NBER WORKING PAPER SERIES

REGULATION AND THE PROVISION OF QUALITY TO HETEROGENOUS CONSUMERS: THE CASE OF PROSPECTIVE PRICING OF MEDICAL SERVICES

Robin Allen

Paul Gertler

Working Paper No. 2269

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 1987

This research was conducted while Robin Allen was on leave as a Victor H. Kramer Fellow at the University of Chicago Law School, Chicago, Illinois. This paper has benefited greatly from the comments of David Dranove, Kathleen Eyre, Mark Satterthwaite, Ken Thorpe, and the participants in seminars at Harvard, Northwestern, and the University of Chicago. Of course, the usual liability disclaimer applies. The research reported here is part of the NBER's research program in Health Economics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #2269
May 1987

Regulation and the Provision of Quality to Heterogenous Consumers: The Case of Prospective Pricing of Medical Serivces

ABSTRACT

This paper analyzes the welfare implications of fixed price regulation in a model in which consumers are heterogenous and a firm can endogenously quality discriminate. The motivation for this analysis is the current move of third party payors (governmental and private insurors) toward prospective pricing of medical services. Our major result is that prospective pricing causes a distributional welfare loss. Specifically, in our model, prospective pricing induces a profit maximizing medical care provider to simultaneously provide a smaller than socially optimal level of quality to more severely ill patients and, surprisingly, a greater than socially optimal amount of quality to less severely ill patients. Further, the distributional welfare loss does not disappear when ethically motivated deviation from profit maximization is allowed.

The inefficient distribution of quality occurs because prospective payment regulation fixes the price across patients with different severities of illness but allows providers to quality discriminate. More complicated DRG pricing rules do not appear to be able to completely avoid this problem. Alternatively, vertical integration of third party payors into the direct provision of medical care is shown to be able to bypass the problem completely. This implies that the recent proliferation of vertically integrated health care organizations such health maintenance organizations, preferred provider organizations, and managed care plans by self-insuring employers are welfare improving.

Robin Allen
US Department of Justice
Anti-Trust Division
Judiciary Center Building
555 4th Street, NW
Washington, DC 20001

Paul Gertler
Department of Health Policy
and Management
Harvard School of Public Health
677 Huntington Avenue
Boston, MA 02115

I. INTRODUCTION

This paper analyzes the welfare implications of fixed price regulation in a model in which consumers are heterogenous and a firm can endogenously quality discriminate. The motivation for this analysis is the current rapid move to prospective pricing of medical services. In 1987 the Medicare program, which is the largest purchaser of health care services in the United States, began paying hospitals a completely prospectively determined fixed fee per patient.[1] In addition, many state Medicaid programs and private insurors have instituted comparable payment systems, and many of the others are actively considering implementing one. Also, several state Medicaid programs, such as New York's, are using a similar method to pay for nursing home care, and extension of the system to physician services is actively being considered by Medicare and many state Medicaid programs.

The move to prospective pricing is in response to the enormous growth in health care (and in particular, hospital) expenditures as a result of retrospective reimbursement of "reasonable costs" and the accompanying high costs of administration.[2] Cost based reimbursement provided little incentive for health care providers to efficiently produce medical care as costs could usually be passed on to the third party payor. In contrast, prospective payment provides incentive for the efficient (minimum cost) production of medical services by allowing providers to keep the difference between the price and

1 See Vladeck (1984) for a complete description of Medicare's prospective payment program.

2 See Schweiker (1982) for a discussion of the rationale used by U.S. Department of Health and Human Services to push for prospective pricing.

variable costs.[3] Further, prospective pricing substantially reduces administrative costs by removing the necessity of annual audits of each hospital's financial statement, and the monitoring of the reasonableness of each patient's bill.

Medicare's prospective payment system pays hospitals a fixed fee per patient depending upon which of the approximately 470 "diagnosis related groups" (DRGs) the patient is placed. Once placed in a DRG, the fee is independent of the services (e.g. number of hospitalized days, number of tests, etc.) provided the patient. The services provided each patient are the quality characteristics of hospital care. Hospital services (quality) are typically tailored to the medical needs of the individual patient (Harris, 1977). Empirical work in Berki (1984), Horn et. al. (1983,1985), and Stern and Epstein (1985) show great variation in severity of illness within DRG categories, implying substantial patient heterogeneity and potential for quality discrimination. In summary, the DRG payment structure fixes the price across patient severity of illness types within DRGs, while allowing medical care providers to quality discriminate (i.e. provide different levels of quality to different patients).

The concern voiced most often about prospective pricing is that it may have an adverse effect on health (i.e. patients will leave hospitals "quicker and sicker"). In models with heterogenous consumers and exogenous quality, Dranove (1986) and Newhouse (1983) argue that prospective payment is an incen-

3 Shleifer (1985) also shows that setting prices based on the average cost of a pool of comparable firms is incentive to choose the cost minimizing level of capital stock.

tive for providers not to admit (transfer) severely ill patients that are likely to be very costly to treat. In a model with endogenous quality and a fixed number of patients (i.e. exogenous quantity), Ellis and McGuire (1986) argue that prospective payment improves efficiency incentives but may also induce providers to reduce quality.

All of these models place strong restrictions on the demand function. Because providers cannot adjust quality in Dranove's and Newhouse's models, the choice set of providers is limited to refusing to treat (i.e. transfer) patients. In essence, quality becomes a binary variable; treating the patient in a high quality hospital or transferring to a low quality hospital. Even though Ellis and McGuire consider endogenous quality, by making quantity ex ogenous they do not allow quality choices to affect the number of patients a provider attracts and thus eliminate any role for a market for patients.

