INCENTIVES IN HMOS

by

Martin Gaynor
H. John Heinz III School of Public Policy and Management
Carnegie-Mellon University
And
National Bureau of Economic Research

James B. Rebitzer
Weatherhead School of Management
Case Western Reserve University
and
National Bureau of Economic Research

Lowell J. Taylor
H. John Heinz III School of Public Policy and Management
Carnegie-Mellon University

12/18/00

Please send correspondence to

Prof. James B. Rebitzer
Department of Economics
Weatherhead School of Management
Case Western Reserve University
Cleveland OH
Phone: 216 295 1419
e-mail: jim-rebitzer@po.cwru.edu

Abstract

With over 78.8 million enrollees, managed care organizations are the dominant form of health insurance in the United States. At the core of every managed care organization is a system of financial and non-financial incentives that encourage physicians to control costs. These incentives have been the subject of high stakes litigation and intense public controversy.

We use unique proprietary data from an HMO network to analyze the effect of incentives on medical costs and quality. We report three findings: (1) costs fall as financial incentives to control costs increase; (2) non-financial features of the incentive system (notably peer-pressure and mutual monitoring among physicians) may also influence costs; and (3) financial incentives can also elicit improvements in measured quality when quality measures are included in the incentive system.

The HMO we study runs a network of independent providers. In this network, primary care physicians are not employees of the HMO. Rather they are independent agents who sign contracts with the HMO. The primary care physicians in the network act as "gatekeepers" who determine the use of resources. Our focus on networks of independent providers is important. These organizations are rarely studied yet they comprise one of the largest segments of the managed care market. Networks of this type are also worthy of special attention because their incentive systems differ in important ways from those found in other types of health care organizations.

Cost data is available for the years 1994-1997 for the approximately 1000 primary care physicians in the network. Quality data is also available for 1997. In addition to this administrative data, we know the incentive contracts written by the HMO during this period. Details of the incentive system and changes in Federal regulations governing physician incentives create both cross-sectional and time series variation in the strength of financial and non-financial rewards for cost and quality. We exploit this variation to estimate the effect of incentives.

Introduction

With over 78.8 million enrollees, managed care organizations are the dominant form of health insurance in the United States. At the core of every managed care organization is a system of financial and non-financial incentives that encourages physicians to control costs. These incentives have been the subject of high stakes litigation and intense public controversy (Freudenheim, 1999; and Rubin et.al. 1999). Critics argue that incentives to reduce costs lead to a dangerous erosion of health care quality. Government, according to this view, should regulate incentive pay arrangements or, at a minimum, force managed care organizations to make their incentive arrangements clear to the public. Defenders counter that, without incentives, physicians will drive health care costs to stratospheric levels without commensurate improvements in health outcomes. Public policy, in their view, is too blunt an instrument to use to shape delicate managerial decisions concerning incentive design.

The public debate around HMO incentives is hampered by insufficient research into the workings of these incentives. Over the past decade economists have devoted increasing attention to the study of incentive systems (Prendergast, 1999). This literature offers important general insights, but it has largely overlooked incentives for health care providers. A parallel, but largely independent, literature focuses on the effects of HMOs on costs, treatment patterns and quality (Miller and Luft, 1994), and on the effect of financial incentives on physician decision making (Gold, 1999; Gaynor and Mark, 1999). Most of the studies in this literature compared physician practice under fee-for-service and managed care plans (e.g., Epstein, et al., 1986; Simon, et al., 1997; Hadley and Mitchell, 1997). Physician incentives within the HMO market segment, however, are varied and complex, and studies that examine the form of these incentives and their effect on outcomes are rare (Gold, et al., 1995b; Dudley, et al., 1998; and Gold, et al.,

1995a,b). Hillman, Pauly and Kerstein (1989) is an important exception to this generalization. Using regression analysis of data from a survey of HMOs, Hillman, et al. (1989) studied the effect of different compensation arrangements on utilization within HMOs. They found that physician payment method had a significant effect on utilization. The study provides strong evidence that contracting differences explain part of the cross-HMO variation in per enrollee hospitalization rates and in the rate of outpatient visits.

A key feature of Hillman, et al.'s study was that it analyzed differences in aggregate utilization *across* HMO's. Our research complements Hillman et.al. by examining the incentive arrangements *within* an HMO. This focus on incentives within a single network allows us to link the detailed structure of the incentive system to the medical costs attributable to a specific primary care provider. Reliance on data from a single HMO has another advantage as well – it eliminates the possibly confounding effects of unobserved, cross plan differences in enrollees or physicians.

The HMO we study runs a network comprised of independent providers. In this type of network, HMOs contract directly with physicians in independent practices (or with associations of such practices) to provide medical services. Our focus on IP networks is important. These rarely studied entities comprise one of the largest segments of the managed care market, roughly 40 percent of total HMO enrollment in 1998 (Gaynor and Mark, 1999). IP networks are also worthy of special attention because their incentive systems differ in important ways from those found in other types of health care organizations.

