

COMPETITION AMONG HEALTH MAINTENANCE ORGANIZATIONS

WILLIAM E. ENCINOSA, III

*University of Michigan
Ann Arbor, MI 48109*

DAVID E. M. SAPPINGTON

*University of Florida
Gainesville, FL 32611*

We develop a model of competition among health maintenance organizations (HMOs) to analyze the effects of market power, scale economies, and asymmetric knowledge of health risk on market outcomes. We find that competition among HMOs may, but need not, ensure socially preferred outcomes. Market power or scale economies can sometimes admit socially preferred outcomes when they would otherwise not arise. Asymmetric knowledge of health risk may or may not be constraining. When it is constraining, a variety of patterns of incomplete health insurance can arise, along with excessive or insufficient treatment and preventive care for either high-risk or low-risk individuals.

1. INTRODUCTION

Health maintenance organizations (HMOs) are now firmly established as major providers of both health insurance and health care services.¹ Therefore, the nature and consequences of market competition among HMOs warrant careful investigation.

We analyze a simple model of HMO competition in this paper. Our model admits a preliminary assessment of how market power may influence outcomes in health-care markets. Such an assessment

We are grateful for the helpful comments and suggestions provided by Jim Burgess, Jacques Cremer, Jon Hamilton, Richard Hirth, Albert Ma, Daniel Spulber, Ted Stefos, an anonymous coeditor and two anonymous referees, and participants of the Second Bi-Annual Conference on the Industrial Organization of Health Care, held in Boston in September 1995. We also thank Rajiv Sharma for valuable research assistance. Support from the U.S. Department of Veterans Affairs and the Robert Wood Johnson Foundation is gratefully acknowledged.

1. Winslow and Anders (1993) report that more than 41,000,000 Americans were enrolled in HMOs as of 1993.

is important because many experts ultimately anticipate only limited competition among a few large HMOs in many regions of the United States.² Our model also admits an analysis of the effects of scale economies on market outcomes. Such an analysis is important because HMO operation typically entails large set-up costs.³ We also examine how asymmetric knowledge of health risks can affect equilibrium outcomes in health-care markets. Such an examination is relevant because experts differ regarding the severity and consequences of such information asymmetries.⁴

This research focuses in part on the question of whether competition among HMOs will ensure that socially preferred outcomes are realized. While society may prefer that different consumers receive different bundles of health services and health insurance, it is not obvious whether competition will admit such differentiation, particularly if low-risk patients subsidize their high-risk counterparts under socially preferred arrangements. We find that while unmitigated competition among HMOs can preclude socially desired cross-subsidies in some settings, it does not necessarily do so. In particular, market power or scale economies can facilitate a congruence between socially preferred and market outcomes.

This research also focuses on the qualitative differences in socially preferred and market outcomes that emerge when patients have better knowledge of their health risks than do HMOs. This superior knowledge may arise because individual patients are better informed about family health histories, recurring symptoms, or personal habits that affect the likelihood of illness. We find that asymmetric knowledge of health risks can, but need not, have important effects both on socially preferred outcomes and on market outcomes. For instance, copayments may arise that impose undesirable income variation on risk-averse patients. Excessive or unduly small levels of preventive care and treatment may also be supplied to limit incentives for misrepresentation of health risks. Whether such distortions arise depend on the nature of patients' preferences regarding the preventive care and treatment provided by HMOs, and on the market power and operating technologies of competing HMOs.

Our analysis proceeds as follows. Section 2 presents the basic elements of the models we analyze. Sections 3 and 4 isolate the effects of

2. See Anders and Winslow (1995) and Weil (1995), for example.

3. The American Medical Association estimates that startup costs for an HMO may exceed \$15,000,000. (See Winslow, 1995.)

4. See Pauly and Langwell (1983). Also see Marquis (1992), Sloan (1992), and Van de Ven and Van Vliet (1995) for a sampling of the different views on this matter.

market power and scale economies by characterizing socially preferred outcomes and market equilibria in the absence of information asymmetries. The effects of asymmetric knowledge of health risks on socially preferred outcomes and on market equilibria are explored in Sections 5 and 6. Directions for further research are outlined in Section 7. Proofs of selected formal results are provided in the Appendix.⁵

Before proceeding, we note that while there are relatively few formal models of HMO competition in the literature, many authors cite the classic model, due to Rothschild and Stiglitz (RS) (1976), of adverse selection in the insurance industry when contemplating the likely effects of competition in the health-care industry.⁶ Because insurance is an important element of the array of services supplied by HMOs, many of the insights offered by RS are relevant to an analysis of HMO competition. However, the central findings of RS may not arise in models of HMO interaction, for three distinct reasons. First, insurance is just one of many products supplied by HMOs. Second, the HMO market may ultimately be characterized by a concentration of market power, which RS did not consider. Third, the technologies of production differ in HMO and insurance markets. Because of these differences, our analysis in Section 6 displays a fundamental departure from RS's main conclusions—that competition among insurance providers under asymmetric information will preclude a single insurance policy for all consumers and will prevent low-risk consumers from obtaining full insurance. We find, in contrast, that if HMO market power or scale economies are sufficiently pronounced, all consumers may receive identical health plans and full insurance in equilibrium.^{7,8}

5. Omitted proofs can be found in Encinosa and Sappington (ES) (1995), which is available from the authors upon request.

6. See, for example, Marquis (1992), Newhouse (1984, 1996), Pauly and Langwell (1983), and Van de Ven and Van Vliet (1995). Pauly (1986) warns against employing conclusions from the theoretical literature on insurance markets to obtain predictions about likely outcomes in health-care markets.

7. More generally, all patients do not necessarily receive the same health plan, but the premium charged to a patient often does not reflect fully the expected cost of treating the patient. This finding contrasts with Pauly's (1984) observation that competition should be expected to result in premiums that are "tailored to risk" (p. 90). Sloan (1992) and Pauly (1986) report limited historical tailoring of premiums to risk in practice.

8. Newhouse (1996) provides a related explanation for the limited tailoring of health plans to health risks that is observed in practice. Newhouse shows that limited tailoring will arise in the RS framework when the transaction costs of writing and implementing multiple insurance contracts are sufficiently pronounced. We present conditions under which the equilibrium in which all patients receive the same efficient health plan is the unique equilibrium, even when the transactions costs of writing multiple contracts is arbitrarily small.