This paper extends previous analysis by combining consumer heterogeneity, endogenous quality, and endogenous quantity into one model. We show that a DRG pricing structure induces a distributional welfare loss. Specifically, in our model, a profit maximizing provider simultaneously supplies a <u>smaller</u> than socially optimal amount of quality to more severely ill patients, and surprisingly, a <u>greater</u> than socially optimal quality to less severely ill patients. The ratio of patients who are over-supplied quality to those who are under-supplied increases with the price.

In addition, these results do not disappear when we allow ethically motivated deviation from profit maximization. We show that such deviation induces providers to reduce the degree of under-provision of quality to more severely ill patients without reducing the over-provision of quality to less severely ill patients. Indeed, even if providers are perfect agents for patients, they still over-provide quality to the less severely ill patients while providing the optimal amount of quality to the more severely ill. In most cases, ethically motivated providers are able to supply the ethical amount of quality to the more severely ill patients only by cross-subsidizing their care with the profits from the treatment of the less severely ill.

Finally, we discuss several possible solutions to this distributional welfare problem. The most promising one appears to be vertical integration of medical care providers with third party payors and insurors. This implies that recent innovations in the health care market such as health maintenance organizations, preferred provider organizations, and managed care plans by self-insuring employers are welfare improving.

II. BEHAVIORAL ASSUMPTIONS AND NOTATION

Individuals seek medical care in response to an illness or accident and for preventive (investment) purposes. The quality of medical care is its expected efficacy (marginal product) in terms of health. The greater the flow of medical services provided to an individual, the higher is quality. The consumption of medical services is assumed to improve an individual's stock of health, but at a diminishing rate (i.e. a positive and diminishing marginal productivity of medical care services).

The productivity of medical services depends on the complexity and severity of the medical problem. The more severe and complex the problem, the greater the marginal productivity of medical services. Let θ measure the complexity and severity of illness.

A provider's demand from the group of individuals with illness type θ is

$$X = X(P,Q;\theta), \tag{1}$$

where X is the number of patients, P is the price charged each patient, and Q is the quality (flow of medical services) provided each patient. The demand function is assumed to be decreasing in price ($X_P \leq 0$), increasing in quality ($X_Q \geq 0$), and due to a diminishing marginal productivity of medical care, concave in quality ($X_{QQ} \leq 0$). We also assume that individuals are less sensitive to price changes for higher quality medical care ($X_{PQ} \geq 0$).

Since the value of medical care rises with severity of illness, the proportion of potential patients within a θ group that purchase care rises with the severity and complexity of illness, holding price and quality constant. On the other hand, the density of potential patients in an θ group declines rapidly with severity and complexity of illness. The decline in density is likely to offset the increase in the proportion of potential patients seeking care, so that demand is likely to be approximately constant over θ types, holding price and quality constant. Therefore, the demand functions are assumed to be constant in θ ($X_\theta = 0$). In addition, more severely ill individuals are assumed to have higher marginal values of quality, so that the demand function is more quality elastic ($X_{Q\theta} \geq 0$) and less price elastic ($X_{P\theta} \geq 0$) for more severe illness types.

Finally, the cost of supplying quality level Q to X patients is

$$C = cXQ, \tag{2}$$

where c is the marginal cost of a unit of quality. The cost function, as specified in (2), exhibits constant returns to scale. The constant returns to scale assumption is consistent with hospital cost function estimates reported in Friedman and Pauly (1981). This assumption is extremely useful in that it implies that the profit and welfare maximization problems are separable in θ, and therefore can be considered separately for each θ type. This makes the problem amenable to standard calculus and bypasses the necessity of solving the more complicated optimal control problem.

III. THE FIRST-BEST

The social optimum is found by choosing price and quality to maximize the sum of consumer surplus and profits. Price and quality are allowed to vary by illness type. Since the cost function exhibits constant returns to scale, the social welfare function is separable by illness type. The social welfare function for illness type θ is

$$W = \int_P^\infty X(v,Q;\theta)dv + PX(P,Q;\theta) - cX(P,Q;\theta)Q, \tag{3}$$

The first-order conditions for each θ are

$$W_P = (P - cQ)X_P = 0 \tag{4}$$

$$W_Q = \int_P^\infty X_Q dv + (P - cQ)X_Q - cX = 0, \tag{5}$$

which reduce to the more familiar conditions,

$$P = cQ \tag{6}$$

$$\left(\int_P^\infty X_Q dv\right)/X = c. \tag{7}$$

Condition (6) requires price to be equal to the marginal cost of an additional patient, and condition (7) requires the marginal contribution of quality to average consumer surplus to equal the unit marginal cost of quality.

One implication of this model is that the welfare maximizing quality provided each patient monotonically increases for more severely ill patient types. Formally, by applying Cramer's rule to the first-order conditions in (4) and (5)

$$\frac{dQ^*}{d\theta} = \frac{-\int_P^\infty X_{Q\theta} dv}{\int_P^\infty X_{QQ} dv - 2cX_Q - c^2 X_P} \geq 0, \tag{8}$$

where Q^* is the welfare maximizing level of quality.[4] The denominator of (8) is the determinant of the hessian of the welfare maximization problem, divided by X_P. Since the determinant of the hessian is positive at the maximum and X_P

---

4 Equation (8) is derived by applying Cramer's rule to (4) and (5) to get

$$dQ^*/d\theta = -\left[W_{PP}W_{Q\theta} - W_{P\theta}W_{QP}\right] \div \left[W_{PP}W_{QQ} - W_{PQ}W_{QP}\right].$$

From (4), and recalling that (4) requires $P = cQ$, $W_{P\theta} = 0$ and $W_{PP} = X_P \leq 0$. From (5), and recalling that $P = cQ$ and $X_\theta = 0$, $W_{Q\theta} = \int_P^\infty X_{Q\theta}\, dv$, $W_{QQ} = \int_P^\infty X_{QQ}$ dv, and $W_{QP} = W_{PQ} = -cX_P \leq 0$. Substitution of these expressions into $dQ^*/d\theta$ and distributing out a X_P in both the denominator and the numerator yields (8).

is negative, the denominator of (8) is negative. The numerator is also negative as result of $X_{Q\theta} \geq 0$, implying that (8) is positive.

