HMOs and Health Externalities:

A Local Public Good Perspective

Allen C. Goodman*

Miron Stano**

August 1998

JEL Classification: H70, I10

Key words: Local Public Goods, Externalities, Health Care, Health Maintenance Organizations (HMOs)

* Department of Economics, Wayne State University, Detroit MI 48202. Phone: 313-577-3235; FAX: 313-577-0149; e-mail: a.goodman@wayne.edu

** Department of Economics, Oakland University, Rochester MI 48309. Phone: 248-370-3291; FAX: 248-370-4275; e-mail: stano@oakland.edu

We are grateful to Sherman Folland, Janet Hankin, Julie Hunsaker, Lee Mobley, Joseph New-house, an anonymous referee, and the journal editor for their helpful comments. Goodman was supported in part by Grants No. DA08711 and DA10828 from the National Institute on Drug Abuse (NIDA) and by a grant from the Blue Cross Blue Shield of Michigan Foundation. Stano was supported by an Oakland University Faculty Fellowship. The usual disclaimer applies.

ABSTRACT

This article introduces a Health Maintenance Organization (HMO) model that characterizes HMO decisions on enrollment, utilization levels, and quality. Facing global budgets dependent on revenue raising capacity, HMOs must determine standards of care and the corresponding premiums to cover costs. Revenues are allocated among HMO providers by gatekeepers who are often subject to strict utilization controls. Our model identifies various uncompensated health and pecuniary externalities that may result from patient disenrollments, thereby affecting treatment costs. As a result of these externalities HMOs may provide less than socially efficient levels of care; and they may also enroll less than socially efficient numbers of members. The model also predicts that HMOs may promote socially cost inefficient treatment options for conditions where multiple options are available.

1. *Introduction*

Despite the advent of "prepaid plans," as Health Maintenance Organizations (HMOs) were first described in the 1920s, the vigorous opposition of organized medicine to any form of pre-payment and to plans not controlled by doctors severely impeded HMOs' growth. Policymakers anticipated rapid enrollment growth after the passage of the HMO Act of 1973, which provided subsidies to federally qualified HMOs and, more importantly, required firms with more than 25 employees to offer the HMO option if one was available in the firm's market area. Through the financial incentives they faced, HMOs were regarded as ideal mechanisms to keep patients heal-thy and to avoid excessive utilization. Although enrollments reached 9 million by 1980, the major period of growth and restructuring of health care delivery came only after 1980 as rapid increases in health care costs finally drove employers to demand private market reform through alternatives to traditional indemnity plans. HMOs and other entities that integrate insurance with the provision of health care enroll about half the U.S. population (Group Health Association of America, 1996) and nearly 75 percent of all workers with job-related insurance (Jensen et al. 1997).

Managed care, as characterized by HMOs, is widely credited with reducing the growth rate of national health care spending in the 1990s, and its continued expansion is supported by many policymakers. Nevertheless the strong public and legislative backlash at managed care suggests that it would be premature to conclude that restructuring has improved the nation's well being.

Our primary objective is to develop the proposition that HMOs have transformed health care from a largely private good, with utilization determined by the familiar physician-patient relationship, to one with qualities often associated with collective goods. The second objective is to advance a theoretical model of quality differences among HMOs to address the concerns associated with a managed

care environment. Here we augment our public goods analysis with insights on product differentiation from the industrial organization field.

The model incorporates two features neglected elsewhere. First, as a result of high disenrollment rates, HMOs will not fully recognize the effects of their treatment decisions on the population's health status. This, together with the capitation payment system that characterizes HMOs, creates a classic external benefit problem that leads to underprovision of health care relative to a social optimum.[1]

The second feature of our model, not generally recognized in the HMO literature, is that treatment therapies for many conditions (with accompanying costs) and treatment benefits may last for many periods. Disenrollment may create another externality problem for conditions with multiple treatment options if HMOs promote treatment choices that minimize current costs, because other parties may be responsible for future treatment costs. Since HMOs cannot fully capture the benefits of health investments in patients, they may underinvest in their members relative to economically efficient levels.

We introduce a simple model with random disenrollment that illustrates the impact of the health externality. We then sketch a multi-period variant where HMOs may respond to disenrollment by choosing treatments with lower up-front costs in preference to equally cost-effective (and possibly more economically efficient) alternatives with higher up-front costs.

2. A Public Good Perspective

Managed care plans share important characteristics with national health systems. HMOs, like many national health systems, rely on primary care gatekeepers who may be salaried government employees (e.g., India), or independent providers reimbursed either on a capitated basis (e.g., United Kingdom) or through fee-for-service (e.g., Canada).

Of more importance than staffing and reimbursement similarities, the availability and quality of services found in national health systems are constrained by global budgets dependent on revenue raising capacity. The fundamental problem facing any national health system is to determine the size and distribution of its health care pie. HMOs similarly need to determine standards of care and the corresponding premiums to cover costs. The resulting pie is allocated among HMO members by gatekeepers who are often subject to strict utilization controls. However, the agency role of physicians becomes more complicated than the traditional one-on-one relationship, in that the HMO's resource limitations require physicians to consider the needs of members other than their own patients. This expanded and more complex agency role mirrors the one in found in national health plans. As is common for social goods, HMOs also need various forms of nonprice rationing to constrain demand.

Unlike monopoly national health plans, however, HMOs in the United States must attract members in a marketplace with many insurance alternatives. It may be useful therefore to conceptualize HMOs as local public goods provided by competing political jurisdictions in order to capture both its collective goods and competitive features.[2] By drawing on the well-developed public finance literature on local public goods that originated with Tiebout (1956), and which has been used to study the selection of employee health benefits (Goldstein and Pauly, 1976), the social efficiency of a system of competing HMOs can be more readily evaluated.

3. *Impacts of HMOs*

---

1. In all cases, we refer to the social optimum as that output at which marginal social benefits equal marginal social costs.