The incentive system of the HMO we study combines three features commonly found in other IP networks. The first is that the network relies on primary care physicians (PCPs) to act as "gatekeepers" to regulate access to medical resources. The HMO devotes considerable

attention to regulating and monitoring the behaviors of these PCPs. It also writes incentive contracts that provide financial rewards to PCPs who successfully control costs and maintain quality. The second feature, *group based incentives,* rewards the performance of panels of primary care doctors rather than individual physicians. These panel-based incentives reduce risk for physicians in the IP network, but they also encourage "free riding," i.e., the process by which one physician benefits from the cost reducing actions taken by others in the panel. In the economics literature, it is well understood that "free riding" undermines the effectiveness of group incentive systems (Weitzman and Kruse, 1990). A less-well studied aspect of group based incentive systems is the informal social processes that take place within the group when incentives are put in place (Hillman, Welch and Pauly, 1992; Kandel and Lazear, 1992; Encinosa, Gaynor and Rebitzer, 2000; Knez and Simester). In panels where physicians are able to monitor actions taken by other physicians, low powered group incentives and peer pressure can produce the same effect as high-powered financial incentives based on individual performance . Empirical studies regarding the interaction of "free-riding" and peer pressure for physicians are scarce, but the phenomenon is likely to be found wherever HMOs rely on group-based incentive contracts (Encinosa, Gaynor and Rebitzer, 2000). In the HMO we study, one of the expected benefits of panel based incentives was that it would cause physicians to monitor each other's actions and, in this way, stimulate cost containment.

A third feature of the incentive system we study is variation in the *independence of physicians*. Some physicians rely heavily on a particular insurer or network for patients, while other physicians are relatively independent because their patient population is made up of enrollees from a variety of different plans. As we discuss below, there are strong psychological

and economic reasons to believe that HMO incentives will be more potent for physicians who rely heavily on that HMO for patients.

This paper analyzes how the mix of financial rewards for "gatekeepers", group-based incentives, and the relative independence of PCPs shape the cost and quality of care. Our empirical investigations rely on a data set covering the years 1994-1997. The data is taken from the internal records of an IP network with roughly 1000 primary care physicians. The HMO's incentive system was first introduced in 1991 and was unchanged through 1996. In 1997, new Federal regulations required that the HMO reduce the intensity of cost-control incentives. We focus on the years 1994-1997 because these are the years for which data are available.

An important feature of the HMO's management style during this time period was that it wrote the same annual contract with all it's PCPs. In principle, the terms of these contracts could change from one year to the next, although in practice no changes were made from 1994-1996. Within any year, however, a single incentive arrangement prevailed.

Empirical analysis of the effect of incentives on costs and quality requires that we observe variation in the intensity of incentives at the same time we are observing costs and quality. How then can we learn about the effect of HMO incentives by studying data from a single HMO during a time period when it's incentives were largely fixed? The answer is to be found in the details of the HMO's group-based incentive system.

In order to spread the risks created by financial incentives, the HMO organized physicians into panels of doctors (PODs). Financial rewards were based on the performance of the *panel*, rather than on the performance of individuals. As a result of the way these panels were constructed, otherwise similar PCPs were grouped into quite different PODs. Panel size ranged from as few as 3 to more than 30 PCPs. In some panels, a few physicians relied heavily

on the HMO enrollees for patients while others in the POD did not. Some panels were made up of physicians in the same primary care specialty, e.g. pediatrics, who were also in the same corporate entity or practice partnerships. Other panels, however, were combinations of diverse specialties and the physicians in the panel lacked a natural connection to one another.

As we will discuss in following sections, the heterogeneity of the panels created variation in the intensity of incentives. It is this variation that we will rely on to examine the determinants of cost and quality for the years 1994-1996. In 1997, new Federal guidelines forced the HMO to reduce the intensity of it's incentive contracts. We also exploit this exogenously induced shift in incentive intensity.

The Incentive System.

Background

At the time of this study, the HMO was concerned with dramatically increasing it's market share in the commercial insurance market.[1] The trick to gaining market share was to sign up employers while, at the same time, offering physicians sufficient financial incentives to join the network. If the network was too small, the insurance would be less valuable to employees and hence to the employers who purchased the insurance product. Similarly, if the fees paid to physicians were too high or if the HMO failed to keep medical costs in line, employers might find the insurance products too expensive.

The HMO we study was closely connected to a university hospital system with a reputation for very high quality care. This association led the HMO to place a high value on

---

[1] The HMO was one of the first HMOs in it's market. In the time period we study, total HMO penetration in the market was roughly 20%.

delivering excellent medical care and attracting the best physicians to the network. The management of the HMO hoped to build an incentive system capable of balancing the tension between the aggressive product pricing needed to attract employers, and the high levels of compensation needed to attract the best physicians to the network.

In discussions with the current and past medical directors, we learned that the incentive system was designed to discourage unnecessary medical tests, specialist referrals, emergency room visits and long hospital stays. One of the medical directors explained the intended effect of the system by way of the following example. Patient X comes into the office complaining of headaches. The patient fears a brain tumor and wants an MRI, but the PCP determines quickly that this is very unlikely to be the cause of the headaches. Under fee for service, the PCP may decide to assuage the patient's fears by ordering an MRI or sending the patient to a neurologist. This action strengthens the physician's relationship with the patient and makes the patient feel better -- at little cost to the patient or the PCP. According to the medical director, the HMOs incentive system, if it was working properly, would cause the PCP to think twice before ordering the test or the referral.

Description of the Incentive System

Incentive Contracts from 1991-1996: PCP's were the HMO's gatekeepers. Every enrollee, as the HMO members were called, was assigned a PCP, and the PCP had to approve all medical procedures or specialist referrals.[2] The PCP was given responsibility for managing the medical budget of the HMO member under her care.

---

[2] Once a specialist was referred by a PCP, the specialist was free to order further tests and procedures without the approval of the enrollee's PCP.