This result follows from patients' marginal value of quality (and hence marginal consumers' surplus) increasing with severity of illness. Individuals with greater severities of illness are willing to pay more for higher quality. As a result, the welfare maximizing amount of quality is larger for more severely ill patients.

IV. DRG PRICE REGULATION

Under DRG price regulation, one fixed price is set for a range of illness types, and by law must accept the DRG price as payment in full. The provider is assumed to take the price as given, but is allowed to supply a different level of quality to individuals of different illness types. No matter how the price is chosen, the provider chooses quality so as to maximize profits. Therefore, the following results are valid for the second-best policy, where the DRG price is chosen to maximize welfare subject to the provider choosing quality to maximize profits.

Let $\hat{P}$ be the price for the DRG comprised of illness types $\underline{\theta}$ through $\overline{\theta}$. Since the cost function exhibits constant returns to scale, the profit function is separable by illness type, and for illness type θ is

$$\Pi = \hat{P}X(\hat{P},Q;\theta) - cX(\hat{P},Q;\theta)Q, \qquad (9)$$

The profit maximizing quality for patients of type θ is given by the condition

$$\Pi_Q = (\hat{P} - cQ)X_Q - cX = 0, \tag{10}$$

which requires the marginal revenue from quality to equal marginal cost.

As in the first-best case, the profit maximizing level of quality monotonically increases with severity of illness. Formally, by applying Cramer's rule to (10) and a little manipulation

$$\frac{d\hat{Q}}{d\theta} = \frac{-(\hat{P} - cQ)X_{Q\theta}}{(\hat{P} - cQ)X_{QQ} - 2cX_Q} \geq 0, \tag{11}$$

where $\hat{Q}$ is the profit maximizing level of quality. The denominator is Π_{QQ} and, by the second order conditions, must be negative at the maximum. The numerator is also negative as $X_{Q\theta} \geq 0$, implying that (11) is positive.

This result is obtained because consumers' marginal value of quality, and therefore the marginal revenue from quality is greater for higher more severely ill patient types. As a result, a profit maximizing provider supplies more quality to patients in the higher θ markets than in the smaller θ markets.

The profit maximizing level of quality for each θ type also increases with price. By applying Cramer's rule to (10) and a little manipulation

$$\frac{d\hat{Q}}{d\hat{P}} = \frac{-[(\hat{P} - cQ)X_{QP} - cX_P]}{(\hat{P} - cQ)X_{QQ} - 2cX_Q} \geq 0. \tag{12}$$

The denominator of (12) is the same as the denominator of (11), and is therefore negative. The numerator is also negative as a result of $X_{QP} \geq 0$ and $X_P \leq 0$, implying that (12) is positive. This result occurs because a higher price implies a higher marginal revenue from quality, and therefore profit maximizing providers supply more quality to each θ market as the DRG price rises.

V. DEVIATION FROM THE FIRST-BEST FOR A GIVEN SEVERITY TYPE

A profit maximizing provider facing DRG price regulation may supply patients of a given illness group a level of quality higher or lower than the first-best depending on the value of the DRG price. Further, by paying a premium, the regulator can induce a provider to supply the first-best quality to a particular θ group, but not to all θ groups. Another way to state this result is that for a particular price, there exists a θ group to which the profit maximizing provider supplies the first-best quality and the associated first-best price is less than the DRG price.

These results are easily demonstrated in figure 1 where the first order welfare maximizing and first-order profit maximizing conditions are pictured for a given θ group. The W_Q=0 and W_P=0 lines represent all the combinations of price and quality that satisfy the welfare maximizing conditions (4) and (5) for a given θ group. These lines are upward sloping with W_Q=0 being steeper than W_P=0.[5] Their intersection gives the welfare maximizing price and

---

5 By the implicit function theorem the slopes of W_P=0 and W_Q=0 are $dP/dQ = - W_{PQ}/W_{PP}$ and $dP/dQ = - W_{QQ}/W_{QP}$, respectively. The second order conditions require $W_{PP} \leq 0$ and $W_{QQ} \leq 0$. From (4) and (5) and recalling that (4) requires $P = cQ$, $W_{PQ} = W_{PQ} = -cX_P \geq 0$. Therefore, W_P=0 and W_Q=0 are both upward sloping, and the second order condition, $W_{PP}W_{QQ} - W_{PQ}W_{QP} \geq 0$, implies that W_Q=0 is steeper than W_P=0.

quality, $[P^*(\theta), Q^*(\theta)]$, for individuals with severity of illness type θ.

The Π_Q=0 line in figure 1 represents all the combinations of price and quality that satisfy the profit maximizing condition (11) for a given θ group. It is upward sloping and everywhere to the left of the W_Q=0 line.[6] Being everywhere to the left of W_Q=0 follows from Π_Q=0 being satisfied at a lower quality for each price.[7] Given the DRG price, the profit maximizing quality $\hat{Q}(\theta)$ supplied individuals with severity type θ is read off the Π_Q=0 line.