2. ELEMENTS OF THE MODEL

We begin by describing the basic elements of the models that we analyze. There are two types of risk-averse patients, denoted 1 and 2. All patients have initial income $y > 0$, and value final wealth (Y) according to the strictly increasing, strictly concave function $u(Y)$. Patients differ according to their health risk. There are $n_1 > 0$ high-risk patients and $n_2 > 0$ low-risk patients. The high-risk patients, who will often be referred to as type-1 patients, are more likely to become ill than their low-risk (type-2) counterparts, *ceteris paribus*. Formally, $0 < q_1(e) < q_2(e) < 1$ for all $e \geq 0$, where $1 - q_i(e)$ is the probability that a type- i patient becomes ill after receiving preventive care e . Preventive care increases at a decreasing rate the probability that the patient is well. Formally, $q'_i(e) \geq 0$ and $q''_i(e) \leq 0$ for all $e \geq 0$, where primes denote derivatives and where the inequalities are strict unless otherwise noted.⁹ In the ensuing analysis, e_i will denote the amount of preventive care received by a type- i patient. The marginal cost of delivering preventive care is normalized to unity. In practice, preventive care may include immunizations, routine examinations, and diet counseling.

When a type- i patient becomes ill, his utility from health drops from \bar{H}_i to \underline{H}_i , where $\bar{H}_i > \underline{H}_i$. Treatment can help improve the patient's health, and thereby increase his utility from health. We will denote by t_i the amount of treatment the type- i patient receives when he is ill.¹⁰ Treatment may take the form of prescribed medication or surgical procedures, for example. Treatment is delivered with constant marginal cost, $c > 0$, and is assumed to increase the sick patient's utility from health at a decreasing rate. Formally, $H'_i(t_i) \geq 0$ and $H''_i(t_i) \leq 0$ for all $t_i \geq 0$ for $i = 1, 2$, where the inequalities are strict unless otherwise noted, and where $\underline{H}_i + H_i(t_i)$ is the type i 's utility from health when he is sick and receives treatment t_i .¹¹ The patient's utility from health is always lower when he is sick than when he is well. Formally, $\bar{H}_i > \underline{H}_i + H_i(t_i)$ for all $t_i \geq 0$. For simplicity, patients' utility functions, $U_i(\cdot)$, are assumed to be separable in wealth (Y) and health (H), i.e., $U_i(Y, H_i) = u(Y) + H_i$. Furthermore, unless otherwise noted, we assume throughout that $u'(0) > (1/c)H'_i(0) > u'(y)$ for $i = 1, 2$. In words, the sick patient's marginal utility of income exceeds his marginal utility of expenditures on health care when his wealth is sufficiently small,

9. Unless otherwise noted, $q'_1(0)$ and $q'_2(0)$ are assumed to be sufficiently large that a strictly positive amount of preventive care is delivered to all patients in equilibrium.

10. The patient's health status—sick or well—is assumed to be observable and contractible.

11. We assume that $H_i(0) = 0$ for $i = 1, 2$.

whereas this relative valuation is reversed when his wealth is sufficiently large. This assumption helps ensure an interior solution to the problems we analyze.

Patients may be charged different amounts according to their health status (and thereby according to the amount of preventive care and treatment they receive). We will denote by p_i^w the total payment the type- i patient makes when he is well, and by p_i^s the total payment he makes when he is sick.¹² Patients select health-care plans to maximize their expected welfare. The expected welfare of a type- i patient who secures a health plan T_i is denoted $W_i(T_i)$:

$$W_i(T_i) \equiv q_i(e_i)[u(y - p_i^w) + \bar{H}_i] + [1 - q_i(e_i)][u(y - p_i^s) + \underline{H}_i + H_i(t_i)], \quad (1)$$

where $T_i \equiv (p_i^w, p_i^s, t_i, e_i)$. A patient's expected welfare is simply his utility when well and his utility when sick, weighted by the probability that each of these states of health occurs. A type- i individual who does not purchase a health-care plan obtains expected welfare $\underline{W}_i \equiv u(y) + q_i(0)\bar{H}_i + [1 - q_i(0)]\underline{H}_i$.

The (extranormal) profit a health-care firm (i.e., an HMO) earns depends on the services it supplies and the fees it imposes. $\pi_i(T_i)$ will denote the unit profit the HMO anticipates on a type- i patient when the patient selects health plan $T_i \equiv (p_i^w, p_i^s, t_i, e_i)$. Formally,

$$\pi_i(T_i) \equiv q_i(e_i)p_i^w + [1 - q_i(e_i)][p_i^s - ct_i] - e_i. \quad (2)$$

An HMO's overall expected profit depends on its unit expected profit on each patient type that is served, on the number of such patients, and on the HMO's fixed costs of production.

We will explore the interaction between two HMOs, labeled A and B . For simplicity, we assume that patients perceive the treatment and care delivered by the two HMOs to be identical. Furthermore, both HMOs are able to credibly commit to deliver all of the treatment and care that they promise to patients.¹³ The two firms also have the same variable operating costs. The firms differ only in their fixed costs of production. Firm A 's fixed costs of production ($F^A \geq 0$) are no larger than those of firm B (F^B). This advantage for firm A may stem from an incumbency advantage (which reduces set-up costs or advertising

12. One could view p_i^w as an initial premium for health insurance and preventive care, and $p_i^s - p_i^w$ as a subsequent copayment for treatment.

13. This assumption may be justified by reputational concerns in a fully intertemporal analysis.

costs), for example. For convenience, we also assume that firm B 's fixed costs of production are always less than the profit a monopoly firm could earn if it had no fixed costs and faced no potential entry.¹⁴

3. THE COMPLETE INFORMATION SOCIAL OPTIMUM

First consider the outcome a social planner would dictate if she had perfect knowledge of patient preferences and production technologies, and for simplicity, if she valued the welfare of each patient equally.¹⁵ The planner would choose health plans to maximize patients' expected welfare, while ensuring nonnegative profit for the least-cost producer. Formally, the planner's problem, [SP], is

$$\text{Maximize}_{T_1, T_2} \sum_{i=1}^2 n_i W_i(T_i) \quad \text{subject to} \quad \sum_{i=1}^n n_i \pi_i(T_i) - F^A \geq 0.$$

Lemma 1 summarizes the key features of the ideal outcome from the social perspective.

LEMMA 1: *At the solution to [SP]:*

- (i) *full income insurance is provided, so $p_i^w = p_i^s \equiv p_i$ for $i = 1, 2$;*
- (ii) *all patients are charged the same premium, so $p_1 = p_2$;*
- (iii) *the efficient level of treatment ($t_i^*(p_i)$) is provided to all patients who become ill, so $u'(y - p_i) = (1/c) H'_i(t_i)$;*
- (iv) *the efficient level of preventive care ($e^*(p_i, t_i)$) is delivered to all patients, so*

$$\frac{\partial W_i(\cdot)/\partial e_i}{\partial W_i(\cdot)/\partial p_i} = \frac{\partial \pi_i(\cdot)}{\partial e_i}, \quad \text{or} \quad - \frac{q'_i(e_i) \Delta H_i(T_i)}{u'(y - p_i)}$$

$$= q'_i(e_i) c t_i - 1,$$

where $\Delta H_i(T_i) = \bar{H}_i - [H_i + H_i(t_i)]$; and

- (v) *the supplier earns zero profit.*

The conclusions reported in Lemma 1 are all quite intuitive. Co-payments ($p_i^w \neq p_i^s$) are not imposed because patients are averse to

14. Formally, we assume that $F^B < \pi^m \equiv n_1 \pi_1(T_1^m) + n_2 \pi_2(T_2^m)$, where $T^m \equiv (T_1^m, T_2^m)$ is the solution to the monopoly problem [M], defined as $\{\text{Maximize}_T \sum_{i=1}^2 n_i \pi_i(T_i) \text{ subject to } W_i(T_i) \geq \underline{W}_i \text{ for } i = 1, 2\}$. This assumption helps ensure that in the equilibria analyzed below, no firm ever earns more profit than it would earn in the hypothetical monopoly setting.