2. Benjamini and Benjamini (1986, P. 226) developed a model to show that "… an HMO plan is a more efficient way of providing medical insurance for a relatively homogeneous group. Different groups may have different plans that best fit their members. People who do not find an HMO close to their ideal (or live far from an HMO medical center) would be better off using the CI [conventional insurance] plan." Indeed, attempts to exclude potentially high-risk clients have analogies to suburbs, which seek to exclude

A burgeoning HMO literature has concentrated on HMO cost, utilization patterns, and quality, relative to fee-for-service (FFS) care. The evidence suggests that HMOs economize by substituting outpatient care for inpatient care, and more generally by using fewer and less costly treatment procedures.[3] It also indicates that HMO enrollees receive more preventive care and that with possible exceptions such as those suffering from depression (Rogers et al. 1993), quality of care is not compromised. A variety of outcome measures indicate that patients with colorectal cancer, acute myocardial infarction, hypertension, and diabetes appear to receive medical and surgical care that is at least comparable to FFS care.[4]

Despite this apparently successful record, HMOs and managed care policies have faced continual attacks in the media, and through Congressional and state legislative hearings. The attacks have focused on gatekeeper and other strategies that can deny care or limit access to expensive therapies. Federal legislation has recently been sponsored to prohibit gag rules, which prevent HMO physicians from discussing the full range of treatment options with their patients. State level activities include proposals to hold plans liable if they fail to cover costs of doctors' recommended treatments and, through "any willing provider" laws, to give patients greater access to physicians or hospitals other than those on their HMO's approved list.

Reform through expansion of managed care and increased competition among managed care plans continues to receive strong support from health care analysts (Ellwood and Enthoven, 1995; Luft, 1996) despite a growing public and legislative backlash. Many believe that gatekeeper

---

low-income, and hence low property value owners. (Hamilton 1975, 1976). We do not pursue this issue here.

3. Miller and Luft (1994) review the post-1980 literature on utilization, cost, and quality. Chernew (1995) examines HMO use of diagnostic tests.

4. See, for example, Retchin and Brown (1990), Carlisle et al. (1992) and Greenfield (1996).

and other utilization management strategies enable HMOs to reduce moral hazard, supplier-in-duced demand, and other forms of excessive utilization associated with traditional FFS insurance.

Other effects associated with HMOs add to their appeal. HMOs have potentially strong competitive impacts in limiting insurance premiums. Some also feel that HMOs will emphasize preventive care through incentives to avoid future treatment costs by keeping patients healthy.[5]

Surprisingly, HMO performance is ambiguous even on basic and widely accepted propositions. HMOs appear to economize by using fewer and less expensive treatments. However, the challenge of determining whether HMOs gain most of their cost advantages through lower inpatient care utilization, through favorable selection of both healthier patients and conservative providers, or through favorable cost arrangements with providers, has not been satisfactorily resolved (Pauly, 1986; Baker and Corts, 1996; Newhouse, 1996).[6]

Similarly, predictions that HMOs provide more preventive care (screening, immunization, and counseling services) are often supported. Nevertheless, Schauffler and Rodriguez (1993) review a substantial number of studies that fail to find any major differences between HMO and FFS use rates for preventive care. Further, a 1995 household survey by Donelon et al. (1996) indicated that patients in managed care plans are no more likely than those in FFS plans to receive physician encouragement to get preventive services.

Finally, the competitive effects of HMOs on the indemnity sector seem more complex than once believed. Baker and Corts (1996) found that HMO market penetration rates of 10 to 15

---

5. Frank and Welch (1985), Feldman et al. (1990), and Wholey, Feldman, and Christianson (1995) examine competitive impacts. Schauffler and Rodriguez (1993) review the issues and empirical findings with respect to preventive care.

6. To complicate matters further, after controlling for selection phenomena, Goldman (1995) found that managed care strategies in large scale public sector programs such as CHAMPUS, fail to reduce inpatient utilization while encouraging large increases in outpatient care.

percent reduce the premiums charged by traditional FFS insurers. In contrast, HMO penetration rates above 20 percent result in substantial increases in FFS premiums through cost-shifting and other practices that dominate the cost-reducing effects of increased competition.

Health care market complexities create difficulties in sorting out the effects of many confounding forces. The problem is compounded by the growing variation among HMOs and other managed care plans, as well as a narrowing of differences between HMOs and indemnity plans as the latter adopt utilization management (Morrison and Luft, 1990). However, we believe that a major impediment to a clearer understanding of HMOs and their effects is the absence of well-developed theoretical models of HMO behavior. With few exceptions, the existing literature consists of empirical tests of ad hoc, intuitive propositions, many of which were developed and systematically categorized in Luft's classic 1981 contribution.[7]

4. Disenrollment

Due to its potential importance, disenrollment merits further discussion.[8] Disenrollment can be voluntary due to switching among available plans, but it is more commonly mandated due to termination, layoff, retirement, or changes in available plans or employee status. Annual dis-enrollment rates of 20 percent or higher are not unusual. Robinson and Gardner (1996), examin-ing a large California firm, find that only 65 percent of enrollees stayed with their FFS plans be-tween 1989 and 1990; 72 percent remained with a major HMO; and 65 percent stayed with other available HMOs. The low retention rates stemmed primarily from involuntary terminations. Nor does the disenrollment rate diminish with length of

7. Exceptions are Debrock and Arnould (1992), Hill and Wolfe (1994, 1997), and Baker and Corts (1995).

8. See Wersinger and Sorenson (1982), Mechanic, Weiss, and Cleary (1983), Hennelly and Boxerman (1983a, 1983b), Lewis (1984), and Harrington, Newcomer, and Presten (1993).

enrollment (Long. Settle, and Wrightson, 1988). Disenrollment rates of 15 or 20 percent per year render the likelihood that a patient will remain with the same HMO for several years to be relatively low.