The relationship between the profit maximizing and welfare maximizing levels of quality depends on the value of the DRG price. If the DRG price equals $\hat{P}$ in figure 1, then the profit maximizing quality level equals the welfare maximizing quality level. Notice that a premium of $\hat{P} - P^*$ must be paid in order to induce the profit maximizing provider to supply the welfare maximizing level of quality. If the DRG price is larger than $\hat{P}$ (say $\hat{P}^1$ in figure 1), then the profit maximizing level of quality is larger than the welfare maximizing level. If the DRG price is set below $\hat{P}$ (say $\hat{P}^0$ in figure 1), then the profit maximizing level of quality is lower the welfare maximizing level.

As shown in figure 1, the members of illness group θ receive the welfare maximizing level of quality from a profit maximizing provider if the DRG price

---

6 By the implicit function theorem the slope of Π_Q=0 is $dP/dQ = -\Pi_{PQ}/\Pi_{QQ}$. The second order conditions require $\Pi_{QQ} \leq 0$, and from (10) $\Pi_{QP} = (P - cQ)X_{QP} - cX_P \geq 0$, both of which imply that Π_Q=0 is upward sloping.

7 This easily demonstrated by comparing the profit maximizing $\Pi_Q = 0$ condition in (10) with the welfare maximizing condition $W_Q = 0$ in (5) and noting that (5) can be expressed as $W_Q = \int_P^\infty X_Q dv + \Pi_Q = 0$. The expression $\int_P^\infty X_Q dv$ is positive, implying that (5) is satisfied at a higher Q than is (10) for the same values of P and θ.

is set at $\hat{P}$. It does not necessarily follow that other illness types will also receive the welfare maximizing quality when the DRG price is set at $\hat{P}$. Specifically, the $W_Q=0$, $W_P=0$, and $\Pi_Q=0$ lines are located in different positions for different θ types so that $\hat{Q}(\theta) \neq Q^*(\theta)$ for other θ types. This point is explored in detail in the next section.

VI. DEVIATION FROM THE FIRST BEST BY SEVERITY TYPE

Under DRG price regulation, profit maximizing providers supply a level of quality lower than the first-best to more severely ill patients, and a level of quality higher than the first-best to less severely ill patients.[8] This result is presented in figure 2a where the first best quality function $Q^*(\theta)$ and the profit maximizing quality function $\hat{Q}(\theta)$ are pictured. The functions are pictured for one DRG which spans the range $[\underline{\theta},\overline{\theta}]$. Both functions have already been shown to be monotonically increasing in θ. The $Q^*(\theta)$ function is steeper than the $\hat{Q}(\theta)$ function, and intersects $\hat{Q}(\theta)$ at $\tilde{\theta}$. A profit maximizing provider supplies patients with severity of illness less than $\tilde{\theta}$ with a greater than socially optimal level of quality, and supplies patients with severity of illness greater than $\tilde{\theta}$ with a less than socially optimal amount of quality.[9] The difference between $\hat{Q}(\theta)$ and $Q^*(\theta)$ is the amount of quality that is over-supplied (or under-supplied) to patients of illness type θ. Note that the difference is decreasing in θ.

---

8 Proof of this result is proved in the appendix

9 There is no guarantee that the $Q^*(\theta)$ line will intersect the $\hat{Q}(\theta)$ line in the $[\underline{\theta},\overline{\theta}]$ range as is pictured in figure 2a. This depends on the value of $\hat{P}$. As we argue later, the way in which Medicare chooses $\hat{P}$ makes it likely that the two functions will cross in the relevant range.

The under- and over-provision of quality is a result of DRG regulation restricting the price to be the same across severity types. Under-provision occurs because, unlike the first-best where more severely ill patients pay more for greater quality, providers cannot by law receive higher payments for higher quality. The relatively low regulated price induces providers to supply a lower than first-best quality to more severely ill types, when in fact, these patients would prefer a higher quality at a higher price. Over-provision occurs because providers find the DRG payment very lucrative for less severely ill patients. The relatively high regulated price induces providers to supply a higher than first-best level of quality to attract more of these patients, even though these patients would prefer a lower quality at a lower price.

A related result is that proportion of patients being over-supplied quality increases with price (i.e. the θ at which $Q^*(\theta)$ equals $\hat{Q}(\theta)$ is increasing in $\hat{P}$). In (12) we demonstrated that the profit maximizing level of quality for each θ is increasing in price. Therefore, an increase in $\hat{P}$, to say P', shifts the $\hat{Q}(\theta)$ line upwards to $Q'(\theta)$ as is pictured in figure 2b. At the new price $Q^*(\theta)$ equals $Q'(\theta)$ at $\theta' > \tilde{\theta}$. Consequently, patients types between $\tilde{\theta}$ and θ' switch from being under-supplied quality to being over-supplied.

Further, there exists a price above which all patients within the DRG (i.e. $\theta \epsilon [\underline{\theta},\overline{\theta}]$) are over-supplied quality, and another below which all patients are under-supplied quality. It is probable that current DRG prices are within the range in which the $Q^*(\theta)$ line crosses the $\hat{Q}(\theta)$ line in the range $[\underline{\theta},\overline{\theta}]$, so that some patients are always under-supplied quality and other are always over-supplied quality. The reason for this is that the DRG prices are set based on the average costs of a large pool of comparable hospitals.[10] The prices reflect the (inefficient) costs associated with the quality provided to the average θ under the old cost based payment system. Since the prices are greater than the cost of efficiently providing quality to the average θ, the prices are high enough so that some less severely ill patients are being over-supplied quality. In addition, some patients will be under-supplied quality as there are always a few very severely ill patients who are outliers in the illness distribution within each DRG.