Each enrollee was assigned an actuarially determined monthly budget based on the enrollee's age and gender as well as the specialty of the primary care physician.[3] This budget was a forecast of the number of dollars needed to cover the costs of all heath care services. The HMO established an account for each enrollee and "credited" to that account the monthly budget number. Health care costs were deducted from the enrollee's account as bills were paid. All health care costs were counted against enrollee credits except for the cost of dental care, pharmacy, mental health and substance abuse services, neonatal intensive care, and any expenses over $15,000 in a year. This last feature, known as a "stop-loss" provision, was very important as it removed incentives to withhold care from the sickest (or most expensive) patients [4]

PCP's were grouped into panels that varied by size and specialty mix. The panels did not need to be corporate entities or partnerships, but such entities were sometimes included in the construction of a POD.[5] The credits and debits for enrollee services were pooled for all PCP's in a panel and payouts were based on the performance of the entire panel. It is important to note that while primary care physicians were responsible for total medical expenditures, their income was determined by fees for the services they themselves performed. These fees were in the neighborhood of 20% of total medical expenditures. PCP services were paid at 125% of

---

[3] There were four specialty categories: pediatrics, internal medicine, family practice, and ob-gyn. The global budget for the HMO was also influenced by historical utilization rates at the HMO and predictions of the rate of medical price inflation.

[4] There is commonly a lag between the time that costs are incurred and the time they are received by the insurer. How then does an insurer decide to allocate these incurred but not recorded (IBNR) expenses? The practice at this HMO was to count in this year's costs, expenses incurred in the calendar year and repaid by May 31st of the subsequent year. Payouts to physicians occurred in June.

[5] Only 11 percent of the PODs in our sample were composed of a single entity for the purposes of receiving payment from the HMO.

Medicare's RBRVS (on the high side in the local market), but 20% of this potential income was "withheld" by the HMO.

If at year's end a panel was in the black, i.e. if total credits for all the POD's enrollees exceeded total debits, the withhold was returned to the panel. Panels with a surplus were also eligible for additional "gain-sharing" pay-outs. During the years 1994-96 the panels split the surplus (credits-debits) with the HMO. [6]

If a panel was in the red, i.e. if the sum of debits exceeded credits, the panel lost one dollar in withheld PCP income for each dollar of deficit (debits-credits). Thus, should the deficit exceed the value of the withheld income, the panel lost 20% of it's fees as well as the possibility for additional incentive "gain-sharing" bonuses.[7]

Figure 1 depicts the payment schedule that prevailed from 1994-1996. For illustrative purposes, the figure is drawn under the assumption that physicians deliver only one type of medical service and the quantity of that service is captured in the variable *S*. It is clear from the drawing that the incentive system is a complex, piecewise linear formula with notches or kinks at two points. As is often true in incentive systems with notches, incentive pressure is not applied equally across all physicians. For panels capable of operating near the target level of costs (S*), the system offers powerful incentives to keep to keep expenses below the allotted credits. In

---

[6] In some cases the withhold was automatically returned to the panel -- notably if the panel was in its first year in the IP network or if it had certain other, unspecified, problems. Details on "problematic" withhold returns are available only for 1996 and 1997, and these problem panels (three in all) were dropped from our sample.

[7] The HMO also withheld 5% of earnings from some specialists referred to by the PCP. No withhold was placed on emergency room, radiology, pathology and anesthesiology services. The withholds were returned to the specialists based on the performance of the referring PCP's panel. Conversations with HMO managers suggest that these specialist withholds were so diffuse and involved such small sums of money, that they could be safely ignored in the analysis of PCP incentives.

contrast, panel's whose style of practice and/or patient mix offer little hope of making the target, experience relatively little incentive pressure. Indeed, a panel that always functions below S** , the system becomes fee for service at a price of 1.05 Medicare's price. In contrast, under a conventional capitation system, incentive pressures are identical regardless of the ability of the panel practice cost-containment.

The "notches" in the incentive system are consistent with the HMO's market strategy of being the high quality (rather than low cost) HMO in the market. The system has a few notable weaknesses however. First, high cost panels (those with expected costs far above S** in Figure 1) can exploit the system by "churning" patients. Secondly, the effectiveness of the system depends critically on the HMO forecasting costs and setting targets appropriately. An unanticipated increase in medical utilization can cause physicians to run up debits so far in excess of credits that they have no hope of having their withhold returned. For panels sufficiently far "under water", the incentives to control costs disappear.[8]

Another feature of the incentive system highlighted in Figure 1 is that incentives to reduce medical utilization are greatest for services that involve physician's other than the primary care physician. Thus the payout to reduce visits to specialists is greater than the payout to reduce similar visits to the PCP's office.

A critical feature of the HMOs incentive system that is not depicted in Figure 1is the construction of the PODs. In their written material, the HMO emphasizes that doctors decide which PODs they wish to join, but the actual construction of the panels seems to have been quite haphazard. No one we spoke with could reconstruct the process by which physicians ended up

---

[8] We do not yet have the monthly data required to directly observe what happens when plans "go under water" during the year, but we hope to collect and analyze such data in subsequent papers.

in the PODs they did and no records were kept regarding the decision process. In some instances, it appears that the HMO simply assigned physicians to panels as means of sharing risk. In other cases, physicians joined panels because they were in the same practice, were located close to or had other social/professional relationships with other panel members. Physicians did move across panels, but as we show in the following empirical analysis, there does not appear to be a significant relationship between costs and the decision to change panels.

Finally, we note that the HMO did not and could not steer patients towards low-cost physicians or PODs in its network. Once an employee signed up with the HMO, she is free to use any physician in the network. Similarly, the HMO did not engage in economic credentialing. This means that as matter of policy and practice, the HMO did not try to force high cost PCPs out of it's network. Once a PCP agreed to join the IP network and once the PCP passed the network's screening process, continued participation in the network was *not* influenced by the cost of services utilized by enrollees.