15. A more general case is examined in ES (1995).

income risk.¹⁶ All patients are charged the same premium because the welfare of all patients is valued equally. Treatment is delivered to all sick patients up to the point where the extra utility that treatment affords a patient per dollar of expenditure is equal to the patient's marginal utility of income. Similarly, preventive care is provided up to the point where its net marginal cost to the supplier is equal to its net marginal benefit per dollar to the patient.¹⁷

One implication of Lemma 1 is recorded in Corollary 1, which refers to Assumptions 1 and 2. Assumption 1 states that low-risk and high-risk patients have identical preferences regarding health status and treatment. Assumption 2 states that preventive care has the same marginal effect on all patients. Together, these assumptions ensure that high-risk patients are more costly to serve than low-risk patients, and so are subsidized by low-risk patients at the social optimum.

ASSUMPTION 1: $\underline{H}_1 = \underline{H}_2; \bar{H}_1 = \bar{H}_2; \text{ and } H_1(t) = H_2(t) \text{ for all } t \geq 0.$

ASSUMPTION 2: $q'_1(e) = q'_2(e) \text{ for all } e \geq 0.$

COROLLARY 1: *Suppose Assumptions 1 and 2 hold. Then $p_1 = p_2, t_1 = t_2,$ and $e_1 = e_2,$ and so $\pi_1(T_1) < \pi_2(T_2)$ at the solution to [SP].*

4. DUOPOLY COMPETITION WITH COMPLETE INFORMATION

We now examine the outcome of unfettered competition between two potential suppliers of health services when the suppliers can costlessly observe each patient's type and are free to offer different health plans to different patients.

Formally, we analyze a duopoly game in which the two firms initially propose health plans simultaneously. All patients then observe the proposed plans and decide which plan (if any) to accept. Each

16. Copayments could be optimal under alternative specifications of patient preferences. If, for example, a patient's utility from wealth Y and health H were $U(Y + H)$ where $U'(\cdot) > 0$ and $U''(\cdot) < 0$, then by setting $p_i^* < p_i^0$, the social planner could help insure the patient against utility variation caused by variations in health.

17. The supplier's net marginal cost of delivering e_i is the difference between the direct marginal cost, 1, and the expected cost savings from reduced treatment costs, $q'_i(e_i)ct_i$. The patient's net marginal benefit from e_i is his marginal rate of substitution of p_i for e_i , which is the ratio of his marginal expected utility from increased e [$q'_i(e_i) \Delta H_i$] to his marginal utility of income [$u'(y - p_i)$]. In the ensuing discussion, we will denote by $t_i^*(p_i)$ and $e_i^*(p_i, t_i)$, respectively, (conditionally) efficient treatment and care levels for type- i patients.

patient accepts at most one plan, and randomizes between plans that offer identical levels of expected utility. $r_j \in [0, 1]$ will denote the fraction of patients that, when indifferent between plans, select firm j 's plan (where $j = A, B$ and $r_A + r_B = 1$). After observing the patients' choices, the firms decide simultaneously whether to enter the market. If a firm enters, it serves all patients that selected its plan. If only one firm enters, that firm can (and will) also serve all patients that initially selected the rival's plan if doing so does not reduce the profits of the entering firm. A firm sinks its fixed costs of production upon entry. We assume that each firm enters the market when it is indifferent between entering and not entering. Our focus is on characterizing the (pure-strategy) Nash equilibria of this complete information duopoly game.¹⁸

Proposition 1 summarizes the effect of competition in the absence of scale economies.

PROPOSITION 1: *Suppose $F^A = F^B = 0$. Then there is a unique (pure-strategy) equilibrium of the complete-information duopoly game in which the two firms offer the same health plan, and firm j serves $r_j[n_1 + n_2]$ patients ($j = A, B$). Each firm makes zero unit expected profit on each patient. Patients face no copayments (so $p_i^w = p_i^s \equiv p_i$ for $i = 1, 2$) and receive efficient treatment and care levels given the premiums they are charged (so $t_i = t_i^*(p_i)$ and $e_i = e_i^*(p_i, t_i)$). Furthermore, if Assumptions 1 and 2 hold, then, relative to the social optimum: (i) the low-risk patients pay a smaller premium and receive more treatment and more preventive care; and (ii) the high-risk patients pay a larger premium and receive less treatment and less preventive care.*

Proposition 1 reports that when the competitors are identical, when patient characteristics are observable and contractible, and when there are no fixed costs of production, competition eliminates cross subsidies, even those that are socially desirable. In other words, competition forces premiums to reflect costs of service. Competition also ensures that the expected welfare of each patient type is maximized, subject to the restriction that the firm must earn exactly zero expected

18. This game and its Nash equilibria are described more formally in ES (1995). Notice that in this game, the two firms compete for patients in Bertrand-like fashion before they incur any sunk costs. Such competition may be representative of situations in which rival HMOs bid for the right to serve the employees of a large corporation, for example. [See Che and Gale (1997) for a useful analysis of how a buyer alliance may optimally structure competition among HMOs.]

profit on each type. Therefore, efficient treatment and care levels are delivered once premiums are (risk-)adjusted to reflect expected costs.¹⁹

In the presence of market power or scale economies, competition need not eliminate cross subsidies. The Bertrand-like competition that we analyze does restrict firm A 's profit to the magnitude of its fixed-cost advantage ($F^B - F^A$), but it does not preclude cross subsidies between patient types. When $F^B > F^A$, any allocation of fixed costs among patients that generates profit $F^B - F^A$ for firm A and precludes firm B from profitably serving any subset of patients can arise in equilibrium. Similarly, when the two HMOs have the same strictly positive level of fixed costs, there are many ways that these fixed costs can be reflected in premiums while dissipating all rents and ensuring that no supplier can profitably serve a subset of patients. These facts are reflected in Proposition 2.