VII. ETHICAL CONSIDERATIONS AND CROSS-SUBSIDIZATION

Even though proprietary hospital chains are rapidly expanding their market shares, it is not clear that all medical care providers can be modeled strictly as profit maximizers. Indeed Pauly (1980) and Ellis and McGuire (1986) stress the role of the physician as a patient's agent, and the conflict between ethical considerations for the patients' welfare and their own profit (income) maximization. In this section we show that ethically motivated deviation from profit maximization causes medical care providers to use profits from the care of less severely ill patients to cross-subsidize the care of more severely ill patients. This cross-subsidization moves the quality of care closer to the first-best for more severely ill patients but does not affect the over-provision of quality to less severely ill patients.

---

10 The pool is constructed to be large enough so that no hospital's actions can significantly influence the pool average costs.

Ethical considerations are introduced by means of a minimum quality constraint. The provider is ethically bound to provide patients at least some percentage of the first-best quality, where this percentage is assumed to be constant over all θ types. This is represented in figure 3a by the $Q^e(\theta)$ line, which is parallel to $Q^*(\theta)$, the first-best quality function. Patients with severity of illness types greater then θ^e in figure 3a receive an increase in quality from the $\hat{Q}(\theta)$ line to the $Q^e(\theta)$ line. Patients with severity less than θ^e are unaffected by the constraint as the profit maximizing quality for them is higher the ethical minimum.

If the minimum quality constraint introduces negative profits for the patients with severity greater than θ^e, then their care must be cross-subsidized with profits from the patients with severity less than θ^e. This notion of cross-subsidization is consistent with the industry complaint that the actual costs incurred by hospitals in caring for very severely ill patients are substantially greater than the DRG price.

If the provider is a perfect agent for patients, then the minimum quality constraint $Q^e(\theta)$ is equal to the first-best quality function $Q^*(\theta)$, as is pictured in figure 3b. In this case, patients with severity of illness greater than $\tilde{\theta}$ receive the first-best quality, and patients with severity less than $\tilde{\theta}$ receive the profit maximizing level of quality given by the $\hat{Q}(\theta)$ line. Hence, even when providers are perfect agents, they still over-supply quality to the less severely ill.

VIII. DRG REFORM

As long as there is severity of illness heterogeneity and endogenous quality discrimination within DRG categories, it will be difficult to find a pricing rule that provides the socially optimal quality and promotes efficiency. One obvious means of reducing the within DRG severity of illness variation is to create more DRG categories. A prominent proposal, suggested by Horn et. al. (1985), would effectively increase in the number of DRGs five fold. The creation of more categories would seriously aggravate the already costly problem of monitoring hospitals to prevent the spurious reclassification of patients into more lucrative DRGs (DRG creep). Alternatively, Dranove (1986) and Shleifer (1985) suggest letting the price depend on exogenous sources of cost variation across providers such as severity of illness. Of course, the validity of this approach requires quality to be exogenous.

Another approach, suggested in Ellis and McGuire (1986), is to let the DRG price be a linear function of the cost of the individual patient's care. As long as the slope parameter is less than one, this structure maintains the efficiency incentives of the fixed rate prospective payment while allowing the price to vary by patient type. Although not perfect, this type of pricing rule may induce providers to supply a quality schedule closer to the first-best. The proposal is problematic in that the informational requirements necessary to implement it are huge as it doubles the number of DRG parameters that must be chosen. The choice of optimal parameters is even more difficult than in the usual principle-agent setting, as ethical deviation from profit maximization and moral hazard incentives introduced by the existence of medical care insurance must be incorporated into the problem. In addition, an incentive scheme which is a linear function of costs puts Medicare back in the costly business of auditing hospital financial statements and monitoring patient bills, and there is still the issue of preventing DRG creep.

In sum, these methods of "fine tuning" the DRG payment scheme will probably improve economic efficiency somewhat, but at the expense of reintroducing substantial administrative burden to regulatory agencies. This seems to defeat one of the major reasons for switching from retrospective cost based reimbursement to DRG style prospective payment. Therefore, rather than devoting further work to analyzing the relative merits of these fine tuning proposals, we consider in the next section how recent invovations in the industry have the potential for the efficient distribution of quality with litle adminstrative cost.

IX. VERTICAL INTEGRATION

There are two fundenmental incentive problems for third party payors (governmental and private insurors)in the health care industry: (1) the optimal design of medical care insurance for consumers which is comnplicated by the familiar problems of moral hazard and advrese selection and (2) the optimal design of provider payment schemes which is the focus of this paper. On the provider side, an inefficient distribution of quality arises from the inability of third party payors to use price incentives to induce medical care providers to implement a first-best quality schedule. The difficulty arises due to the conflicting incentives facing third party payors and medical care providers. Insurors cannot <u>ex ante</u> offer prospective subscribers the first-best quality schedule because they only imperfectly control providers' quality schedules, $\hat{Q}(\theta)$, through ex post incentives embedded in the methods they use to pay providers. Therefore, competition in the insurance market could at best induce insurors to offer the second-best quality schedule, $\hat{Q}(\theta)$.