The Rules Change in 1997: In 1997 the Federal government introduced new guidelines governing physician incentive plans. Under these rules, physicians could have no more than 25% of their earnings "at risk". In order to bring its system into conformity with the new Federal guidelines, the HMO had to weaken it's incentive payout formula.

Under the new system the HMO introduced in 1997, the HMO continued to pay at 125% of Medicare's RBRVS, but only 10% of income was withheld. Panels were rank ordered on the basis of the size of the surplus in their patient care accounts (credits-debits) and those with a surplus had their withheld income returned . The HMO also abandoned the old incentive bonus under which PODs split their patient care account surpluses with the HMO. In place of this "gain-sharing" bonus, the POD substituted the following system. Panels whose surplus put them

in the top quartile of panels with a surplus received a bonus of 7.5% of their earnings. Those in the second and third quartiles had a bonus of 5 percent of their earnings and those in the 4th quartile received a bonus of 2.5 percent of their earnings. In sum, the system introduced in 1997 had the effect of cutting the minimum payout for achieving cost targets in half (from 20% to 10%) and capping the "gain sharing" payout for cost reduction at 7.5% of earnings.

Analysis of the Incentive System

Having described the salient features of the incentive system, we now describe our strategy for identifying the effect of incentives on costs. [9] We focus our attention on the problem of moral hazard in teams.

Free-Riding and Average Costs in a POD: Ceteris paribus, as the number of group members increase, the link between individual actions and group performance is weakened. Therefore, as group size increases, members know that others in the group find it less attractive to take costly private actions to improve group performance. This knowledge about others' incentives reduces each individual's incentives to improve group performance. Thus, in equilibrium, larger groups will have lower performance levels than smaller groups. The effect of free-riding on group performance can be expressed as follows. *For a given number of HMO members, panels with more physicians will have higher average costs than otherwise identical panels with fewer physicians.*

---

[9] This analysis is derived more formally in a preliminary model of the HMO's "incentive" system. This model relies on a number of restrictive assumptions and focuses on the "withhold" aspect of the system. We believe that many of our restrictive assumptions can be relaxed and that we can incorporate the full structure of the system without altering our key findings. Future drafts of the paper will incorporate a more general and realistic model into the main body of the text.

The motivation for the preceding hypothesis came from the conventional economic analysis of moral hazard in teams. An alternative motivation for this hypothesis can be drawn from the informal social processes that take place within the group or panel (Hillman, Welch and Pauly, 1992; Kandel and Lazear, 1992; Encinosa, Gaynor and Rebitzer, 2000). In panels where physicians are able to monitor actions taken by other physicians, peer pressure can induce feelings of shame or guilt that can have the same effect on performance as high-powered financial incentives. The HMO we study was careful to distribute to every member of the POD information on the performance of all the other physicians in the panel. In settings where physicians are able to easily interpret this information and where frequent social interactions make peer pressure meaningful, we would expect that informal group interactions may undo some of the effects of free-riding. Under the reasonable assumption that mutual monitoring and the associated sanctions/rewards are more effective among physicians in small groups, we would expect smaller PODs to have less free-riding than larger PODs.

A Closer Look At the Incentive Problem: Economic models of moral hazard in teams typically involve group members who can chose to undertake costly actions that improve group performance. A clear understanding of the incentive problem within PODs requires that we also understand the kind of "costly actions" that PCPs must take to control medical utilization costs.

We recently attended a small meeting of physicians in the HMOs' network. This meeting was part of a regular series of meetings whose purpose was to discuss ways to keep medical utilization costs down. In this meeting a relatively inexperienced physician asked the other, more experienced physicians, how one goes about controlling costs due to medical utilization. The lively ensuing discussion centered around the following three strategies. First, keep patients out of the emergency room by offering extended hours, sophisticated answering services, and

appropriate phone-medicine. Second, teach patients to be take responsibility for intelligently managing their own health care -- especially patients with such chronic diseases as congestive heart failure, diabetes and asthma. Third, reduce unnecessary referrals to specialists. Sometimes this last strategy requires offering speedier access to the PCP than to the specialist. Thus, to cite an example taken from the discussion, if a patient calls in with an ankle sprain and wants to be referred to a sports medicine specialist, the PCP can discourage the referral by offering to see the patient quickly, before the specialist is available. Alternatively the PCP can discourage specialist visits by learning more about the specialty. Rather than, for example, referring all patients with chest pains immediately to a cardiologist or referring all patients with acne or rashes immediately to a dermatologist, PCPs can get additional training to help identify when such referrals are truly necessary.

The important point about the preceding list of cost-reduction strategies is that some of them involve fixed costs (learning more about cardiology or dermatology) and others involve variable costs (teaching each patient with asthma how to manage their asthma at home). To the extent that fixed costs are important, one would expect to find that PODs with a large number of HMO members would benefit most from undertaking cost-reduction actions. For this reason, econometric investigations of the importance of moral hazard in PODs requires that we condition our estimates on the number of HMO members the POD sees.[10]

Even if fixed costs were not important for most cost containment activities, other incentive considerations would still suggest a relationship between the number of HMO members seen by the POD and average POD medical costs. One of the general results from the

literature on incentive design is that the power of an incentive instrument to shape an agent's behavior increases as the "precision" of the performance measure increases (Milgrom and Roberts,). Thus, if the principal is relying on a noisy measure of agents' actions, the effectiveness of the incentive increases with the ratio of signal to noise in the performance measure.

