more general environment, and so we invoke the name.

Figure 2 illustrates the substitution and income effects in this model. First, to identify the substitution effect, we isolate the impact of the change in b on the slope of the constraint. The curve labeled $s^*(q)$ is derived by shifting $y = b_1 s(q)$ downward until it is tangent to utility isoquant v_0. When the provider maximizes utility given this budget set, he chooses point B and earns utility v_0. Since the provider's maximized utility is the same as when the surplus retention rate is b_0, rotating the constraint from $y = b_0 s(q)$ to $y = s^*(q)$ represents a change in b in which the changes in the provider's behavior can be attributed only to the change in the slope of the budget set. This is equivalent to the construction of a compensated price change in consumer theory.[8] Since $s^*(q)$ is steeper than $b_0 s(q)$ through point A, the provider reacts to the increase in b, which is analogous to an increase in the relative price of quality, by substituting toward surplus and away from quality. Thus quality is necessarily lower at point B than at point A—the substitution effect (on quality) is always negative.

We next turn to the income effect. Increasing b expands the budget set outward, increasing the real consumption opportunities available to the provider. Taking the substitution effect as fixed, this shifts the constraint upward from $s^*(q)$ to $b_1 s(q)$. In this model, upward shifts in the constraint lead the provider to increase both quality and retained surplus, and so the provider responds to this increase in purchasing power by increasing quality.[9]

In this model, the income and substitution effects oppose each other. Whether a provider increases or decreases quality when her surplus retention rate rises depends on which effect—substitution or income—is stronger. Interestingly, it turns out that we can tell which effect predominates by looking at well known properties of the UFM function, $u(y)$. In particular, the degree of concavity of that function determines whether an increase in b leads the provider to increase or decrease quality. A natural measure of concavity is the Arrow-Pratt coefficient of relative risk aversion. Even though uncertainty plays no role in our analysis, this metric is well-suited to our purposes.[10] Since $bs(q)$ is the provider's retained surplus, the decision maker's Arrow-Pratt coefficient of relative risk aversion for UFM function u () is given by:

$$\rho = -bs(q)\frac{u''(bs(q))}{u'(bs)(q)}.$$

Proposition 1. *Suppose $q^* > 0$. If $\rho > 1$, then an increase in the surplus retention rate, b, leads the provider to increase quality. If $\rho = 1$, then the provider's quality choice is unaffected by an increase in the surplus retention rate. If $\rho < 1$, then an increase in the surplus retention rate leads the provider to decrease quality.*

Proof: Denoting the dependence of q^* on b by $q(b)$, substituting $q(b)$ into (3), and totally differentiating with respect to b yields the following expression for $\frac{dq}{db}$, the response of q^* to a change in surplus retention:

$$\frac{dq}{db} = \frac{-s'(q(b))u'(bs(q(b))) - s'(q(b))bs(q(b))u''(bs(q(b)))}{(bs''(q(b))u'(bs(q(b))) + (bs'(q(b)))^2 u''(bs(q(b))))}. \qquad (6)$$

[8]Mathematically, $s^*(q)$ is the new constraint shifted downward by the change in retained surplus when quality is held constant at the level of point A. That is, let $A = (q_A, y_A)$. Then, $s^*(q) = b_1 s(q) - s(q_A)(b_1 - b_0)$.

[9]We prove that upward shifts in the constraint lead the provider to increase quality and surplus (a property akin to being "normal goods") in the Appendix.

[10]The Arrow-Pratt coefficient is often used as a measure of curvature in riskless environments. For example, see the discussion of intertemporal consumption problems in Gollier (2001).

The denominator of (6) is negative by concavity of $s()$ and $u()$. Substituting in the definition of ρ, the numerator has the same sign as $s'(q(b))(\rho - 1)$, which has the opposite sign of $\rho - 1$. Hence $\frac{dq}{db}$ is positive if $p > 1$ and negative if $p < 1$. □

Why should the degree of concavity be the critical variable? The key insight is that the importance of the income effect relative to the substitution effect increases as concavity of the UFM function increases. Thus, when $u()$ is not very concave, the substitution effect dominates, and increasing the rate of surplus retention leads the provider to decrease quality. On the other hand, when concavity is extreme, the income effect dominates, and increasing b leads the provider to increase quality.