A natural reaction to a competitive insurance market is for insurors to vertically integrate into the direct provision of health care. Vertically integrated insuror-providers who are able to offer the first-best quality schedule $Q^*(\theta)$, have a competitive advantage. Indeed, vertically integrated insuror-providers such as health maintenance organizations, preferred provider organizations, and managed care plans by self-insuring employers are proliferating throughout the health care industry. Unlike the unintegrated market, competition in a market of vertically integrated insuror-providers could induce the first-best quality schedule. Indeed, the market for health insurance is becoming substantially more competitive with the emergence of health maintenance organizations, preferred provider organizations and managed care plans by self-insuring employers.

Vertical integration of third party payors and providers removes the necessity of imperfect regulatory pricing of providers, and therefore, in combination with a competitive market for insurance mitigates the type of distributional welfare losses discussed here. The familiar problems in designing optimal insurance are still present, but are no longer complicated by the imperfect control of providers by third party payors.

X. CONCLUSIONS

The purpose of this paper is point out some of the difficulties and complexities that patient severity of illness heterogeneity and endogenous quality discrimination add to the medical care provider pricing problem. Our major result is that DRG style prospective pricing causes a distributional welfare loss. Specifically, it induces medical care providers to simultaneously provide a smaller than socially optimal level of quality to more severely ill patients and, surprisingly, a greater than socially optimal amount of quality to less severely patients. This inefficient distribution of quality occurs because DRG price regulation fixes the price across patients with different severities of illness. More complicated DRG pricing rules do not appear to be able to completely avoid this problem.

Vertical integration of third party payors into the direct provision of medical care may be a solution to this problem. In an unintegrated market the first-best quality schedule in unavailable to insurors. They can at best offer potential subscribers the second-best quality schedule that results from the imperfect control of providers through incentives embedded in payment mechanisms. With vertical integration the first-best quality schedule is available to insurors, and therefore, there is no longer a need to design second-best provider payment schemes. Competition among vertically integrated insuror-providers can induce the provision of the first-best quality to patients in every illness group. Thus, the recent proliferation of vertically integrated health care organizations such health maintenance organizations, preferred provider and managed care plans by self insuring employers are welfare improving.

APPENDIX

This appendix provides proof of the proposition that a profit maximizing provider supplies a lower than socially optimal level of quality to more severely ill patients and a greater than socially optimal level of quality to more severely ill patients. This proposition is illustrated in figure 2a. Since both $Q^*(\theta)$ and $\hat{Q}(\theta)$ have already been shown to be monotonically increasing in θ (see sections III and IV respectively), this proposition is proven by demonstrating that the slope of $Q^*(\theta)$ is everywhere greater than the slope of $\hat{Q}(\theta)$ (i.e. showing that (8) is greater than (11)). In sections V and VI, we established that, for a reasonable range of prices, there exists at least one $\theta \epsilon [\underline{\theta}, \overline{\theta}]$ where $Q^*(\theta)$ equals $\hat{Q}(\theta)$. Let any severity type at which the two functions are equal be denoted by $\tilde{\theta}$. If the slope of $Q^*(\theta)$ is greater than the slope of $\hat{Q}(\theta)$ at each $\tilde{\theta}$, then $Q^*(\theta)$ crosses $\hat{Q}(\theta)$ only once and is everywhere steeper than $\hat{Q}(\theta)$.

Demonstrating that (8) is greater than (11) at any $\tilde{\theta}$ is made easier by rearranging (8). By (6), cQ can substituted into (8) for P. Then, assuming that third order derivatives of the demand function are negligible, we can write (8) as

$$\frac{dQ^*}{d\theta} = \frac{-(\alpha - cQ)X_{Q\theta}}{(\alpha - cQ)X_{QQ} - 2cX_Q - c^2X_P} \leq 0. \tag{13}$$

where α is defined by $X(\alpha,Q;\theta)=0$ for the Q and θ at which (13) is evaluated.

Now we show that (13) is larger than (11) at $\tilde{\theta}$. We have already established that at $\tilde{\theta}$, Q^* equals $\hat{Q}$ and P^* is less than $\hat{P}$. Since Q and θ are identical and third order derivatives of the demand function are assumed to be negligible, the values of cQ, $X_{Q\theta}$, and X_{QQ} are the same in (11) and (13). The additional result that P^* is less than $\hat{P}$ along with our demand function assumption $X_{QP} \geq 0$, implies that $2cX_Q$ is smaller in (13) than in (11). Since $2cX_Q$ is in the denominators of (11) and (13), this difference acts so as to make (13) larger than (11).

The denominators of (11) and (13) are both negative. The first two terms in the denominators of both expressions are identical expect that $\hat{P}$ in (11) is replaced with $\alpha > \hat{P}$ in (13). The third term in the denominator of (13) has no counterpart in (11). This term, $(- c^2X_P)$, is positive, and its addition to the denominator of (13), the rest of which is negative, acts to increase (13) relative to (11).

The remaining difference between (11) and (13) is that $\hat{P}$ in (11) is replaced with $\alpha > \hat{P}$ in (13). If $\hat{P}$ equalled α, then, based on our results so far, (13) would be bigger than (11). Therefore, since (11) can be shown to be increasing in $\hat{P}$ and (13) can be shown to be increasing in α, α larger than $\hat{P}$ implies that (13) is larger than (11).

REFERENCES

Berki, S., M. Ashcraft, and W. Newbrander (1984) "Length of Stay Variations within ICDA-8 Diagnosis Related Groups." Medical Care, Vol.22, pp.126-142.

Dranove, D. (1987) "Rate Setting by Diagnosis Related Groups." Mimeo, Graduate School of Business, University of Chicago.

Friedman, B. and M. Pauly (1981) "Cost Functions for a Service Firm with Variable Quantity and Stochastic Demand." Review of Economics and Statistics, Vol. 53, pp.620-624.