In our context, the key performance measure is average medical costs. Average medical costs are noisy performance measures because they are determined by an unobserved combination of the cost-containing practices adopted by physicians and by random events specific to individual patients. These random events make it possible that panels adopting aggressive cost-control measures will have a budget deficit, and that panels who pay no attention to costs will end up with a budget surplus. Reducing the importance of these random factors will have the effect of linking rewards and sanctions more closely with actions. As the number of HMO enrollees signing with a POD increases, the effect of random shocks cancel one another out and the ratio of signal to noise increases. It follows that HMO incentives have a stronger influence on behavior in PODs that attract many HMO enrollees.

HMO Membership and Average Costs for Individual PCPs: Physicians with relatively large share of their panel's members will, ceteris paribus, have more influence on POD performance than other physicians. It follows from the logic of moral hazard in teams, that physicians with large numbers of members (relative to the panel) have greater incentives to control costs than others. The existence of substantial fixed investments in the practice of "cost containment" would serve to reinforce this relationship. *As a result we expect to find a negative*

---

[10] Note that a negative association between enrollees and average costs cannot be due to the HMO steering members to low cost physicians because, as we note above, the HMO had no way to

relationship between the number of HMO members a physician has and the average costs incurred by these members.

An Empirical Analysis of The Effect of Incentives on Costs

For the years 1994-1996, we identify the effect of incentives by comparing average costs in PODs where free-riding is likely to be more severe with PODs where free-riding is likely to be less severe. Specifically, we examine the following hypothesis:

- *Hypothesis 1* For a given number of HMO enrollees, panels with more physicians will have lower average costs than panels with fewer physicians.

In 1997, changes in Federal law required that the HMO dramatically reduce financial incentives for cost-containment. This exogenously induced change in incentive intensity allows for an auxiliary test of hypothesis 1. We state this test in the form of the following hypothesis:

- *Hypothesis 1a:* The reduction in incentive intensity in 1997 should weaken the relationship between number of PCPs and average medical costs.

The logic of moral hazard in teams suggests that, ceteris paribus, the incentives for an individual PCP to control costs increases with the share of HMO members the physician has. This relationship would be reinforced if the practice of "cost containment" entailed substantial fixed costs. This reasoning leads us to the following hypothesis: [11]

- *Hypothesis 2:* As the number of HMO enrollees seen by an individual PCP increases (relative to the panel), the average costs incurred by these enrollees falls.

In what follows we present empirical investigations of these hypotheses.

---

influence which PCPs its physicians chose.

[11] Changes in physician ID numbers between 1996 and 1997 make it impossible to track individual PCPs between 1996 and 1997. We therefore cannot estimate a hypothesis 2a analogous to hypothesis 1a.

Hypotheses 1 and 1a: The Determinants of Costs at the Panel Level

We have data on average medical expenditures per enrollee by panel for the years 1994-1997.[12] We use these data to construct the variable *Total Medical Expenditures*. Although *Total Medical Expenditures* is a measure of annual average costs per member, it is conventional in the industry to express this number on a monthly basis. Thus, a single HMO member who stays with the HMO for a year is said to contribute 12 member-months to the POD's total membership. If this member incurs expenses of $1080 in a year, *Total Medical Expenditures* for that member is then calculated as $90 per member per month or $90 pmpm.

We use *Total Medical Expenditures (pmpm)* as a dependent variable in our examination of hypotheses 1, and 1a. Specifically, we estimate the following equation for POD *i* in year *t* :

(1) $$\textit{Total Medical Expenditures (pmpm)}_{it} = \alpha + \alpha_1 \textit{Log Physicians in POD}_{it} + \alpha_2 \textit{Log Member-months}_{it} + \alpha_3 \textit{Log Physicians in POD}_{it} * 1997 + \alpha_4 \textit{Log Membermonths}_{it} * 1997 + \alpha_5 X_{it} + \varepsilon_{it}$$

Log Physicians in POD is the log of the number of physicians in panel. *Log Member-months* is a measure of the average number of HMO enrollees in a POD during a year. [13] *Log Physicians in POD*1997* and *Log Membermonths *1997* represent interactions of the physician and member-month variables with a dummy variable equal to 1 if the year is 1997. Finally, X is a column vector of variables and α_5 is a row vector of parameters that capture the influence of other factors determining medical expenditures. We represent the mean zero, normally distributed error term in equation (1) by ε . Free-riding in groups (hypothesis 1) suggests implies

---

[12] The HMO actually collected medical expenditures on a monthly basis, but we were allowed access only to the annual data used to calculate incentive payouts.

[13] Member-months is the standard quantity measure of enrollments in the insurance industry. An enrollee who joins the HMO mid-way through the year is said to have added 6 member-months while an enrollee who was present the entire year counts for 12 member months.

that $\alpha_1 > 0$ while the Federally mandated reduction of incentive intensity in 1997 (hypothesis 1a) suggests that $\alpha_3 < 0$.

Column 1 of Table 1 presents pooled-time series estimates of equation (1) for the years 1994-1997.[14] Consistent with hypothesis (1), we observe that PODs with more physicians have lower total medical expenditures. This coefficient is economically as well as statistically significant. An increase in POD size from 10 to 12 physicians is associated with an increase in costs of $6.58 per member per month. This is an increase of 7.3% over the sample mean - implying an elasticity of costs with respect to number of physicians of 0.4. Consistent with hypothesis 1a, we also observe that the coefficient on *Log Physicians in POD*1997* is -29.67. Point estimates suggest that in 1997, an increase in POD size from 10-12 physicians led to an increase in costs of only $1.18 per member per month.

Not surprisingly, given our discussion of the fixed costs involved in cost-containment, the coefficient on *Log Member-months* is negative and statistically significant. Thus PODs with 20% more members than the mean panel have, average costs that are $2.00 lower (per member per month). The small and statistically insignificant coefficient on *Log Membermonths*1997* indicate that this relationship was not much different in 1997 than in 1994-1996.