Some of the financial and quality of care problems associated with patient disenrollments are well known. Patients and providers face the difficulties of maintaining continuity of care and complete medical records. HMOs face the added financial burdens resulting from higher patient recruiting costs, disruption of cash flows, and upward pressure on premiums for continuing members if lower risks are more likely to disenroll. The literature has also recognized that dis-enrollment diminishes HMOs' incentives to provide preventive care (Robinson and Gardner, 1996). After all, why emphasize preventive care for a patient who is not likely remain a member, when that care provides the greatest return in the form of averted future treatment costs?

Elsewhere, Stano (1996) has hypothesized that disenrollment can also influence choice of treatment when multiple treatment alternatives are available. There are many conditions for which continuing care can substitute for high cost surgery or other invasive and capital-intensive forms of care. For example, treatment alternatives for prostate cancer include "watchful waiting," surgery, external radiation, and "high tech" seed therapy (radioactive seed implants) with external radia-tion. Mental health and substance abuse problems, among the most common sources of hospital admission (Lemrow et al. 1990), provide a more widespread and generic example of substitution possibilities. Here, drugs and lower level professional therapy can often be used to replace inten-sive inpatient psychiatric care.[9] Many other common acute and chronic conditions, including those of the heart, back and joints, involve choices among similarly wide ranges of alternatives.

In particular, disenrollments encourage HMOs to substitute continuing care for up-front capital

intensive services. In some cases, an HMO, acting in its profit-maximizing interests, might opt for a socially less efficient alternative. With rapid technological changes producing new treat-ment options with substantial impacts on health care spending, HMO decision-making must be carefully examined. This need becomes even more relevant as the U.S. and other countries, exper-iencing rapid growth of national health care costs, pursue reforms based on managed competition.

While the potential loss of patients may influence treatment decisions of FFS providers as well as HMOs, the capitation method of payment to HMOs renders the disenrollment problem particularly important. FFS providers are paid for each unit of care. Aside from uncollectibles, they are not at risk of losing money on services provided currently or in the future.

In contrast, by integrating insurance with the provision of health care, the HMO receives a fixed payment per enrollee to cover costs in the current period, and over time, for those who remain enrolled. Thus, unlike FFS care, where payment in every period is very likely to cover costs, the HMO must consider the timing of expenditures and the financial losses of overspending on patients who may disenroll. One way for an HMO to "self-insure" against long term losses attributable to disenrollment is to economize on care for those currently enrolled.

5. A Competitive Model of HMOs

We propose an HMO model that addresses several salient issues. An HMO will wish to determine both the number of consumers to serve and the level of service to provide. One might consider an analogy to urban suburbs, which set their public service levels to satisfy residents who are willing to pay for the services packages provided. Our analysis borrows both from local public finance and from

---

9. Managed care has been widely attacked in the media for these kinds of substitutions. See, for example, Pollock (1995) and Hymovitz (1995).

industrial organization. Local public goods packages are determined by consumer preferences and by provider cost conditions. Industrial organization provides a framework for analyzing quality differentials under a variety of market structures.[10]

We argue that HMO treatment costs are related to member health status, which is a func-tion of care received at the HMO and elsewhere, as well as other factors including diet, environ-ment, and life style. People's long term health relates in part to short term decisions about how much care to provide. This follows directly from the long term nature of human capital, most par-ticularly health capital (Grossman, 1972a, 1972b). Thus treatment decisions at one HMO may affect treatment costs at other HMOs, an externality that must be addressed in characterizing HMO market solutions, and optimal levels of health care.[11]

We start with a model in which firms face random enrollment and disenrollment related to changes in tastes, changes in population, or employment conditions (related to those having insur-ance coverage). This simple model illustrates key externality factors related to health status and health care in an HMO setting, showing how the unpriced externality leads to a socially inefficient level of care. We finish by examining HMO treatment decisions in an intertemporal context, and show a systematic HMO bias toward continuing low cost care rather than "big ticket" technology.

a. Random Disenrollment

Begin with an HMO providing coverage (services) x_i per enrollee to its insured k_i who are

10. Wildasin (1986, Section 4) provides a more complete discussion of public finance literature. Goldstein and Pauly (1976) use a related approach in discussing workers who choose among firms, each of which offers a single health insurance plan.

11. This is analogous to the "spillover" argument for public provision of education, suggesting that an educated public benefits those (taxpayers) who pay for, but do not receive, the education.

representative of the population.[12] HMOs vary by levels k and x, with members changing HMOs due to changing tastes for x, or due to employment-related conditions. Hence k_i (or alternatively price p_i) and x_i indicate decision variables for HMO_i. We assume (following Salop 1979) that there are enough HMOs in the market (i.e., the costs of entry are not too high) so that there is some competition among them.[13]

It is appropriate to presume that the HMO recognizes that its pricing decisions may influence its enrollment k. The HMO faces inverse demand curve $p = P(k, x)$, where $P_k < 0$, and $P_x > 0$ (subscripts k, x, and y all refer to partial derivatives). Suppose that there are a large number of consumers with unit demands (they join this HMO, join another one, purchase from providers in the fee-for-service sector, or make no purchases at all). As noted in Figure 1, $P(r, x)$ is the maximum annual premium (including the employer's contribution) that consumer r is willing to pay for a unit of HMO coverage of quality x.

(Figure 1 - Inverse demand curve facing HMO)

Let C^i represent the HMO's total annual cost to provide expected level x_i of health services to its k_i enrolled members. Total cost C^i is related positively to x_i, positively to number of person-years of membership k_i, and negatively to "health" y of those who happen to be members at the time. Having k_i as an argument in the cost function allows increasing and/or decreasing returns to scale for both k_i and x_i.[14]

Assume that health y is positively related to the level of services by all providers. Since at any time in the future, these individuals may be members, average health status y of the population equals $S\, k_i$

12. Assume that the structure of the firm represents a "staff model" HMO (in which physicians are firm employees), although the model also applies to other forms of integrated service delivery.