PROPOSITION 2: *When $F^B > F^A$, there exist a continuum of pure-strategy Nash equilibria in the complete-information duopoly game in which firm A serves all patients. If $F^B = F^A > 0$, two such continua of equilibria exist, one in which firm A serves all patients and the other in which firm B serves all patients. At every equilibrium: (i) the firm that serves all patients earns profit $F^B - F^A$; (ii) no copayments are levied (so $p_i^w = p_i^s = p_i$ for $i = 1, 2$); and (iii) efficient treatment and care levels are delivered (so $t_i = t_i^*(p_i)$ and $e_i = e_i^*(t_i, p_i)$). At equilibrium $s \in [0, 1]$, the health plan selected by type- i patients is the solution to the problem $[W_i(s)]$, defined as $\{\text{Maximize}_{T_i} W_i(T_i) \text{ subject to: (1) } n_i \pi_i(T_i) = f_i(s)F^B; \text{ and (2) } W_i(T_i) \geq \underline{W}_i\}$, where $f_2(s) = s$ and $f_1(s) = 1 - s$. Furthermore, at those equilibria where the producer earns a higher (respectively, a lower) unit expected profit $\pi_i(\cdot)$ on type- i patients than at the social optimum, these patients are charged a higher (respectively, a lower) premium and receive less (respectively, more) treatment and care than at the social optimum.*

Proposition 2 reports that competition ensures efficient (although not necessarily socially optimal) outcomes even in the presence of market power or scale economies. Duplicative fixed costs are avoided, since a single supplier serves all patients. Furthermore, once the fraction of firm B 's fixed costs that will be recovered from the unit profit generated by serving each type of patient is determined, patient welfare is maximized, subject to ensuring the specified unit profit from each patient

19. Higher premiums increase the marginal utility of income and therefore reduce efficient care and treatment levels, *ceteris paribus*. Consequently, when competition imposes higher premiums on high-risk patients, they receive less preventive care and less treatment when ill.

type. Firm *A* is indifferent among the many ways to allocate fixed costs across patients in our model, so continua of equilibria can arise.

The flexibility that competing firms have in assigning fixed costs of production across patient types admits the possibility that both types of patients will receive the same health plan in equilibrium. This possibility arises, for example, when: (i) patients' preferences are sufficiently similar (so Assumptions 1 and 2 hold); and (ii) firm *B*'s fixed costs are sufficiently large that firm *B* cannot profitably attract only the low-risk patients from firm *A* when firm *A* offers to all patients the health plan that maximizes the expected welfare of high-risk patients while generating zero expected profit on these patients. Condition (ii), along with a convenient upper bound on F^B , is stated formally in the Appendix as Assumptions 3–5.

PROPOSITION 3: *Suppose assumptions 1–5 hold. Then there is a (unique) pooling Nash equilibrium of the complete-information duopoly game at which all patients receive the same health plan (so $T_1 = T_2$). This health plan entails no copayments (so $p_i^w = p_i^s \equiv p_i$) and provides efficient treatment and care levels (so $t_i = t_i^*(p_i^s)$ and $e_i = e_i^*(p_i, t_i)$). Only one firm serves all patients, and that firm earns profit $F^B - F^A$ at this equilibrium.*

To interpret Proposition 3, suppose firm *A* offered to all patients the same (efficient) health plan that generates expected profit $F^B - F^A$ for firm *A*. Since low-risk patients are less costly to serve than high-risk patients, this uniform health plan would generate more than $(1/n)$ ($F^B - F^A$) in profit on each low-risk patient. If firm *B*'s fixed costs of production (F^B) were sufficiently small, firm *B* could profitably attract just the low-risk patients with a health plan that they prefer to the single plan offered to all patients by firm *A*. In contrast, when F^B is sufficiently large, firm *B* will be unable to profitably attract only the low-risk patients, because its average cost of serving just these patients will be too large. Most importantly, as F^B increases, the reduced scale economies that firm *B* suffers when it serves only the low-risk patients eventually outweigh the cost savings from serving a lower-cost (but smaller) population. Consequently, when F^B is sufficiently large, firm *B* will not be able to upset the pooling equilibrium identified in Proposition 3, which we will refer to as the s^p equilibrium.

As we show in Section 6, this (s^p) equilibrium with no risk adjusting (so all patients receive the same health plan despite differences in their expected treatment costs) may persist even when patients are privately informed about their health risk. Furthermore, the s^p equilibrium will be the unique equilibrium under the identified assumptions if firms incur strictly positive (but even arbitrarily small) transactions

costs when they offer more than a single health plan.²⁰ Absent transactions costs, a continuum of equilibria characterized by partial risk adjusting can arise in which high-risk patients pay more than low-risk patients for the health insurance they receive, but the premium differential is less than the corresponding difference in expected treatment costs. The central observation here is that scale economies can provide HMOs with leeway to implement little or no risk adjusting, even in competitive markets.²¹

5. THE SOCIAL OPTIMUM WITH ADVERSE SELECTION

Now consider the adverse-selection setting, where patients have better knowledge of their health risk than do health-care suppliers (perhaps because of superior knowledge of past health problems or personal activities that affect one's health, for example).²² Formally, each patient alone knows his health risk (i.e., his type). The two suppliers and the social planner know that there are n_i patients of type i , but cannot discern any individual patient's type.²³

The socially preferred outcomes in this adverse-selection setting can be characterized by solving the following problem of the social planner, denoted [SP-A]:

$$\text{Maximize}_{T_1, T_2} \sum_{i=1}^2 n_i W_i(T_i) \tag{3}$$

subject to

$$\sum_{i=1}^2 n_i \pi_i(T_i) - F^A \geq 0 \tag{4}$$

and

$$W_i(T_i) \geq W_i(T_k) \quad \text{for } k \neq i, \quad \text{for all } i, k = 1, 2. \tag{5}$$

20. The uniqueness holds because all other equilibrium configurations entail greater total cost. Thus, Newhouse's (1996) insight extends in this setting to the case of arbitrarily small transactions costs.

21. High-risk patients may even be afforded more treatment, more preventive care, and a lower premium than their low-risk counterparts. This possibility arises when Assumptions 1 and 2 hold, and Assumptions 3–5 hold as strict inequalities. See ES (1995) for details.

22. Marquis (1992) provides evidence of adverse selection resulting from asymmetric knowledge of health risks in practice.

23. HMOs can often identify directly and discriminate against severe health risks in practice. See Light (1992), Sloan (1992), and Van de Ven and Van Vliet (1995), for example.

The social planner again seeks to maximize patients' expected welfare while ensuring nonnegative profit using the least-cost technology. The self-selection constraints (5) indicate that in this adverse selection setting, patients must be induced, not simply ordered, to select the treatment plan that the social planner would like them to select.