In order to capture differences in expected costs due to enrollee age and gender and physician specialty, we include the HMOs target expenditures for the POD, *Target Expenses (pmpm)*. We find that the HMO is doing a pretty good job predicted expenditures from one year to the next. The coefficient of 0.99 on *Target Expenses* suggests that a $1 increase in expected POD expenses correlates with a $0.99 increase in actual POD expenses in the following year.

The equation in column (1) also includes the variable *Years that POD is in Sample Since 1994* This variable is important because during the time of this study, the HMO was expanding from urban settings into suburban areas. The positive coefficient on *Years POD is in Sample Since 1994* indicates that the older, more urban, panels were likely treating more expensive populations of enrollees. Finally the estimates in column (1) include year fixed effects with 1994 being the omitted year.

Column (2) in Table 1 introduces a new dependent variable, *Medical Expenditures Net of Stop Loss*. Under the HMO's incentive system, physicians are not held responsible for expenditures above $15,000 per patient in a year. We would, therefore, worry about our interpretation of the coefficients on *Log Physicians in POD* and *Log Physicians in POD*1997* if they changed dramatically when we net out "stop-loss" expenditures. We observe instead, however, that the coefficients in column (1) are virtually identical to those in column (2).

Column (3) introduces another dependent variable, *Fraction Withhold Returned.* The HMO cares about regulating *Total Medical Expenditures*. The physicians, in contrast, care about whether they can win back their withhold by containing costs. Consistent with moral hazard in panels, we find that as the number of physicians in a POD increases, the panel is less likely to win back their withhold. Point estimates indicate that an increase in panel size from 10 to 12 physicians reduces the fraction withhold returned by 0.04 or roughly 10% of the mean. The coefficient on *Log Physicians in POD*1997* is large (0.12) and positive, but imprecisely measured and not statistically significant. Our interpretation of this result is that physicians

---

[14] As we note in the text accompanying Table 1, our estimates are from a time-series /cross-section panel and we have adjusted standard errors to account for possible within POD correlation of errors and cross-panel heteroscedasticity.

responded to the weakened incentives in 1997 by reducing cost containment actions, but not by so much as to influence their chances of having their (much reduced) withhold returned.

Estimates in columns (4)-(6) of Table 1 add additional control variables to capture heterogeneity across PODs and their enrollee populations. We introduce a dummy variable indicating whether the panel is composed entirely of physicians in the same specialty and a vector of dummy variables equal to 1 if a POD was a member of one of the large networks of providers in the local area (*POD in Network 1 – 5*). We also include variables that capture the specialty composition of the panel (*Fraction POD in: Family Practice, Internal Medicine, or Pediatrics* and *POD Only OB/GYN)*. Finally, we introduce the variable % *Medicaid > 10 in 1997* . This variable indicates wheter more than 10% of the enrollees in a POD in 1997 were in the HMOs Medicaid "line of business".[15] There were no Medicaid enrollees in the HMO until 1995 and only a very small number in 1996. By 1997, however, there were enough Medicaid enrollees that the HMO began tracking them by panel. If Medicaid enrollees are more expensive enrollees, then PODs that attracted relatively large numbers of them in 1997 may be different from other PODs in other unobserved ways. For example, if the physicians in a panel are located close to an inner-city neighborhood where incomes are low, we would expect them to have both a higher proportion of Medicaid patients as well as a higher proportion of poverty related medical expenses in their non-Medicaid population. We find only limited some support for this supposition in Table 1. The coefficient on % *Medicaid> 10 in 1997* is positive but imprecisely measured and not statistically significant.

---

[15] We chose the 10% cutoff because HMO management used this same cutoff in "adjusting" quality measures by POD. We discuss quality measures later in this section of the paper.

The addition of numerous control variables in columns (4)-(6) had no effect on the signs and statistical significance and little effect on the magnitudes of the coefficients on our key variables of interest, *Log Physicians in POD* and *Log Physicians in POD*1997*. These results suggest that our findings are not likely to due to correlation between the number of physicians in a POD and other, observed characteristics of the panel, it's physicians or it's patient population.

An alternative approach to controlling for POD heterogeneity is to estimate cost equations with POD level fixed effect variables. These fixed effect estimates are presented in columns (7)- (9) of Table 1. The coefficients on the key variables, *Log Physicians in POD* and *Log Physicians in POD*1997* are not qualitatively changed from those in equations (1) - (6), although the magnitude of the effect of number of physicians on costs increases in absolute value. Thus, in our fixed effect estimates, a POD that increases the number of physicians from 10 to 12 will, for the years 1994-1996, experience an increase in *Total Medical Expenditures (pmpm)* of $8.28 per member per month This effect falls to $3.29 per member per month in 1997.

We have so far interpreted the fixed effect coefficients on *Log Physicians in POD* and *Log Physicians in POD * 1997* in columns (7) - (9) as reflecting changed behaviors due to the changing degree of moral hazard in teams. An alternative interpretation is that it is selection rather than behavioral changes that are driving the result. It may be, for example, that *Log Physicians in POD* has a positive correlation with *Total Medical Expenditures (pmpm)* because panels that reduce the number of physicians do so by kicking out their most expensive physicians. Alternatively, it may that panels with growing numbers of physicians are hiring new members who tend to be more expensive than incumbents. It would not be surprising if the

HMOs incentive system had selection effects, and in studies of other incentive systems, the primary effect of incentives is via selection (Lazear, ; Encinosa, Gaynor and Rebitzer,).