13. Determining the optimal number of firms, and the impacts of entry and/or exit, refers more explicitly to vertical differentiation in the product market and is beyond the scope of this paper.

14. By analogy, consider a "production run" of quantity k_i cars, each with a vector of character-istics **x** including size, power, and carrying capacity (among other elements). We follow Rosen (1974) in his treatment of market equilibrium for composite goods.

x_i divided by population K. A constant in the denominator is unimportant in our analysis, so we normalize it as 1; hence $y = \Sigma k_i x_i$.

Profits for HMO_i can be defined as:

$$\pi_i = k_i P^i (k_i, x_i) - C^i (k_i, x_i, y) \quad (1)$$

with $C^i_k > 0$, $C^i_x > 0$, and $C^i_y < 0$ (with C^i_{kk}, C^i_{xx}, and C^i_{yy} all positive).

The first decision (optimizing with respect to k_i) leads to the condition:

$$\partial \pi_i / \partial k_i = P (k_i, x_i) + k_i P^i_k(k_i, x_i) = C^i_k \quad (2a)$$

which states that the premium (i.e. marginal revenue product per enrollee) must equal the marginal cost of the enrollee. Since an individual HMO is small compared to the entire population, the HMO does not recognize its impact through k_i on population health status y.

The second decision is how much x_i to provide. Again, here through x_i, the HMO does not recognize its impact on population health status y. Thus, for any k_i, the maximization is:

$$\partial \pi_i / \partial x_i = k_i P_x (k_i, x_i) = C^i_x. \quad (2b)$$

This equality indicates that the marginal revenue of providing one more unit of quality (services) to the entire enrollment k_i equals the marginal cost of providing that additional unit.

The result is familiar to students of monopoly or monopolistic competition. The firm with some market power is concerned with the impact of output on market price and may thus serve a different amount of consumers than the perfect competitor, irrespective of the health status of the population that it serves.

That this occurs is indicated by the following depiction of the societal welfare function:

$$W(k,x) = [\int_0^k P(r, x)\, dr - pk\,] + [pk - C (k, x, y)] \quad (3)$$

$$= \textit{Consumer surplus} + \textit{Producer surplus}$$

Ignoring the health externality, first order conditions for social optimum are:

$$\partial W/\partial k \equiv P(k, x) = C_k \qquad (4a)$$

Price for the marginal consumer = marginal cost.

$$\partial W/\partial x = \int_0^k \partial P(r, x)/\partial x \, dr - C_x = 0$$

$$\equiv \int_0^k \partial P(r, x)/\partial x \, dr = C_x . \qquad (4b)$$

where $\partial P(r, x)/\partial x$ = marginal willingness to pay for an additional unit of quality by consumer r.

$\int_0^k \partial P(r, x)/\partial x \, dr$ = Gross marginal surplus with respect to quality.

So, total benefit = *MC* of supplying additional unit of quality to all k consumers.

Tirole (1988) proves that the incentive to provide quality is related to the marginal willingness to pay for quality for the marginal consumer in the case of the monopolist, and for the average consumer in the case of a social planner. These two can be easily compared only if the output is the same in both arrangements. If not, the (somewhat weaker) conclusion is that since the marginal consumer is not generally representative of the average consumer the market structure is likely to produce bias in product selection at any given output.

b. The Health Externality

The previous section indicated that profit-maximizing HMOs might not recognize the impact of health services on health *y*. It is useful to compare the market optimum of marginal revenue equaling marginal cost, to an optimum derived alternatively by a "planner" who is aware of the externality of each HMO's level of *x* on health *y*, or by an entrepreneur who seeks to maximize the joint profits of a group of two or more HMOs.

Consider a situation with two firms (with a ready generalization to *n* firms). The entrepreneur seeks to maximize total profits *T*, or with appropriate substitutions:

$$T = \pi_1 + \pi_2 = k_1 P^1 (k_1, x_1) + k_2 P^2 (k_2, x_2) - C^1 (k_1, x_1, y) - C^2 (k_2, x_2, y) \quad (5)$$

To maximize profits, the entrepreneur must consider the impacts of health services x_1 and x_2 on the health of others elsewhere. Differentiating (5) with respect to x_1 and x_2 yields:

$$k_1 P^1_x (k_1, x_1) = C^1_x + k_1 (C^1_y + C^2_y), \text{ and} \quad (6a)$$

$$k_2 P^2_x (k_2, x_2) = C^2_x + k_2 (C^1_y + C^2_y). \quad (6b)$$

This is comparable to (2b), but $C^i_y < 0$ implies that the social optimum leads to more treatment x_i than the competitive monopolistic market.

(Figure 2 - Externality model with random disenrollment)

This is easily shown in Figure 2. Without the externality, HMO_1 optimizes at point *A*, giving level x_1^{mkt}. The optimal level of x_1 at point *B* is x_1^{opt}, as noted by the downward shift in the right hand side by the factor $k_1 (C^1_y + C^2_y)$, which is unambiguously negative. This indicates an inefficiently small level of HMO care *x*, and by implication a substitution of non-HMO and/or non-health care inputs (such as the patient's own time) for the HMO care.

Alternatively, consider the social planner optimizing (3) with respect to *y*, as well as *x* and *k*. Following (4b), optimization with respect to x_i (and *y*) yields:

$$\int_0^k \partial P(r, x)/\partial x \, dr = C_x + k_i \sum_i C^i y \quad (7)$$

Optimization with respect to k_i yields:

$$P(k, x) = C_k + k_i \sum_i C^i y \quad (8)$$

It is important to view our optimizations in the context of the current health system. The HMO optimum sets marginal revenue $k_i P^i_x$ equal to marginal cost C^i_x. Under a traditional FFS arrangement,

there is a separation of insurance from health care provision. With limited or zero coinsurance, the individual decision, given the insurance coverage, is likely to lead (through moral hazard) to inefficiently high levels of x. With marginal private benefits less than the marginal costs of providing the care, the standard "deadweight loss" occurs.[15]

We have established that the health externality exists in the context of either a monopolistic provider or a utility maximizing social planner. In either circumstance, failure to address the externality leads to inefficiently small numbers of consumers enrolled in HMOs, and inefficiently low quality for those who are members.