The self-selection constraints do not always constrain the social planner. For instance, when all patients value treatment and preventive care in the same manner (so that Assumptions 1 and 2 hold), the social optimum will entail identical premiums, preventive care, and treatment for all patients. When the social planner offers only this single ideal health plan to all patients, adverse-selection concerns do not arise.

The social planner may also be unconstrained by her limited information about patient type when the two types of patients value treatment and care very differently. To illustrate, Lemma 2 considers an extreme case where preventive care reduces the likelihood of illness only for the low-risk patients, while treatment only improves the health of high-risk patients who are ill.²⁴ In this setting, neither type of patient values the service provided to the other type of patient, and so neither has an incentive to purchase the health plan intended for his counterpart.

LEMMA 2: *Suppose $0 = H'_2(t) < H'_1(t)$ for all $t \geq 0$ and $0 = q'_1(e) < q'_2(e)$ for all $e \geq 0$. Then at the solution to [SP-A]:*

- (i) *there are no copayments, and each patient is charged the same premium (so $p_1 = p_2$); and*
- (ii) *all patients receive the treatment and care that is efficient for them (so $t_1 = t_1^*(p_1) > 0 = t_2^*(p_2) = t_2$ and $e_2 = e_2^*(p_2, t_2) > 0 = e_1^*(p_1, t_1) = e_1$).*

More generally, private knowledge of health risk may constrain the social planner, and inefficient pricing, treatment, and/or care may result. To provide just one illustration, suppose high-risk patients become sick more often and suffer a greater loss in utility when they do become sick, but are otherwise identical to low-risk patients. In this setting, the social planner would like to charge all patients the same premium but provide more preventive care to high-risk patients. However, since low-risk patients value preventive care, they will prefer the health plan intended for high-risk patients. To deliver additional preventive care solely to high-risk patients, the social planner raises

24. These extreme assumptions are useful, but not necessary, to make the more general point that the self-selection constraints need not bind at the solution to [SP-A] when patients' valuations of treatment and care are sufficiently different.

the expected price of the increased care, and charges the patient more when he is well than when he is sick (through a partial rebate of the initial premium when the patient becomes ill, for example). This payment arrangement is particularly unattractive to the low-risk patient, who is less likely to become ill and so is relatively unlikely to receive the partial premium rebate. The social planner also provides more than the efficient level of care to high-risk patients. Doing so renders the health plan intended for high-risk patients differentially attractive to them, since they value the extra care more than their low-risk counterparts. These conclusions are summarized in Lemma 3.²⁵

LEMMA 3: *Suppose $q'_1(e) = q'_2(e)$ for all $e \geq 0$, $H'_1(t) = H'_2(t)$ for all $t \geq 0$, and $\bar{H}_1 - \underline{H}_1 > \bar{H}_2 - \underline{H}_2$. Then at the solution to [SP-A]:*

- (i) *high-risk patients always pay more for health care than low-risk patients, and pay the most when they are well (so $p_1^w > p_1^s > p_2^w = p_2^s \equiv p_2$);*
- (ii) *the efficient level of treatment is provided to all sick patients, so low-risk patients receive less treatment (i.e., $t_1 = t_1^*(p_1^s) < t_2^*(p_2) = t_2$); and*
- (iii) *high-risk patients receive more than the efficient level of preventative care, while low-risk patients receive the efficient level of care (so $e_1 > e_1^*(p_1^w, p_1^s, t_1)$ and $e_2 = e_2^*(p_2, t_2)$).*

6. DUOPOLY COMPETITION WITH ADVERSE SELECTION

Having illustrated some of the many new effects that private knowledge of health risk can introduce even in the social optimum, we turn now to the question of whether competition among health-care providers in the presence of adverse selection will give rise to socially optimal outcomes. We model competition here exactly as in the complete-information setting. To review briefly, the two potential producers initially propose health plans simultaneously. Each patient then selects at most one plan. After observing the patients' choices, the firms decide simultaneously whether to enter. If a firm enters, it incurs its fixed costs of production and delivers the health plan it promised. We refer to this interaction when each patient's risk type is known only to that patient

25. Since all patients value treatment (t) identically on the margin, efficient treatment levels are delivered to all patients. However, since the high-risk patients pay more for their plan, they have a higher marginal utility of income when sick. Consequently, as Lemma 3 reports, the efficient treatment level is smaller for the high-risk patients. Other qualitative distortions that can arise in socially preferred outcomes in the adverse-selection setting are analyzed in ES (1995).

as the asymmetric-information duopoly game. Our focus is on (pure-strategy) Nash equilibria of this game.

Whether competition among HMOs will ensure socially optimal outcomes depends upon patient characteristics and production technologies. If the two potential producers have no fixed costs of production, and if the patients' valuations of the multiple services offered by the firms are sufficiently distinct (e.g., if only high-risk patients benefit from treatment and only low-risk patients value preventive care), then competition will ensure efficient treatment and care levels and the absence of copayments, but will not ensure the socially preferred premium levels.

PROPOSITION 4: *Suppose (i) $F^A = F^B = 0$, (ii) $0 = H'_2(t) < H'_1(t)$ for all $t \geq 0$, and (iii) $0 = q'_1(e) < q'_2(e)$ for all $e \geq 0$. Then there is a unique (pure-strategy) equilibrium of the asymmetric-information duopoly game in which firm j serves $r_j[n_1 + n_2]$ patients and makes zero unit expected profit on each patient. No copayments are charged ($p_i^w = p_i^s \equiv p_i$ for $i = 1, 2$), efficient treatment levels are provided ($t_1 = t_1^*(p_1) > t_2^*(\cdot) = 0 = t_2$), and efficient preventive care is delivered ($e_2 = e_2^*(p_2, t_2) > e_1^*(\cdot) = 0 = e_1$).*

Proposition 4, like Proposition 1, emphasizes that in the absence of scale economies, competition among identical suppliers will drive premiums to reflect costs, and thereby eliminate cross subsidies. Proposition 4 also points out that when patients' valuations of different health services are sufficiently distinct, the same (efficient, but not socially optimal) outcomes that arise under complete information also arise under asymmetric information. The information asymmetry is of no consequence here because the patients' preferences are so diverse that high-risk patients do not find it attractive to misrepresent their innate characteristics in order to secure the health plan intended for low-risk patients.

When patients' valuations are more congruent, however, private knowledge of health risk can cause the market outcome to diverge from the socially optimal outcome in additional ways. To illustrate, recall that under Assumptions 1 and 2 all patients value treatment in identical fashion, and preventive care affects each patient's likelihood of illness in the same manner. These assumptions ensure that high-risk patients are systematically more costly to serve than low-risk patients, just as in RS's model of the insurance industry where some consumers are systematically more prone to accidents than others.²⁶ Consequently, as RS have shown, if there are no fixed costs of production, a (pooling)

26. This congruity is demonstrated more precisely in ES (1995).

equilibrium in which all patients purchase the same health plan cannot arise. An identical plan for all patients that yields zero expected profit generates positive rent on the low-risk patients. Consequently, a competitor can profitably return some of this rent to low-risk patients (without attracting high-risk patients), and thereby eliminate the zero-profit pooling equilibrium.