To get a better feel for the intuition, consider the bookend case of a provider with a linear UFM function. This provider's utility isoquants are linear. Therefore, since $s^*(q)$ is a vertical translation of $b_1 s(q)$, point C must lie directly above point B, i.e., there is no income effect. However, the substitution effect persists, and therefore an increase in b leads the provider to decrease quality. When the provider's UFM function is concave, but only slightly so (i.e., $\rho < 1$), the income effect is positive, but the substitution effect continues to dominate, and increasing b decreases quality.

At the other extreme, when the provider's UFM function is very concave (i.e., the provider's utility isoquant through point A bends sharply), a small increase in surplus significantly reduces the marginal utility of surplus. In this case, the substitution effect is very small (points A and B are close together), whereas the income effect is large. Thus, when the provider's UFM function is very concave, the income effect dominates and the increase in b leads to an increase in quality. By extension, the income effect continues to dominate whenever the provider's UFM function is sufficiently concave.

To complete the analysis of the solution to PMP, we briefly comment on corner solutions where minimal quality, i.e., $q^* = 0$, is optimal. When (4) fails to hold (i.e., when $bs'(0)u'(bs(0)) \leq -1$), the decision maker chooses $q^* = 0$ and $y^* = bs(0)$. This "corner solution" situation could arise because either the marginal cost of providing additional quality ($bs'(0)$) or the marginal disutility of decreasing surplus ($u'(bs(0))$) is large in magnitude.[11] This suggests that the decision maker may choose minimum quality when b is either large (i.e., near 1) or small (i.e., near 0). When b is near 1, the marginal cost of providing additional quality, $bs'(0)$, is large. On the other hand, $u'(bs(0))$ is large when surplus is small, and therefore when b is small.

2.2. Implications for decision making

Proposition 1 establishes that a provider may either increase or decrease quality in response to an increase in surplus retention, a result that has potentially important implications. For example, naïve intuition might suggest that a decrease in the tax rate imposed on for-profit providers would increase quality because the provider values quality and decreasing the tax rate increases the provider's ability to purchase quality. However, if the decision-maker's

[11]Note that the "corner solution" of $q^* = 0$ would never arise if $s'(0) > 0$, since (4) always holds when $s'(0) > 0$. Intuitively, if the provider is rewarded with higher surplus for quality above the minimum (over some initial range of q), then the provider would always choose $q^* > 0$. Our assumption that surplus is concave in quality, therefore, can encompass not only capitation and other forms of supply-side cost sharing with $s'() < 0$, as assumed, but also fee-for-service payment, as long as the reward for additional quality is eventually outweighed by his cost.

marginal utility from surplus declines sufficiently slowly (i.e., $u()$ is sufficiently close to linear), this need not be the case. Similarly, a decrease in the claims of outside parties on the provider's surplus may also lead to a decline in quality.

Our primary interest in ρ is as a measure of the curvature of the decision maker's UFM function, not as a measure of his risk attitude. Nevertheless, estimates of ρ from the literature on choice under uncertainty can be helpful in understanding the kind of behavior we should expect to observe in the world and, in particular, whether we should expect to see actual decision makers respond to an increase in b by decreasing quality. A number of studies have attempted to estimate the relative risk aversion of actual decision makers. Gollier (2001) suggests ρ between 1 and 4 as a plausible range, and this is supported by Barsky et al. (1997) who estimate ρ somewhere just above 4. However, estimates such as these tend to be very sensitive to the size of the hypothetical lotteries used to elicit preferences (Kandel and Stambaugh, 1991). This suggests that, to the extent that these studies accurately estimate the curvature of individual decision makers' utility-for-money functions, we might expect individual decision makers such as physicians to react to an increase in the surplus retention rate by increasing quality.

From the theory of choice under uncertainty we expect firms and institutional decision makers to be less risk averse—and therefore to have less curved UFM functions—than individual decision makers. The fact that an institution's capacity can be scaled much more readily than an individual's provides an additional reason why curvature should be lower for institutions. This suggests that institutional decision makers are more likely to respond to an increase in the rate of surplus retention by decreasing quality than are individuals. That is, in the case of institutional decision makers, the counterintuitive effects of seemingly beneficial policy changes (e.g., lowering tax rates, decreasing workers' claims on surplus, etc.) might be a common occurrence.

A natural question to ask is: what conditions will affect the concavity of decision makers' utility-from-surplus functions and therefore the manner in which they respond to variation in the rate of surplus retention? In the remainder of this section, we briefly discuss four factors that affect the curvature of the utility-from-surplus function: the provider's level of financial resources, exposure to background risk, breakeven or "target surplus" concerns, and (in the context of a more dynamic model) liquidity constraints. The analysis of each case supports the same basic idea: financially secure providers are more likely to have utility-from-surplus functions that are less concave, and therefore they are more likely to respond to an increase in the rate of surplus retention by decreasing quality.