In our analysis, inefficiency in k indicates that some number of consumers use providers who offer higher cost care than is available through the HMO. Inefficiency in x indicates that the consumers in the HMOs use either more non-HMO care, more non-health goods, or more leisure time (following Grossman) to produce their desired levels of health.

c. Disenrollment and the Tort System

An efficient tort system, along with patients who are well informed about quality of care, will mitigate the externality effects. The HMO will have difficulty "under-providing" care due to the threat of both disenrollment and the financial penalties imposed by the malpractice market. In the limiting case of perfect information and a perfectly efficient legal system, the externality will be fully internalized. However, there are significant limitations on the discipline that can be exerted by legal remedies and patients in today's health care environment.

On the issue of tort laws, although there is agreement that the U.S. medical malpractice system

15. Even under the standard HMO contract, a consumer facing zero marginal out-of-pocket cost (although possibly considerable time cost) may wish to consume care beyond the efficient level of

has reduced negligence (White 1994), employer sponsored managed care plans are exempt under ERISA (the Employee Retirement Security Act of 1974) from usual state law tort claims for medical malpractice and wrongful death.[16] Aside from this exemption, the enforcement costs of tort liability rules (to patients, providers/insurers, and the courts) must be considered (Olsen, in press) as well as other imperfections that erode the efficiency of the malpractice market as a deterrent to inappropriate care. These errors include liability judgments when there is no negligence (type I error) and absence of liability when there is negligence (type II error). A revealing indication of the imperfections in the litigation system comes from the Harvard Medical Practice Study which found that a substantial proportion (27.6 percent) of adverse effects in hospitalizations result from negligence (Brennan et al. 1991). Of the cases involving negligence, only 1.53 percent resulted in malpractice claims (Localio et al. 1991).[17]

As for the discipline resulting from patients' monitoring of quality and their ability to switch plans, the development of quality indicators and "report cards" is still evolving (Eddy 1998, Kertesz 1998). The most prominent and ambitious approach to quality evaluation is the Health Plan Employer Data and Information Set (HEDIS) maintained by the National Commission on Quality Assurance (NCQA), a voluntary accreditation organization for HMOs. HEDIS represents over 50 plan performance measures with five components including quality of care and access and satisfaction (NCQA, 1998). Quality indicators include rates of Cesarean section, low birth weight, and tobacco use, where low rates are preferable, and preventive care measures, such as childhood immunization rates and screening rates for

---

marginal private benefit equaling marginal cost. Analyses of the conflicts between HMO and patient interests are beyond the scope of this paper.

16. The legal protection to managed care organizations and employers is the object of intense scrutiny in current Congressional efforts to promulgate a "patients' health care bill of rights".

17. In a follow-up of malpractice claims, Brennan, Sox and Burstin (1996) found substantial errors of both type I and type II. The severity of the patient's disability rather than negligence was the only statistically significant determinant of payment.

breast cancer and cholesterol, where high rates are preferable. There is little hard evidence on how and the extent to which patients use report cards. However, Chernew and Scanlon (1998) recently found that "employees do not ap-pear to respond strongly to plan performance measures" (p. 19**),** and surveys suggest that recommendations by family and friends are more important than ratings developed by experts (Blendon et al. 1998).

6. Disenrollment and Treatment Choice

We have established that disenrollment can affect the levels of services x_i provided. How-ever, by distinguishing only among *levels* of x_i, we have not distinguished among alternative treatment methods. In this section, we consider the longer term consequences that potential dis-enrollment can have on HMO treatment practices. We will show that in the presence of expected disenrollment, HMOs will tend to use "low-tech" treatments with smaller up-front costs, even when the present discounted value (PDV) of the costs equals the PDV for "high-tech" treatments.

Consider an infinitely-lived HMO that serves overlapping cohorts of customers for either one or two periods. In this variant, the HMO receives constant revenue per patient each period, so its goal is to minimize costs. Costs are modeled for Cohort 1 entering in Period 1:

- For "high-tech," possibly capital-intensive procedures, the HMO incurs costs M_1 in Period 1, and 0 in Period 2.
- For low-tech, less capital-intensive procedures, the HMO incurs costs m_{11} and m_{12} in Periods 1 and 2, where m_{ij} refers to cohort i at time j.

For Cohort 1, the HMO would plan "high-tech" procedures if:

$$M_1 < m_{11} + m_{12} / (1 + r). \qquad (9)$$

where r is the appropriate interest rate for discounting.

However, suppose that potential HMO members use expenditures M_t and m_{tt} as indicators of HMO quality. Thus, probability g of disenrollment and incurring the Period 2 costs is related to (M_t - m_{tt}); holding m_{tt} constant, the larger M_t, the smaller g, and the less likely that people will disenroll. So the HMO uses high-tech procedures for Cohort 1 if:

$$M_1 < m_{11} + (1 - g)\, m_{12} / (1 + r). \tag{10}$$

To maintain revenue, Cohort 2, of size g, must be enrolled in Period 2. The Cohort 2 HMO decision is similar to the Cohort 1 decision if:

$$M_2 < m_{22} + (1 - g)\, m_{23} / (1 + r). \tag{11}$$

So, for any two period sequence, the firm chooses the high-tech treatment if:

$$M_1 + M_2/(1 + r) < m_{11} + (1 - g)\, m_{12} / (1 + r) + m_{22}/(1 + r) + (1 - g)\, m_{23}/(1 + r)^2. \tag{12}$$