This is not necessarily the case in the presence of scale economies, however. A firm that attempts to serve only the low-risk patients faces higher average fixed costs of production than another firm that serves all patients. If this cost disadvantage is sufficiently pronounced, it may be impossible for a firm to profitably attract only the low-risk patients when a competing firm offers the same health plan to all patients. Consequently, a pooling equilibrium with efficient pricing and treatment may exist even when patients' characteristics are structured to parallel the characteristics that RS examine. Furthermore, if the two firms are identical so that competition eliminates all profit, the single health plan that is offered will be the socially optimal plan. These conclusions are recorded formally in Proposition 5.

PROPOSITION 5: *Suppose Assumptions 1–5 hold. Then the s^p equilibrium of the complete-information duopoly game is an equilibrium of the asymmetric-information duopoly game. At this equilibrium, all patients receive the same health plan. The plan entails no copayments and provides efficient treatment and care levels (so $p_i^w = p_i^s \equiv p_i$, $t_i = t_i^*(p_i^s)$, and $e_i = e_i^*(p_i, t_i)$ for $i = 1, 2$). If $F^A = F^B$, this plan is the socially optimal plan. If $F^B > F^A$, this plan imposes a higher premium and provides less treatment than the socially optimal plan.*

Proposition 5 indicates that competition and adverse selection need not result in different health plans for different types of patients, or in inefficient levels of treatment and care. In fact, if there are any transactions costs (however small) associated with offering different health plans (recall Newhouse, 1996) and if Assumptions 1–5 hold, then the s^p equilibrium with no risk adjusting will be the unique equilibrium of the asymmetric-information duopoly game.

Absent transactions costs, other equilibria arise in which patients with different health risks select different health plans, and these plans may not ensure efficient treatment. Proposition 6 illustrates this point. The statement of the proposition refers to T^s , which is the solution to problem [WA-s], defined as: $\{\text{Maximize}_{T_1, T_2} n_1 W_1(T_1) + n_2 W_2(T_2)$ subject to, for $i, k = 1, 2$: (1) $W_i(T_i) \geq W_i(T_k)$ for $k \neq i$; (2) $n_i \pi_i(T_i) = f_i(s)F^B$, where $f_2(s) = s \in [0, 1]$ and $f_1(s) = 1 - s$; and (3) $W_i(T_i) \geq \underline{W}_i\}$. The statement also refers to health plans (T_1^1 and T_2^m) that are defined in Assumptions 3 and 4 in the Appendix. Proposition 6 also

refers to assumption 6, which serves to simplify the model by ruling out copayments and ensuring that preventive care is never provided in equilibrium (since it has no effect on patient health).

ASSUMPTION 6: $q'_1(e) = q'_2(e) = 0$ for all $e \geq 0$, and copayments are not permitted.

PROPOSITION 6: Suppose $F^A = F^B > 0$, and Assumptions 1, 5, and 6 hold. Then for any $n_1 > 0$, there exists an $n_2 > 0$ such that for any $n_2 \geq n_2$ and $n_2 \pi_2(T_1^1) \leq F^B \leq n_2 \pi_2(T_2^m)$, T^1 is a separating Nash equilibrium of the asymmetric-information duopoly game in which one firm serves all patients and makes zero profit. At this equilibrium, the low-risk patients select a health plan that provides efficient treatment (so $t_2 = t_2^*(p_2)$). The high-risk patients select a health plan with a higher premium (so $p_1 > p_2$) and more than the efficient level of treatment (i.e., $t_1 > t_1^*(p_1)$).

Proposition 6 illustrates that even when the two potential suppliers have identical technologies and even when the only relevant differences among patients are those analyzed by RS, distortions that are qualitatively different from those identified by RS can emerge in equilibrium.²⁷ If, for reasons outside of our model, the supplier decides to charge low-risk patients a premium that reflects both the variable costs of serving them and all fixed costs of production, the low-risk patients may prefer a health plan intended for high-risk patients, whose premium does not incorporate any fixed costs. To prevent low-risk patients from selecting the health plan intended for high-risk patients, the supplier can increase the level of treatment promised to high-risk patients (even beyond its efficient level) and raise the premium accordingly. Because high-risk patients are more likely to become ill, they value a promise of increased treatment when ill more highly than do low-risk patients. Consequently, providing an inefficiently large level of treatment to high-risk patients (and charging them accordingly) can best discourage low-risk patients from selecting the health plan intended for high-risk patients.

Most importantly, a supplier's flexibility in assigning fixed costs of production across patient types while precluding profitable operation by a rival can cause self-selection constraints to bind in different ways than they bind in RS's analysis. The result can be distortions that differ qualitatively from those identified by RS.

In concluding, we emphasize that although the fixed costs and

27. This result may be surprising in light of the fact that Stiglitz (1977) identifies the same qualitative distortions in a monopoly setting (where scale economies are present) that RS identify in a competitive setting with no scale economies.

multiple services in our model admit conclusions that differ qualitatively from those drawn by RS, important similarities can also emerge. As in RS, there may be no Nash equilibrium in our asymmetric information duopoly model if fixed costs of production and the proportion of high-risk patients are both sufficiently small.²⁸ The distortions identified by RS can also emerge if the incumbent assigns relatively few of the fixed costs of production to the low-risk patients. Therefore, the useful insights provided by RS should not be ignored. Instead, the importance of the distinguishing features of the RS model should be recognized before the model is employed to predict the outcome of competition among HMOs.

7. CONCLUSION

The primary purpose of this research was to develop a simple model of HMO competition, in order to analyze the effects of market power, scale economies, and asymmetric information in health-care markets. We found that competition among HMOs can, but need not, cause market outcomes to diverge from socially preferred outcomes. Market power or scale economies can provide HMOs with some leeway in designing profitable health plans. This leeway may be employed to limit the extent to which health-care premiums are adjusted to reflect patient risk. Asymmetric knowledge of health risk, coupled with market power or scale economies, can also result in incomplete insurance for either high-risk patients or low-risk patients. However, the rich set of policy instruments available to HMOs can sometimes enable them to render inconsequential information asymmetries that would be problematic in insurance markets.

The fact that a great variety of findings emerge from a relatively simple model of HMO interaction reinforces the obvious fact that the market for health services is an intricate and complex one that defies simple characterization. Significant additional research is clearly in order, particularly in the following four directions.