Harris, J. (1977) "The Internal Organization of Hospitals." The Bell Journal of Economics pp. 467-482.

Horn, S. (1983) "Measuring Severity of Illness; Comparisons Across Institutions." American Journal of Public Health Vol. 72, pp. 25-31.

Horn, S. and P. Sharkey (1984) "Measuring Severity of Illness to Predict Patient Resource Use within DRGs." Inquiry, Vol. 20, pp. 314-321.

Horn, S. et. al. (1985) "Inter-hospital Differences in Severity of Illness." New England Journal of Medicine, Vol. 313, pp. 325-333.

Ellis, R. and T. McGuire (1986) "Provider Behavior Under Prospective Payment: Cost Sharing and Supply." Journal of Health Economics", Vol.5, pp.129-151.

Newhouse, J.(1983) "Two Prospective Difficulties with Prospective Payment, or, its Better to be a Resident than a Patient with a Complex Problem." Journal of Health Economics, Vol. 2, pp. 269-174.

Pauly, M. (1980) Doctors and Their Workshops (University of Chicago Press, Chicago, Illinois)

Schweiker, R. (1982) "Report to Congress: Hospital Prospective Payment for Medicare." U.S. Department of Health and Human Services, Washington D.C.

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

Vladeck, B. (1984) "Medicare Payments to Diagnosis Related Groups." Annals of Internal Medicine, Vol. 100, no. 4, pp.576-590.

FIGURE 1

FIRST-BEST AND DRG PRICE REGULATED EQUILIBRIA FOR A GIVEN SEVERITY OF ILLNESS (Θ)

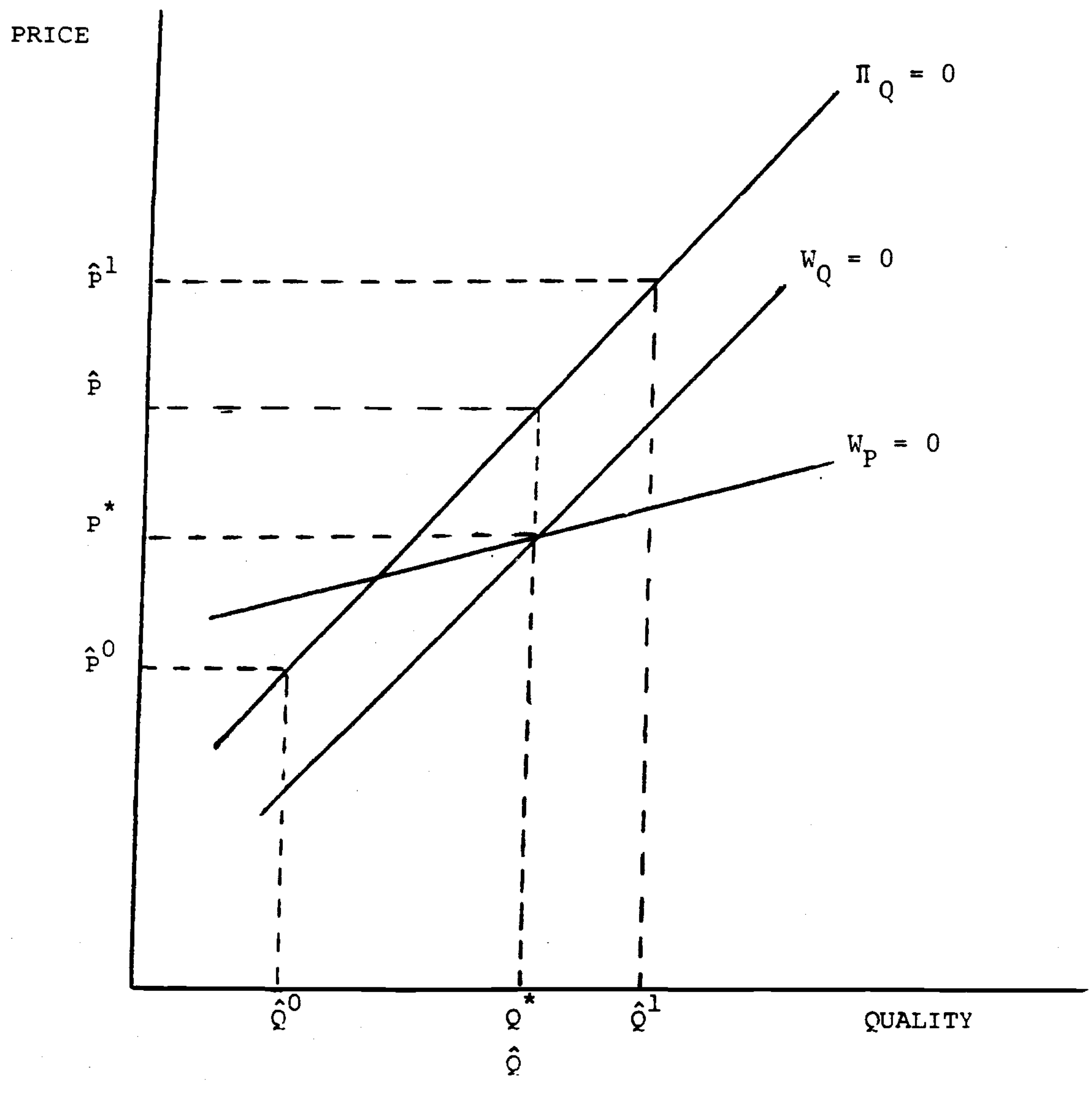
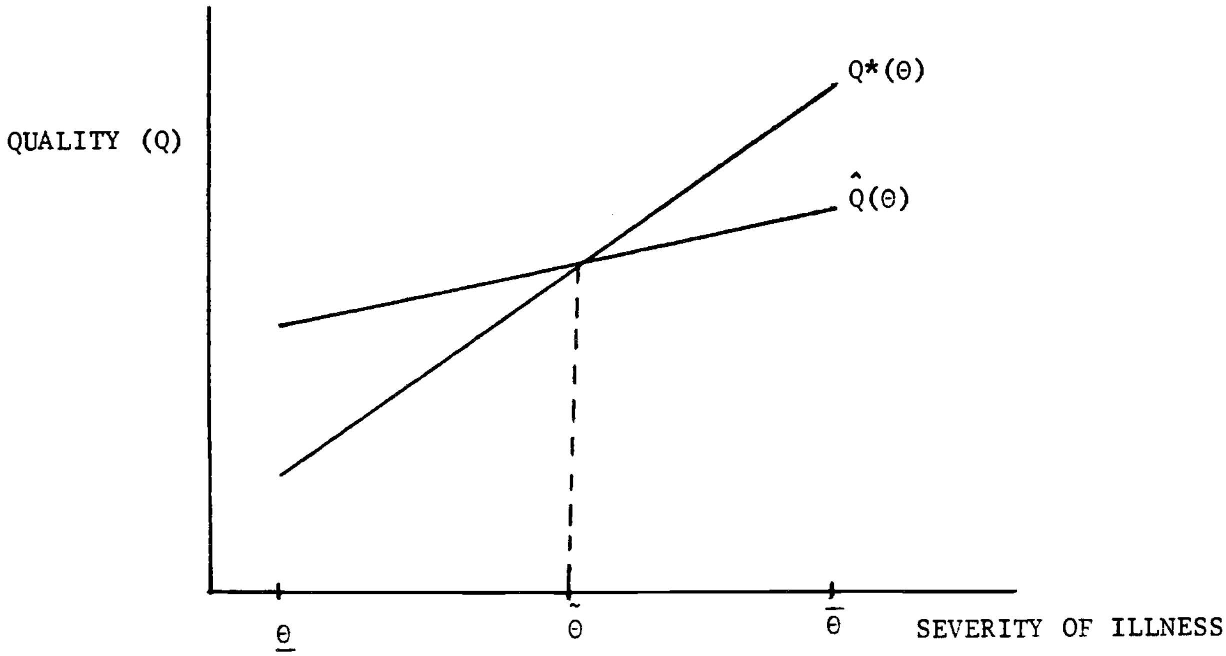