Over time, the HMO will be indifferent between high-tech and low-tech options if:

$$\sum_{1}^{\infty} M_t\,(1+r)^{1-t} = \sum_{1}^{\infty} [m_{tt}\,(1+r)^{1-t} + (1-g)\, m_{tt+1}\,(1+r)^{-t}]. \tag{13}$$

Assuming that all $M_t = M$, and that all $m_{ij} = m$, expanding both sides, and taking limits as $t \to \infty$, the HMO chooses high-tech if:

$$M < m\,[1 + (1-g)/(1+r)], \tag{14a}$$

and low-tech if:

$$M > m\,[1 + (1-g)/(1+r)]. \tag{14b}$$

Thus the higher the disenrollment rate g, the more important is the disenrollment effect. If $g = 0$, the HMO faces the standard investment criterion, comparing first period costs with discounted future costs. Its decision here will be economically efficient.

With the likelihood that increased competition through increased choice will raise g, man-aged

competition and other competitive strategies must be more carefully examined. If *g* increases, *m [1 + (1-g)/(1+r)]* decreases, and continuing care *m* becomes the more financially viable option even if the PDVs are equal, and even if treatment *M* is more economically efficient in producing health. In effect, HMOs self-insure against future disenrollment by reducing current costs through (low cost) continuing care rather than high-tech treatment.[18]

7. *Conclusions*

The growing role of HMOs has emphasized the need for a theoretical framework for understanding managed care organizations and helping policymakers develop more effective strategies for health system reform. This article models HMO decisions on enrollment and service intensity, recognizing that in many circumstances health care provides an unpriced external bene-fit. By producing where marginal private revenue equals marginal private cost, profit-maximizing HMOs will likely underprovide services or the quality of the services relative to efficient levels.

It is not surprising, then, that many studies have found HMO cost-savings relative to FFS plans. FFS plans encourage overutilization to the point where marginal private benefits can be far less than marginal cost. HMOs are widely believed to discourage this deadweight loss and other forms of overutilization such as supplier-induced demand.

Aside from the inefficiently low overall utilization levels and enrollments, a second exter-nality resulting from patient disenrollment may distort the HMO's treatment choice for conditions with alternative treatment options. HMOs will be biased toward therapies with low up-front costs because other payers will be responsible for the future treatment costs of disenrolled members. Increased competition, by

18. If low cost service signifies inadequate service, the externality may be at least partially internalized through consumer mobility among HMOs. However, the transactions costs of changing providers, and

making more insurance and provider options available, may increase disenrollment rates (Long, Settle, Wrightson, 1988) and thus aggravate this inefficiency as well.

Our utilization predictions are consistent with the rising conflict between HMOs and their members. On the one hand, complaints about inadequate levels of services and the unavailability of expensive, high-tech treatment options are more common. As noted previously, legislation aimed at regulating HMO practices is becoming more widespread.

On the other hand, if conflicts between HMOs and their members relate at least in part to competition and to the externality problems described here, policies to internalize the externalities may be wiser. Such policies might include permitting some forms of cooperation among insurers, reforming tort law to increase the efficiency of the malpractice markets, and improving the dis-semination of quality information to health care consumers and managed care providers. Alternatively, arrangements to minimize uncompensated external effects may evolve among insurers even without government intervention or encouragement.[19]

In sum, there are still many unknown consequences associated with HMOs and other managed care plans. Advocates of system reform through managed care and managed competition must view health system reorganization with caution. There may be no simple way to reduce costs while maintaining the quality of health and health care over the long term.

---

such institutional constraints as limited open-enrollment periods (only once per year for a short time) tend to limit endogenous consumer mobility among HMOs.

19. The Coase Theorem suggests that in the absence of transactions costs, all externalities will be internalized through negotiations, leading to an efficient allocation. It is unlikely, however, that the costs of negotiating or policing agreements among large numbers of HMOs and other providers would permit this solution.

References

Baker L.C., and K.S. Corts. 1995. The effects of HMOs on conventional insurance premiums: Theory and evidence. National Bureau of Economic Research, Working Paper 5356.

_____. 1996. HMO penetration and the cost of health care: Market discipline or market segmentation. *American Economic Review* 86: 389-394.

Benjamini Y., and Y. Benjamini. 1986. The choice among medical insurance plans. *American Economic Review* 76:221-227.

Blendon, R.J., M. Brodie, J.M. Benson, et al. 1998. Understanding the managed care backlash. *Health Affairs* 17:80-94.

Brennan, T.A., L. Leape, N.M. Laird, et al. 1991. Incidence of adverse events and negligence in hospitalized patients: Results of the Harvard Medical Malpractice Study I. *New England Journal of Medicine* 324:370-376.

Brennan, T.A., C.M. Sox, and H.R. Burstin. 1996. Relation between negligent adverse events and the outcomes of medical malpractice litigation. *New England Journal of Medicine* 335:1963-1967.

Carlisle, D.M., A.L. Siu, E.B. Keeler, et al. 1992. HMO vs fee-for-service care of older persons with acute myocardial infarction. *American Journal of Public Health* 82:1626-1630.

Chernew, M. 1995. HMO use of diagnostic tests: A review of the evidence. *Medical Care Research and Review* 52:196-222.

Chernew, M., and D.P. Scanlon. 1998. Health plan report cards and insurance choices. *Inquiry* 35:9-22.

Debrock, L., and R.J. Arnould. 1992. Utilization controls in HMOs. *Quarterly Review of Economics and Finance* 32:31-53.

Donelon, K., R.J. Blendon, J. Benson, et al. 1996. All payer, single payer, managed care, no payer: Patients' perspectives in three nations. *Health Affairs* 15:254-265.

Eddy, D.M. 1998. Performance measurement: Problems and solutions. *Health Affairs* 17:7-25.