First, different patient characteristics should be considered. For instance, patients' preferences may not be separable in income and health. Furthermore, different patients have different income levels, and limited ability to pay for medical services or health insurance can

28. A Nash equilibrium fails to exist in RS's model only when n_2 is sufficiently larger than n_1 . In contrast, in the setting of Proposition 6, our model admits a separating Nash equilibrium whenever n_2 is sufficiently larger than n_1 , provided F^B is of appropriate magnitude.

cause additional discrepancies between socially preferred outcomes and market outcomes. Patients also generally differ according to the severity of their condition when they become ill.

Second, additional differences among HMOs should be explored. HMOs typically have different variable and fixed costs of production. Furthermore, HMOs often specialize in the delivery of particular health-care services. HMOs can also differ in the quality of the services they deliver, in their reputations, and in their knowledge of patients' health risks.^{29,30}

Third, alternative depictions of competition among HMOs warrant investigation. In contrast to the simplifying assumption of our model, it is generally prohibitively costly for HMOs to specify in advance all relevant aspects of health care that they will deliver.³¹ Furthermore, the possibility always exists that an HMO may renege on a promise about the service it will deliver. For these reasons and others, patients may be reluctant to switch HMOs, or to leave a fee-for-service arrangement to enroll in an HMO. A patient's choice of an HMO may also depend critically on the particular physicians associated with different HMOs.

Fourth, the many different ways in which governments intervene in health-care markets warrant careful attention.³²

APPENDIX

ASSUMPTION 3: $F^B \leq \hat{I} \equiv n_1 \pi_1(T_2^m) + n_2 \pi_2(T_2^m)$ when Assumption 2 holds, where T^m is the solution to the monopoly problem [M], defined in footnote 14.

ASSUMPTION 4: $F^B \geq n_2 \pi_2(T_1^1)$ when Assumption 2 holds, where T_1^s is the solution to problem $[W_i(s)]$, defined in Proposition 2.

ASSUMPTION 5: $W_2(T_1^1) \geq \underline{W}_2$ when Assumption 2 holds.

Assumption 3 ensures that firm B 's fixed costs do not exceed the

29. See Gaynor and Ma (1994), Ma (1994), Ma and Burgess (1993), and Wolinsky (1995) for models of health-care markets in which quality differentiation plays a key role.

30. In our model, if firm A knows patients' types while firm B does not, then when $F^A = F^B = 0$, both firms will earn zero profit serving high-risk patients, while firm A will attract all low-risk patients and earn an information rent in doing so. This rent may also allow firm A to operate as the sole supplier even when $F^A > F^B$.

31. See Newhouse (1984), Ma and McGuire (1995) and Chalkley and Malcomson (1995) provide formal analyses of settings where critical elements of health-care services are not contractible.

32. See Newhouse (1992) for useful observations on this point.

variable profit firm A would secure from delivering to all patients the component of the profit-maximizing health plan designed for low-risk patients. This assumption helps to ensure that all patients prefer the health plan they receive in equilibrium to no health plan. Assumption 4 places the lower bound on F^B described in the text. Assumption 5 ensures that this lower bound is strictly less than the upper bound specified in Assumption 3.

Proof of Proposition 3: Suppose $F^B > 0$. The first-order conditions of the program $[W_i(s)]$ for $i = 1, 2$ reveal that $\pi_2(T_2^s)$ is increasing in s and that $\pi_i(T_1^s)$ is decreasing in s for $i = 1, 2$. Since $\pi_2(T_2^0) = \pi_1(T_1^1) = 0$, continuity and the maintained assumptions ensure that there exists a unique $s^p \in [0, \infty)$ such that

$$\pi_2(T_2^{s^p}) = \pi_2(T_1^{s^p}). \quad (\text{A1})$$

By Assumptions 1 and 2, each $T_i^{s^p}$ for $i = 1, 2$ must satisfy the same first-order condition:

$$cu'(y - p_i) = H'(t_i) = \frac{cq'(e_i)}{1 - cq'(e_i)t_i} [\bar{H}_i - \underline{H}_i - H_i(t_i)]. \quad (\text{A2})$$

(A1) and (A2) imply that $T_1^{s^p} = T_2^{s^p}$.

Next, note that entry by the rival will not be precluded if $s^p > 1$, because $n_2 \Pi_2 > F^B$ in this case. Furthermore, $s^p \in [0, 1]$ if and only if $\pi_2(T_2^1) \geq \pi_2(T_1^1)$. [Recall $\pi_1(T_1^1) = 0$ and $\pi_1(T_1^s) < 0$ for all $s > 1$.] Assumption 4 ensures that this inequality holds, since $n_2 \pi_2(T_2^1) = F^B$. That is, $s^p \in [0, 1]$ if and only if

$$F^B \geq n_2 \pi_2(T_1^1). \quad (\text{A3})$$

Next, we verify that $T_i^{s^p}$ satisfies the individual rationality constraints:

$$W_i(T_i^{s^p}) \geq \underline{W}_i \quad \text{for } i = 1, 2. \quad (\text{A4})$$

Note that $W_i(T_i^{s^p}) \geq \underline{W}_i$ if and only if $\pi_i(T_i^{s^p}) \leq \pi_i(T_i^m)$ for $i = 1, 2$. Thus, a necessary condition for the existence of an F^B which satisfies both (A3) and (A4) is that $\pi_2(T_1^1) \leq \pi_2(T_2^m)$. This inequality is guaranteed by Assumption 5.

Finally, we identify the largest value of F^B for which (A4) is satisfied under the s^p health plan. Notice that (A4) holds at s^p if and only if

$$n_1 \pi_1(T_1^{s^p}) + n_2 \pi_2(T_2^{s^p}) \leq n_1 \pi_1(T_2^m) + n_2 \pi_2(T_2^m). \quad (\text{A5})$$

This conclusion follows from the necessary conditions associated with

the monopoly problem [M] defined in footnote 14. These conditions imply that $n_1 \pi_1(T_2^m) + n_2 \pi_2(T_2^m) < n_1 \pi_1(T_1^m) + n_2 \pi_2(T_1^m)$. Assumption 3 guarantees that inequality (A5) holds. Therefore, for each value of F^B for which $n_2 \pi_2(T_1^1) \leq F^B \leq n_1 \pi_1(T_2^m) + n_2 \pi_2(T_2^m)$, there exists a unique pooling Nash equilibrium of the complete-information duopoly game at $s^p \in [0, 1]$ in which all patients choose the same health plan. \square

Proof of Proposition 4: The proof is straightforward, once it is verified that no self-selection constraints bind in equilibrium. This fact follows from the observation that at the identified equilibrium, (1) $W_1(T_1) \geq \underline{W}_1 > \underline{W}_1 - [u(y) - u(y - p_2)] = W_1(T_2)$ and (2) $W_2(T_2) \geq \underline{W}_2 > \underline{W}_2 - [u(y) - u(y - p_1)] = W_2(T_1)$. \square

Proof of Proposition 5: First notice that the rival firm is unable to make a profit and attract some or all patients from the incumbent at the s^p equilibrium. If the rival firm could do so, the s^p equilibrium would not be an equilibrium of the complete-information duopoly game.