Financial Resources. Individual decision makers typically exhibit decreasing relative risk aversion. That is, the rate at which their marginal utility of surplus decreases (as measured by ρ) is higher at low wealth levels than at high wealth levels.[12] Thus, for example, we might expect poor physicians to respond to an increase in b by increasing quality, but wealthy physicians to respond to the same change by decreasing quality.

Background risk. Background risk also affects the effective curvature of decision makers' UFM functions. Under reasonable circumstances, decision makers' preferences are believed to exhibit "risk vulnerability," which means that adding zero-mean background risk has the effect of making the decision maker behave as if his UFM function were more concave (i.e.,

[12] See Gollier (2001).

increasing ρ).[13] Thus we might expect providers operating hi highly risky environments to react to an increase in the surplus retention rate by increasing quality, while those facing less background risk would react to the same change by decreasing quality.

Breakeven Concerns. Frequently, decision makers are faced with the need to achieve a certain critical level of surplus. If the decision maker is unable to achieve this goal, he will face the possibility of a large penalty. For example, a firm that cannot break even over the long run will go out of business; a nonprofit reaping red ink may lose contributions; a decision maker who has large losses may risk dismissal. Even if such undesirable events occur only probabilistically, they may be important enough in expectation to significantly affect the decision maker's behavior.[14] Because of this, we would expect the decision maker's marginal utility from surplus to be very high near the breakeven point and to fall rapidly as wealth increases beyond the breakeven level. In other words, ρ will be high near the surplus target or breakeven level and much lower once surplus moves significantly beyond this level. The result is that, once again, we expect decision makers whose current asset levels place them near the breakeven point to respond to an increase in b by increasing quality, while those who safely satisfy their requirements to respond by decreasing quality.

A simple model of a decision maker facing a breakeven constraint, which can be thought of as the limiting case of a decision maker with breakeven or target income concerns, illustrates these effects. The predictions of our model are consistent with the behavior that arises in this limiting case (as well as in the more general case where concavity increases greatly in the region of the breakeven constraint).

Let $s(q) = r - c(q)$, where r is a fixed reimbursement, and $c(q)$ is the provider's increasing and convex cost of providing quality. The provider has linear utility for money, with overall utility function:

$$v(q) = q + b(r - c(q)),$$

and he faces breakeven constraint $b(r - c(q)) \geq f$, where $f > 0$ represents the target income, the surplus required to break even.[15]

If r is large and f is small, the breakeven constraint does not bind, and the optimal quality choice satisfies:

$$bc'(q^*) = 1.$$

Totally differentiating with respect to b yields:

$$\frac{dq}{db} = \frac{-c'(q^*)}{bc''(q^*)} < 0,$$

[13]The concept of risk vulnerability is due to Gollier and Pratt (1996). A related concept is proper risk aversion (Pratt and Zeckhauser, 1987). See Gollier (2001) for a concise discussion of both concepts.

[14]Frank (1990) also discusses "the possibility that significant bankruptcy costs may lead an otherwise risk neutral firm to avoid risk" (p. 925). In his model, firms behave more cautiously at low levels of aggregate demand, leading to self-sustaining recessions. See Gollier (2001) for a discussion of the relationship between breakeven/liquidity constraints and risk aversion.

[15]The results continue to hold if the provider has concave utility for money.

which agrees with Proposition 1. For a risk-neutral provider in relatively good times (i.e., when the breakeven constraint does not bind), an increase in b induces a decrease in q.

Now consider the case where the breakeven constraint does bind. In this case, q^* satisfies:

$$b(r - c(q^*)) = f,$$

and

$$\frac{dq}{db} = \frac{(r - c(q^*))}{bc'(q^*)} > 0.$$

Hence, when times are difficult enough that the breakeven constraint binds, the provider responds to an increase in b by increasing q. This stands to reason, since the constraint only binds when the provider would like to increase quality but cannot do so and still break even. An increase in b relaxes the constraint and allows the provider to choose higher quality.

Liquidity Constraints. In a dynamic model in which a firm faces a liquidity constraint, similar effects may arise. Even if the constraint does not bind currently, the possibility that the constraint might bind in the future may lead the decision maker to decrease current quality in order to increase reserves in the future. An increase in surplus retention decreases the likelihood of needing these reserves, and therefore makes the decision maker more willing to supply current quality. On the other hand, if there is little threat of the liquidity constraint binding in the future, the decision maker will respond to an increase in surplus retention by decreasing quality, since in this case the substitution effect, rather than the income effect, dominates.