Ellwood, P.M., and A.C. Enthoven. 1995. 'Responsible choices': The Jackson Hole group plan for health reform. *Health Affairs* 14:24-39.

Feldman, R., H. Chan, J. Kralewski, et al. 1990. Effects of HMOs on the creation of competitive markets for hospital services. *Journal of Health Economics* 9:207-222.

Frank R.G., and W.P. Welch. 1985. The competitive effects of HMOs: A review of the evidence. *Inquiry* 22:148-161.

Goldman, D.P. 1995. Managed care as a public cost containment mechanism. *Rand Journal of Economics* 26:277-295.

Goldstein, G.S., and M.V. Pauly. 1976. Group insurance as a local public good. In *The Role of Health Insurance in the Health Services Sector*, R. Rossett Ed. New York: Neal Watson.

Greenfield, S. 1996. Comparison by systems and specialties of medical outcomes in patients with hypertension and non-insulin dependent diabetes. *American Journal of Managed Care* 2:535-538.

Grossman M. 1972a. *The Demand for Health: A Theoretical and Empirical Investigation*, New York: Columbia (for the National Bureau of Economic Research).

_____. 1972b. On the concept of health capital and the demand for health. *Journal of Political Economy* 80:223-255.

Group Health Association of America. 1996. *1995 Sourcebook on HMO utilization data.* Washington, DC.

Hamilton, B.W. 1975. Zoning and property taxation in a system of local governments. *Urban Studies* 12:205-11.

_____. 1976. Capitalization of interjurisdictional differences in local tax prices. *American Economic Review* 66:743-53.

Harrington, C., R.J. Newcomer, and S. Presten. 1993. A comparison of S/HMO disenrollees and continuing members. *Inquiry* 30:429-440.

Hennelly, V.D., and S.B. Boxerman 1983a. Out-of-plan use and disenrollment outgrowths of dissatisfaction with a prepaid group plan. *Medical Care* 21:348-358.

_____. 1983b. Disenrollment from a prepaid group plan: A multivariate analysis. *Medical Care* 21:1154-1167.

Hill, S.C., and B.L. Wolfe. 1994. Testing the HMO competitive strategy: An analysis of its success and failure in Wisconsin. Presentation at the Meetings of the Western Economic Association, Vancouver, British Columbia, July 3, 1994.

_____. 1997. Testing the HMO competitive strategy: An analysis of its impact on medical resources. *Journal of Health Economics* 16:281-286.

Hymowitz, C. 1995. In the name of Freud, why are psychiatrists complaining so much? *Wall Street Journal*, December 21:A1.

Jensen, G.A., M.M. Morrisey, S. Gaffney, and D.K. Liston. 1997. The new dominance of managed care: Insurance trends in the 1990s. *Health Affairs* 16:125-136.

Kertesz, L. 1998. Making sense of health plan data. *Modern Healthcare*, April 20:99-106.

Lemrow, N., D. Adams, R. Coffey, and D. Farley. 1990. The 50 Most Frequent Diagnosis-related Groups (DRGs), Diagnoses, and Procedures: Statistics by Hospital Size and Location. U.S.

Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research (DHHS Publication No. (PHS) 90-3465), Washington, DC.

Lewis, K. 1984. Comparison of use by enrolled and recently disenrolled populations in a health maintenance organization. *Health Services Research* 19:1-22.

Localio, R.A., A.G. Lawthers, T.A. Brennan, et al. 1991. Relation between malpractice claims and adverse events due to negligence: Results of the Harvard Medical Practice Study III. *New England Journal of Medicine* 325:245-248.

Long , S., R.F. Settle, and C.W. Wrightson. 1988. Employee premiums, availability of alternative plans, and HMO disenrollment. *Medical Care* 26:927-938.

Luft, H.S. 1981. *Health Maintenance Organizations: Dimensions of Performance*. New York: John Wiley & Sons.

_____. 1996. Modifying managed competition to address cost and quality. *Health Affairs* 15:23-38.

Mechanic, D., N. Weiss, and P.D. Cleary. 1983. The growth of HMOs: Issues of enrollment and disenrollment. *Medical Care* 21:338-347.

Miller, R.H., and H.S. Luft. 1994. Managed care plan performance since 1980: A literature analysis. *Journal of the American Medical Association* 271:1512-1519.

Morrison, E.M., and H.S. Luft. 1990. Health maintenance organization environments in the 1980s and beyond. *Health Care Financing Review* 12:81-90.

National Commission for Quality Assurance (NCQA). 1998. HEDIS 3.0. Washington, DC.

Newhouse, J.P. 1996. Reimbursing health plans and health providers: Selection versus efficiency in production. *Journal of Economic Literature* 34:1236-1263.

Olsen, R.N. In press. The efficiency of medical malpractice: A new appraisal. *Research in Law and Economics*.

Pauly, M.V. 1986. Taxation, health insurance, and market failure in the medical economy. *Journal of Economic Literature* 24:629-675.

Pollock, E.F. 1995. Managed care's focus on psychiatric drugs alarms many doctors. *Wall Street Journal*, December 1, p. A1.

Retchin, S.M., and B. Brown. 1990. Management of colorectal cancer in Medicare health maintenance organizations. *Journal of General Internal Medicine* 5:110-114.

Robinson, J.C., and L.B. Gardner. 1996. Involuntary health plan switching: Case study of a corporate health benefits program. *Medical Care Research and Review* 53 (1996): 225-239.

Rogers, W.H., K.B. Wells, L.S. Meredith LS, et al. 1993. Outcomes for adult outpatients with depression under prepaid or fee-for-service financing. *Archives of General Psychiatry* 50:517-525.

Rosen, S. 1974. Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy* 82:34-55.

Salop, S. 1979. Monopolistic competition with outside goods. *Bell Journal of Economics* 10:141-156.