Notice further that the incumbent firm cannot earn more than $F^B - F^A$ by offering any other separating or pooling contract. Therefore, the incumbent weakly prefers the s^p equilibrium to any other equilibrium of the asymmetric information duopoly game. \square

Proof of Proposition 6: Since $F^A = F^B$, it is evident that any candidate for an equilibrium of the asymmetric information duopoly game must solve problem [WA-s]. $T^1 \equiv (T_1^1, T_2^1)$ will be an equilibrium of this game if it is not Pareto-dominated (i.e., if there does not exist another solution to [WA-s] that is preferred to T^1 by both types of patients). We will show that for any given $n_1 > 0$ and $\pi_2(T_1^1) \leq \beta \leq \pi_2(T_2^m)$, there exists an $n_2(\beta) > 0$ such that T^1 is not Pareto-dominated when $n_2 \geq n_2(\beta)$ and $F^B = \beta n_2$. To do so, first notice that because $\beta \leq \pi_2(T_2^m)$, the individual rationality constraints (3) in [WA-s] are satisfied at T^1 . Now, let $\Gamma \equiv \{T_1^s | s^p \leq s \leq 1\}$, $I \equiv \{T | W_2(T) = W_2(T_2^1)\}$, and $I^p \equiv \{T | W_2(T) = W_2(T^{s^p})\}$. The necessary conditions for a solution to [WA-s] reveal that Γ is bounded from above by I and from below by I^p in the p - t plane, and that all three curves coincide only at T_1^1 . Furthermore, as $n_2 \rightarrow \infty$, I^p converges pointwise to I . Therefore, Γ converges to I pointwise in s as $n_2 \rightarrow \infty$.

Next, note that $W_1(T) \leq W_1(T_1^1)$ (with equality only at T_1^1) for $T \in I$ with $\pi_1(T) \geq 0$. Since Γ converges to I pointwise in s as $n_2 \rightarrow \infty$, there exists an $n_2(\beta) > 0$ such that $W_1(T) \leq W_1(T_1^1)$ (with equality only at $s = 1$) for $T \in \Gamma$ with $\pi_1(T) \geq 0$, for all $n_2 \geq n_2(\beta)$. Therefore, T^1 is not Pareto-dominated by T^s for any $s \geq s^p$ when $n_2 \geq n_2(\beta)$.

When $s < s^p$, the necessary conditions for a solution to [WA-s] reveal that $W_1(T_1^s) < W_1(T_1^{s^p}) < W_1(T_1^1)$. Consequently, T^1 is not Pareto dominated for any $s \in [0, 1)$.

Therefore, T^1 is a Nash equilibrium at β when $n_2 \geq \hat{n}_2(\beta)$. The proof is completed by defining $n_2 \equiv \max_{\beta} n_2(\beta)$. This maximum exists because $[\pi_2(T_1^1), \pi_2(T_1^m)]$ is compact. \square

REFERENCES

- Anders, G. and R. Winslow, 1995, "The HMO Trend: Big, Bigger, Biggest," *The Wall Street Journal*, B1–B4.
- Chalkley, M. and J. Malcomson, 1995, "Contracting for Health Services with Unmonitored Quality," Discussion Paper No. 9510, University of Southampton.
- Che, Y. and I. Gale, 1997, "Buyer Alliances and Managed Competition," *Journal of Economics and Management Strategy*, 6, 175–200.
- Encinosa, W. and D. Sappington, 1995, "Competition among Health Maintenance Organizations," Mimeo, University of Michigan.
- Gaynor, M. and C.A. Ma, 1994, "Insurance, Vertical Restraints and Competition," Mimeo, Johns Hopkins University.
- Light, D., 1992, "The Practice and Ethics of Risk-Rated Health Insurance," *Journal of the American Medical Association*, 267, 2503–2508.
- Ma, C.A., 1994, "Health Care Payment Systems: Cost and Quality Incentives," *Journal of Economics and Management Strategy*, 3, 93–112.
- and J. Burgess, 1993, "Quality Competition, Welfare, and Regulation," *Journal of Economics*, 58, 153–173.
- and T. McGuire, 1995, "Optimal Health Insurance and Provider Payment," ISP Discussion Paper 59, Boston University.
- Marquis, M.S., 1992, "Adverse Selection with a Multiple Choice among Health Insurance Plans: A Simulation Analysis," *Journal of Health Economics*, 11, 129–151.
- Newhouse, J., 1984, "Cream Skimming, Asymmetric Information, and a Competitive Insurance Market," *Journal of Health Economics*, 3, 97–100.
- , 1992, "Pricing and Imperfections in the Medical Care Marketplace," in Peter Zweifel and H.E. Frech, eds., *Health Economics Worldwide*, Dordrecht, The Netherlands: Kluwer Academic Publishers, 3–22.
- , 1996, "Reimbursing Health Plans and Health Providers: Selection Versus Efficiency in Production," *Journal of Economic Literature*, 34(3), 1236–1263.
- Pauly, M., 1984, "Is Cream-Skimming a Problem for the Competitive Medical Market," *Journal of Health Economics*, 3, 87–95.
- , 1986, "Taxation, Health Insurance, and Market Failure in the Medical Economy," *Journal of Economic Literature*, 24, 629–675.
- and K. Langwell, 1983, "Research on Competition in the Market for Health Services: Problems and Prospects," *Inquiry*, 20, 142–161.
- Rothschild, M. and J. Stiglitz, 1976, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, 90, 629–650.
- Sloan, F., 1992, "Adverse Selection: Does It Preclude a Competitive Health Insurance Market?" *Journal of Health Economics*, 11, 353–356.
- Stiglitz, J., 1977, "Monopoly, Non-linear Pricing and Imperfect Information: The Insurance Market," *Review of Economic Studies*, 44, 407–430.

- Van de Ven, W. and R. Van Vliet, 1995, "Consumer Information Surplus and Adverse Selection in Competitive Health Insurance Markets: An Empirical Study," *Journal of Health Economics*, 14, 149-169.
- Weil, T., 1995, "Managed Health Care: A Utility-Style Monopoly?" *Public Utilities Fortnightly*, February 1, 14-15.
- Winslow, R., 1995, "AMA Sets Up Program to Help Doctors Fund Their Own Managed Care Plans," *The Wall Street Journal*, February 15, B6.
- and G. Anders, 1993, "Medical Industry Scrambles to Keep Up With Changes," *The Wall Street Journal*, September 13, A7.
- Wolinsky, A., 1995, "Regulation of Duopoly: Managed Competition vs Regulated Monopolies," Mimeo, Northwestern University.