FIGURE 2

FIRST-BEST AND DRG QUALITY BY SEVERITY OF ILLNESS

PANEL A



PANEL B

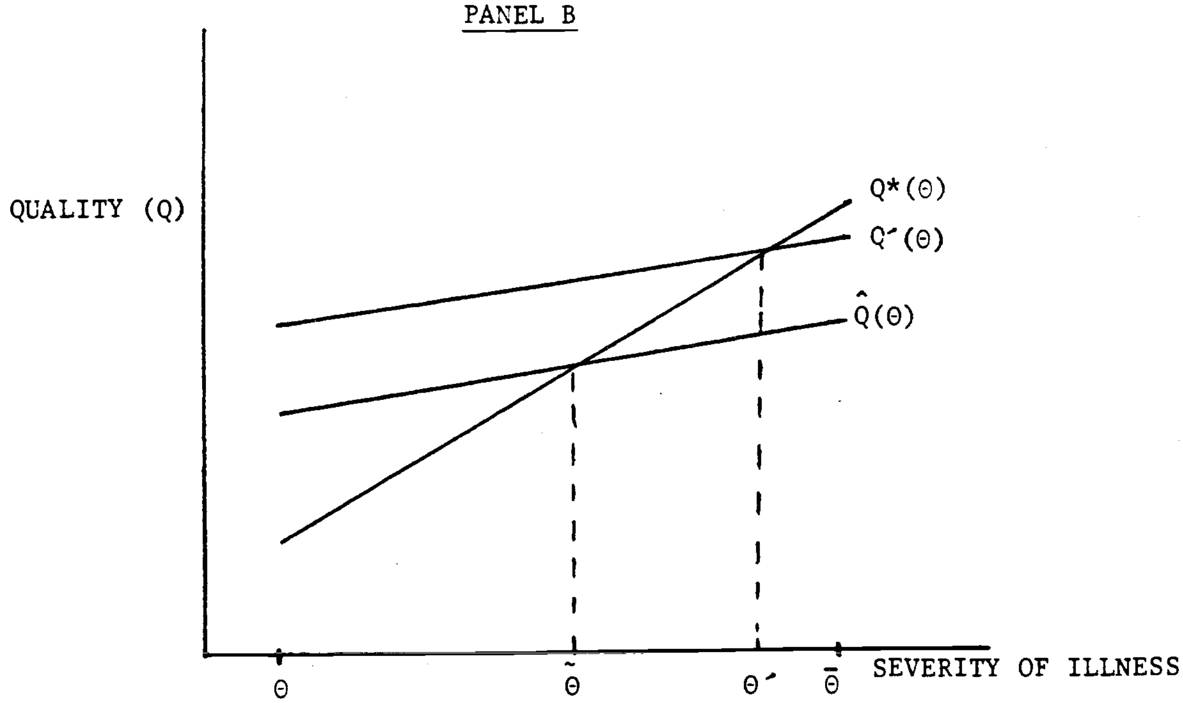


FIGURE 3

ETHICALLY MOTIVATED DEVIATIONS FROM PROFIT MAXIMIZATION