Table 2 offers an investigation of the importance of selection effects. It relies on individual cost data for the HMO's primary care physicians. For the years 1994-1996 the HMO's data system allows us to track the movements of individual physicians as they enter the network and then move across panels.[16] We use this data in Panel A to see if those leaving shrinking groups have costs that are, on average, higher than others in their panel. We observe that the 298 physicians who exited a panel at the end of 1994 or 1995, had costs that were, on average, 2 percent higher than other PCPs in the panel they are leaving. This difference between movers and stayers is both small in magnitude and statistically insignificant. To put the Table 2 , Panel A figures in perspective, remember that the point estimates in column 1 of Table 1 suggest that a panel decreasing from 12 to 10 members will experience a 7% reduction in *Total Medical Expenditures*. If the leavers in such a panel have costs that are, on average, 2 % higher than the stayers, the selection effect of reductions in physicians in the panel should be only 0.4%.

In Panel B of Table 2, we compare total medical expenditures between incumbent physicians and new entrants in the panel. Of the 746 new entrants in the years 1995 or 1996, we observe that entrants have costs that are, on average, 1% higher than those in the panel they are entering. This difference is far too small to account for the results in Table 1. It is also statistically insignificant.

Taken together, the results in Tables 1 and 2 offer strong support for hypotheses 1 and 1a. Ceteris paribus, panels with more physicians have higher costs, and the weakening of

---

[16] Changes in physician ID codes made it difficult to link individual physicians in 1997 to those in the 1994-1996 data.

incentives in response to new Federal regulations weakened the effect of panel size on costs. We can extend our analysis of moral hazard by analyzing how variation in the number of physicians in a POD influences different types of medical expenditures. In Table 3, we analyze various sub-components of total medical expenditures for 1994-1996, the years when this data are available. As a robustness check on the importance of dropping data from 1997, column (1) of Table 3 re-estimates the equation in column (4) of Table 1 but restricts the sample to the years 1994-1996. The sign and standard error of the coefficient on *Log Physicians in POD*, are very close to those estimated over the full sample of years (see column 4 of Table 1). This suggests that any differences in results between Tables 1 and 3 are not likely to be due to the absence of data from 1997.

Under the HMO's incentive system, physicians can reduce costs in two ways. First, they can reduce the amount of income that accrues to them via fees by, for example, reducing the number of office visits they have with patients. Alternatively physicians can cut out expenses that accrue as fees to other health care providers, e.g. by reducing specialist referrals. Given the payout formula in the HMO's incentive contracts, we can expect physicians to prefer the second method of reducing costs to the first.

Column (2) of Table 3 presents estimates of the determinants of *Expenditures Not Paid to Self (pmpm)*. This variable measures the average, per member per month medical costs of procedures and treatments that do not earn revenues for the primary care physician. We find that the coefficient on *Log Physicians in POD* is positive, statistically significant, and similar in magnitude to the coefficient in column (1). In contrast, the dependent variable in column (3) measures expenditures that ***do*** earn income for the primary care physician. The coefficient on *Log Physicians in POD* in column 3 is positive, but small in magnitude and not statistically

significant. Taken together, the results in columns (2) and (3) suggest that in PODs with fewer physicians (and consequently higher powered incentives), physicians cut expenditures on services and procedures that generates fees for other physicians. Physicians do not seem to cut expenditures on services that generate fees for themselves.

Columns (4) and (5) of Table 3 compare the determinants of costs due to in-patient and out-patient services and procedures. The dependent variable in column (4), *Total In-Patient (pmpm)*, tracks costs due to procedures conducted in a hospital on an in-patient basis. These costs are determined by the resource based relative value scales (RBRVS) codes that Medicare uses for reimbursement. We observe that the coefficient on *Log Physician in POD* in column (4) is small and not statistically different from zero. Column (5) estimates the determinants of out-patient expenses. The variable *Total Out Patient* includes costs incurred in a hospital (on an out-patient basis), a clinic, or a doctor's office. In equation (5) of Table 3, we find that *Log Physicians in POD* has a positive and statistically significant effect on out-patient costs. The magnitude of this effect is large. Increasing panel size from 10 to 12 physicians increases outpatient costs by $1.71 per member per month or roughly 5% above the mean. The implied elasticity is 0.26.

Taken together, the results in columns (4) and (5) suggest that the HMO's incentives do not have much influence on in-patient expenses, but are quite important for out patient expenses. This conclusion makes sense. Patients admitted to hospitals are likely to be sicker than other patients and therefore are less likely to have purely elective procedures. Patients in a hospital are also more likely to have their care decided by a specialist rather than the by the primary care physicians and it is the PCPs who are most effected by the HMO's incentive system. In contrast, out-patient procedures are more likely to be elective and more likely to be conducted by the PCP.

We would therefore expect the incentive system to work most powerfully on *Out Patient Expenses*.

Hypothesis 2: The Determinants of Costs at the PCP Level

We have so far focused our analysis on hypotheses 1 and 1a. These hypotheses concern the determinants of costs incurred by the physician panel as a whole. Hypothesis (2), in contrast, examines the costs incurred by members treated by an individual PCP. Specifically, hypothesis (2) states that as the number of HMO enrollees seen by an individual PCP increases the average costs incurred by these enrollees falls. We have two theoretical reasons for expecting this negative relationship between member-months and average costs for a PCP. First, increases in a PCP's share of the panel's HMO enrollments increases the influence that the PCP's actions have on panel performance. Under the HMOs group incentives, this naturally increases incentives to control costs. Secondly, if adopting a cost-containment practice style entails significant up-front investments, the returns to these investments increase with the number of HMO enrollees managed by a PCP.