2.3. Quality responses to other changes

The behavior described in Proposition 1 is specific to changes in the surplus retention *rate*, and its complex nature arises because in such situations the income and substitution effects push in opposite directions. Other important changes in the provider's environment give rise to unambiguous comparative statics. In this section we consider two such changes, an increase in productive efficiency (i.e., a downward shift in the firm's marginal cost curve) and a "ratcheting up" of performance incentives (i.e., an upward shift in the marginal revenue of quality). To illustrate, consider the case where $s(q) = ar(q) - kc(q)$. For this case, $r(q)$ is a non-negative, strictly increasing, and strictly concave function representing the firm's per-patient revenue as a function of quality, and $c(q)$ is the strictly positive, strictly increasing, and strictly convex cost of producing quality. An increase in $a > 0$ represents an increase in the firm's reward for increasing quality (e.g., performance incentives), and an decrease in $k > 0$ represents a gain in productive efficiency. The provider's utility function is:[16]

$$v(q) = q + u(ar(q) - kc(q)).$$

Proposition 2. *If a and k are such that the provider chooses positive quality, then:*

[16]In this representation, an increase in the surplus retention rate, b, corresponds to a simultaneous increase in a and k.

(i) an increase in productive efficiency (i.e., a decrease in k) leads the provider to increase quality, and
(ii) an increase in performance incentives (i.e., an increase in a) leads the provider to increase quality.[17]

Proof: First, consider an increase in k. Differentiating $v(q)$ with respect to q yields the optimality condition (for an interior solution):

$$u'(s^*)(ar'(q^*) - kc'(q^*)) = -1, \qquad (7)$$

where $s^* = ar(q^*) - kc(q^*)$. The second-order condition is:

$$u'(s^*)(ar''(q^*) - kc'(q^*)) + (ar'(q^*) - kc'(q^*))^2 u''(s^*) < 0.$$

Recognizing the dependence of q^* on k and totally differentiating (7) with respect to k, yields:

$$\frac{dq^*}{dk} = \frac{c'(q^*)u'(s^*) + (ar'(q^*) - kc'(q^*))c(q^*)u''(s^*)}{u'(s^*)(ar''(q^*) - kc'(q^*)) + (ar'(q^*) - kc'(q^*))^2 u''(s^*)} < 0$$

The numerator of dq^*/dk is positive since $c'(q)$ and $u'(s^*)$ are both positive and $(ar'(q^*) - kc'(q^*))$ and $u''(s^*)$ are both negative. The denominator is negative by the second-order condition. Part (i) of the proposition follows from noting that an increase in productive efficiency corresponds to a decrease in k.

Next, consider an increase in a. Taking q^* in (7) as a function of a and totally differentiating with respect to a yields:

$$\frac{dq^*}{da} = \frac{-r'(q^*)u'(s^*) - r(q^*)(ar'(q^*) - kc'(q^*))u''(s^*)}{u'(s^*)(ar''(q^*) - kc'(q^*)) + (ar'(q^*) - kc'(q^*))^2 u''(s^*)} > 0.$$

The numerator is negative since $r'(q^*)$ and $u'(s^*)$ are positive, $r(q^*)$ is positive, and $(ar'(q^*) - kc'(q^*))$ and $u''(s^*)$ are negative. The denominator is negative by the second-order condition. Hence dq^*/da is positive. □

The comparative statics of an increase in productive efficiency or pay-for-performance incentives are unambiguous because in these cases the income and substitution effects push in the same direction.[18] In fact, in these cases both are positive. For example, increasing a lowers the effective price of increasing quality (since there is a greater revenue increase to offset the increased cost of producing additional quality), which gives the provider an incentive to increase quality. At the same time, increasing a also increases the level of surplus associated with any level of quality, thereby shifting out the provider's budget constraint. This increase in real income also induces the provider to produce additional quality. An increase in productive efficiency unambiguously leads to an increase in quality because the income and substitution

[17] If a and k are such that the provider chooses $q^* = 0$, then the provider may respond to an increase in a or a decrease in k by continuing to provide minimal quality.

[18] Although we state the results for a particular type of shift in the provider's revenue and cost functions (i.e., increases respectively in a and k), the results in Proposition 2 hold in any model in which increasing productive efficiency decreases both cost and marginal cost and increasing performance incentives increases both revenue and marginal revenue. Proposition 2 generalizes immediately to such cases.

effects are both positive. That is, a decrease in k makes the provider wealthier and reduces the amount of surplus that must be forfeited in order to provide additional quality. Both forces push the decision maker toward increasing quality.