Schauffler, H.H., and T. Rodriguez. 1993. Managed care for preventive services: a review of policy options. *Medical Care Research and Review* 50:153-198.

Stano, M. 1996. An alternative framework for evaluating the efficiency of managed care. *American Journal of Managed Care* 2:639-644.

Tiebout, C.M. 1956. A pure theory of local expenditures. *Journal of Political Economy* 64: 416-424.

Tirole, J. 1988. *The Theory of Industrial Organization*, Cambridge MA: MIT, 1988:100-102.

Welch, W.P. 1986. The elasticity of demand for health maintenance organizations. *Journal of Human Resources*: 21:252-266.

Wersinger RP, and A.A. Sorensen. 1982. Demographic characteristics and prior utilization experience of HMO disenrollees compared with total membership. *Medical Care* 20:1188-1196.

White, MJ. 1994. The value of liability in medical malpractice. *Health Affairs* 13:75-87.

Wholey, D.,R. Feldman, and J.B. Christianson. 1995. The effect of market structure on HMO premiums. *Journal of Health Economics* 14:81-105.

Wildasin, D.E. 1986. *Urban Public Finance*. London: Harwood Academic Publishers.

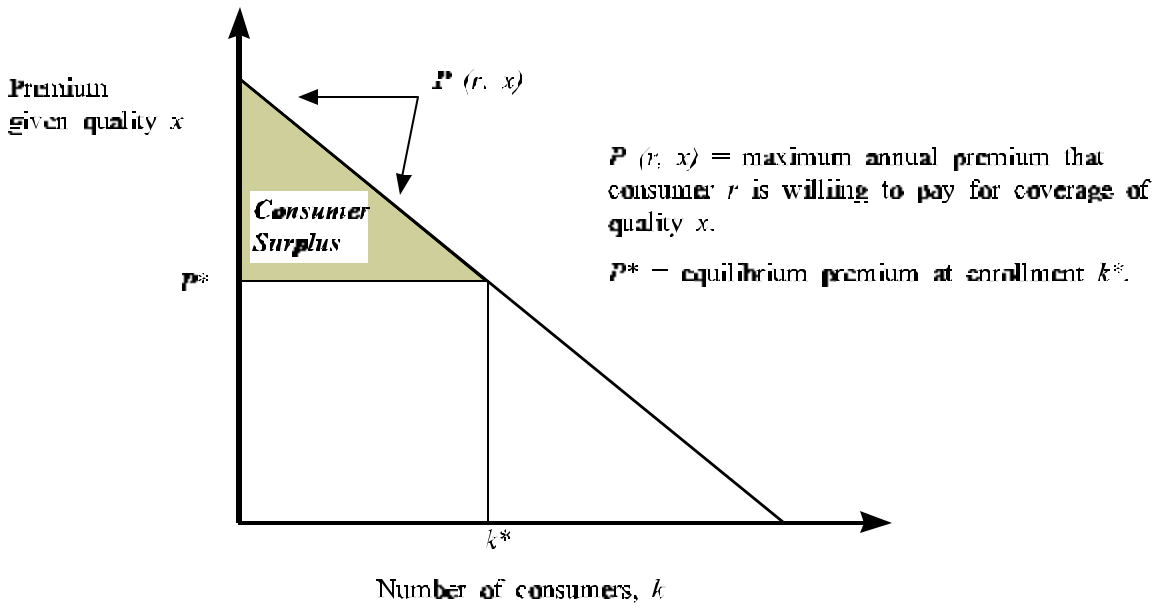

Figure 1. Inverse demand curve facing HMO

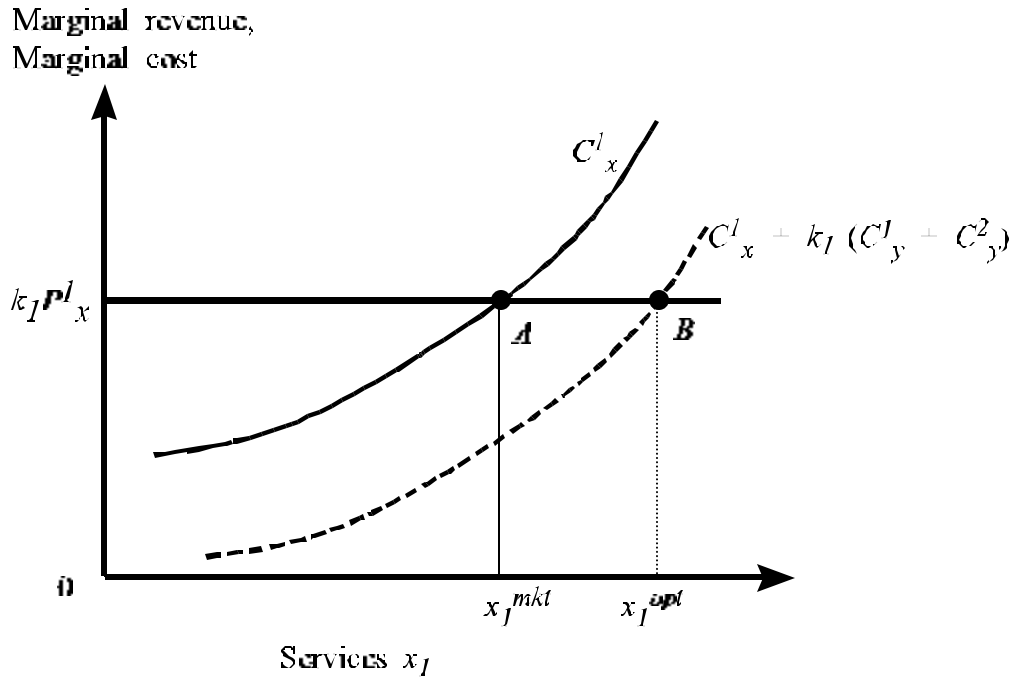

Figure 2. Externality model of HMO with random disenrollment