It is also straightforward to derive the quality implications of other changes in the environment. For example, an increase in per-patient revenue r increases quality (Ma, 1994). That is, if $v = q + u(r - c(q))$, then $\frac{dq^*}{dr} > 0$. In this case, increasing the quality-independent payment to the provider has only an income effect, and this income effect induces the provider to increase quality. Finally, better provider "agency" for patients also unambiguously leads to higher quality. It is easy to show that if $v = \alpha q + u(bs(q))$, then $\frac{dq^*}{d\alpha} > 0$. Since this type of exogenous change does not affect the level of financial surplus associated with any level of quality, it only exerts a substitution effect. Since increasing αincreases the providers relative preference for quality, an increase in αinduces an increase in quality. Thus, professional norms that promote behaving in the best interest of patients can curb the less savory actions of profit-seeking provider organizations or health plan managers, such as stinting on quality for unprofitable clients or services (Fuchs, 1996; McGuire, 2000; Newhouse, 2002).

3. Discussion

This paper considers how the fraction of financial surplus retained by a provider affects the quality of the services the provider offers. This relationship is more complicated than it might at first appear. We have shown that providers with low curvature in their utility-for-money functions, as might be expected for those in good financial condition, might respond to a decrease in outside claims on financial surplus by decreasing quality, not increasing it. Such a response arises because increasing the rate of surplus retention effectively transforms the provider into a more efficient surplus producer: less quality must be sacrificed to achieve the same increase in surplus. If the effect of this change in the relative prices of quality and surplus is strong enough, it may overcome the fact that the positive wealth effect associated with the increase in surplus retention induces the provider to choose higher quality. If so, the net effect will be a decrease in quality.

While our results arise in a somewhat stylized model, they do suggest that the effects of policy interventions aimed at changing the rate of surplus retention have effects on quality that are difficult to predict, and sometimes go in the opposite direction than would have been naïvely imagined. Thus, lowering taxes on for-profit providers may have the adverse effect of decreasing quality.

Conversely, raising the wages of hospital employees in response to the pressures of a flush year, though diminishing the hospital's reserves, may actually bolster quality. In both instances, the curvature of the utility function for surplus complicates predictions.

Appendix: The income effect is positive

In this Appendix, we provide a brief argument for why the income effect of an increase in the surplus retention rate (identified in the discussion of Fig. 2) must be positive.[19] To see why the income effect must lead the provider to increase quality, let $B = (q_B, y_B)$ and $C = (q_C, y_C)$, and consider Fig. 3. Since $s^*(q)$ is a downward translation of $b_1 s(q)$, $s^{*'}(q_B) = b_1 s'(q_B)$,

[19]This result can also be shown algebraically by writing $b_1 s(q)$ as $s^*(q) + k$ and implicitly differentiating the solution to the provider's problem when the constraint is $y = s^*(q) + k$ with respect to k.

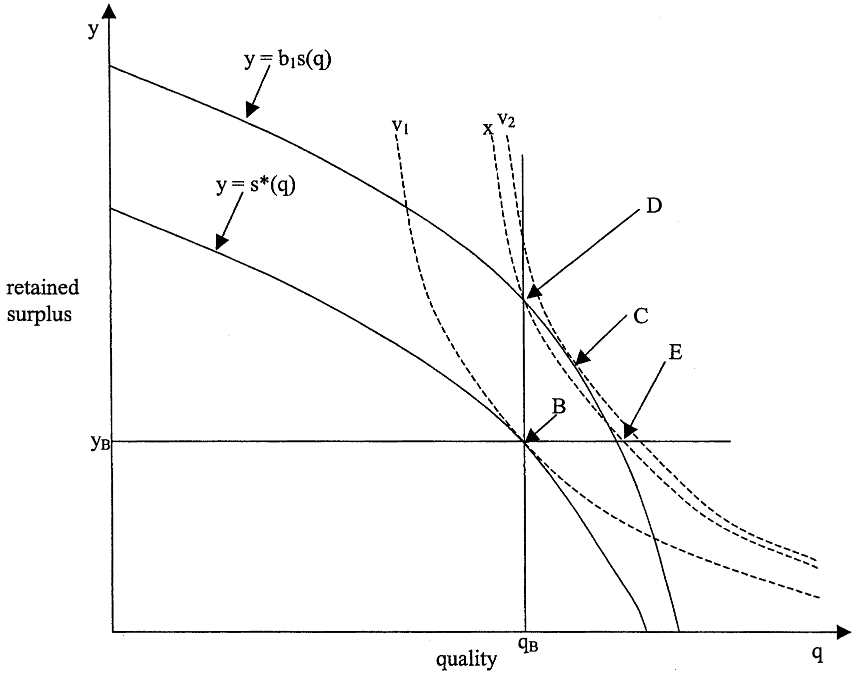

Fig. 3 The income effect is uniformly positive

and the constraints have the same slope at points B and D. On the other hand, quasilinear utility implies that utility isoquants are horizontal translations of each other. Therefore the iso-utility curve through point D (labeled x) has the same slope at point E as constraint $y = b_1 s(q)$ has at point D, and therefore, by convexity of the isoquants, point D cannot be an optimum. Further, we know (also by convexity) that isoquant x is steeper than constraint $b_1 s(q)$ at point D, and therefore that the solution to the provider's optimization problem when the constraint is $y = b_1 s(q)$ must lie further to the right along the constraint than point D. Hence the income effect must increase quality.

Acknowledgment The authors thank Tom McGuire, Joe Newhouse, and two anonymous referees for helpful comments and the AHRQ for financial support under grant number 1 R03 HS013362-01A1.

References

Barsky, R., F. Juster, M. Kimball, and M. Shapiro. (1997). "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study." *Quarterly Journal of Economics* 112(2), 537–579.

Choné, P. and C-T. A. Ma. (2005). "Asymmetric Information from Physician Agency: Optimal Payment and Health Care Quality." Mimeo, Boston University.

Ellis, R., and T. G. McGuire. (1990). "Optimal Payment Systems for Health Services." *Journal of Health Economics* 9, 375–396.

Frank, Jeff. (1990). "Monopolistic Competition, Risk Aversion, and Equilibrium Recessions." *Quarterly Journal of Economics* 105(4), 921–938.

Fuchs, V. R. (1996). "Economics, Values, and Health Care Reform." *American Economic Review* 86(1), 1–24.

Gollier, C. (2001). *The Economics of Risk and Time.* Cambridge, MA: MIT Press.

Gollier, C., and J. Pratt. (1996). "Risk Vulnerability and the Tempering Effect of Background Risk." *Econometrica* 64(5) 1109–1123.
Kandel, S. and R. Stambaugh. (1991). "Asset Returns and Intertemporal Preferences." *Journal of Monetary Economics* 27(1), 39–71.
Lakdawalla, D. and T. Philipson. (1998). "Nonprofit Production and Competition." National Bureau of Economic Research Working Paper 6377.
Ma, C. A. (1994). "Health Care Payment Systems: Cost and Quality Incentives." *Journal of Economics & Management Strategy* 3(1), 93–112.
McGuire, T. G. (2000). "Physician Agency." Chapter 9 in A. J. Culyer and J. P. Newhouse, (eds.), *Handbook of Health Economics,* Vol. 1A (Amsterdam: Elsevier North-Holland), pp. 461–536.
McGuire, T. G. and M. V. Pauly. (1991). "Physician Response to Fee Changes with Multiple Payers." *Journal of Health Economics* 10, 385–410.
Newhouse, J. P. (1970). "Toward a Theory of Nonprofit Institutions." *American Economic Review* 60, 64–74.
Newhouse, J. P. (2002). *Pricing the Priceless: A Health Care Conundrum.* Cambridge, MA: MIT Press.
Pauly, M. V. (1987). "Nonprofit Firms in Medical Markets." *AEA Papers and Proceedings* 77(2), 257–262.
Philipson, T. and D. Lakdawalla. (2001). "Medical Care Output and Productivity in the Nonprofit Sector." In *Medical Care Output and Productivity*. David M. Cutler and Ernst R. Berndt (eds.) NBER Studies in Income and Wealth, vol. 62. Chicago and London: University of Chicago Press, pp. 119–137.
Pratt, J. and R. Zeckhauser. (1987). "Proper Risk Aversion." *Econometrica* 55(1), 143–154.
Rowland, C. (2006). "Medicaid Cash Lifts Hospital Profit." *The Boston Globe*, pg. E1.
Sloan, F. A. (2000). "Not-For-Profit Ownership and Hospital Behavior." Chapter 21 in *Handbook of Health Economics*, Culver and Newhouse (eds.), pp. 1141–1174.