The HMO provided us with cost data for individual primary care physician for the years 1994-1996. We used this data to estimate the following equation:

(2) $$\textit{Total Medical Expenses (pmpm)}_{it} = \beta + \beta_1 \textit{Target}_{it} + \beta_2 \textit{Log Member-months for Individual Physician}_{it} + \beta_3 \textit{Log Member-months of Other Physicians in POD}_t + \beta_3 Z + \phi_{it}$$

where $\textit{Total Medical Expenditures(pmpm)}_{it}$ is the per member per month total of debits in the patient care accounts of enrollees under the care of physician *i* at time *t. Target* is the total credits in these same patient care accounts; $\textit{Log Own Member-months For Individual Physician}_{it}$ is the natural log of physician *i*'s member-months at time *t* and *Log Member-Months of Other Physicians in POD* is the log of member-months for all physicians in the panel other than *i* at

time *t*. The letter *Z* represents a vector consisting of physician fixed effects and year effects. Hypothesis (2) implies that $\beta_2 < 0$. If there are no significant up-front investments in cost-containment, the logic of moral hazard in teams would also suggest that β_3 is positive. This follows because increases in other PCP's member months reduces each individual PCP's influence on total panel costs. Alternatively, if there are significant up-front investments in cost-containment, than the membership belonging to other PCPs in the panel may be relatively unimportant in determining physician *i's* actions.

Column (1) of Table 4 estimates equation (2) for the years 1994-1996. We observe that β_2 is negative and statistically significant.[17] The size of the coefficient indicates that the economic effect of own member months is also large. A 20 percent increase in a PCP's member-months from one year to the next reduces per member per month total medical expense by roughly $2.00. This result supports hypothesis (2). The coefficient on *Log Number of Other Physicians in POD* is negative but it is measured imprecisely and it is not statistically different from zero.

Columns (2) and (3) in Table 4 re-estimate the cost equation in column (1) with dependent variables that track in-patient and out-patient costs respectively. As was true in the in the panel level estimates in Table 3 we find incentive effects for out-patient but not in-patient costs. The coefficient on *Log Member-Months for Individual Physician*, is negative and statistically significant for out-patient expenses (column 3), but not in-patient expenses (column 2).

---

[17] All the OLS regressions in Table 4 are calculated using robust standard errors for within physician clustering of error terms.

Column (4) re-estimate the equation in column (1) but adds another right hand side variable, *Log Number of Other Physicians in POD*. In this specification, as before, an increase in *Log Member- Months For Individual Physician* has the effect of reducing a physician's *Total Medical Expenditures*. The coefficient on *Log Member-Months for Others in POD* is negative (and strongly statistically significant) while the coefficient on *Log Number of Other Physicians in POD* is positive, significant and roughly twice the magnitude of *Log Member-Months for Others in POD*. These two coefficients, taken together, suggest that increasing the number of *other* physicians in the panel by 20% and increasing their member-months by 20% will lead to a $7.80 increase in total medical expenditures for an individual physician.

In our view, the results in column (4) may be best explained by processes of peer pressure and mutual monitoring that take place in groups. As the number of *other physicians* in the panel increases, each individual physician has less opportunity to monitor the actions of others in the panel. The accompanying reduction in peer pressure makes it less likely that other physicians in the panel are taking the costly actions needed to control costs. This, in turn, reduces the incentives that each individual physician has to devote time and energy to reducing their own medical expenditures. The result is that increases in the number of *other* physicians increases each individual physician's average costs. The results in column (5) and (6) indicate that similar moral hazard considerations are important for the determinants of out-patient expenditures, but not in-patient expenditures.

Incentives for Quality

Our results in Tables 1-4 are broadly consistent with hypotheses (1), (1a) and (2). We interpret these findings as indicating that the HMO's incentive system is effectively motivating the primary care physician "gate-keepers" of the network to reduce costs Much of the

controversy surrounding HMO incentive systems involves the claim that cost-control efforts require physicians to sacrifice the quality of medical care. In this section we investigate the determinants of care quality in the HMO and the ability of the incentive system to influence care quality while restraining costs.

The HMO first incorporated quality measures into its incentive system in 1997. As discussed above, this change was part of a general overhaul of the system to bring it into compliance with the then new Federal regulations governing physician incentive plans.

Under the new system, quality incentives were linked to cost-control incentives. Panels with a surplus in their enrollee accounts were rank ordered based on care quality and bonuses were paid out according to rank. Panels in the top quartile for quality received a bonus of 7.5% of earnings. Those in the second and third quality quartile received a bonus of 5% of earnings and those in the 4th quartile received 2.5% of earnings. A key feature of this system is that quality payouts were conditional on having successfully kept patient costs below the targeted levels for that year.

In constructing quality measures, the HMO gave PODs points based on two types of quality measures. The first type of measure focused on preventative care and included such things as the percent of women aged 50-65 who had mammograms or the percent of children aged 3-5 who had annual well-child evaluations. This type of measure is referred to as a HEDIS measure because the indicators were derived from a set of standardized quality measures put forward by the Health Plan Employer Data and Information Service (HEDIS). [18] In addition to the selected HEDIS measures of preventative care, the HMO constructed a second quality

---

[18] HEDIS is sponsored, supported and maintained by the National Committee for Quality Assurance (NCQA), the organization responsible for accrediting managed care organizations.

measure based on indicators of patient satisfaction, the accessibility of the PCPs after hours and direct office inspections. Details on the plan's HEDIS and NON-HEDIS quality measures are presented in Data Appendix 2. As we describe in the appendix, the NON-HEDIS measures proved to be both difficult to collect and of doubtful relevance to panel quality. For this reason, we focus our analysis on the HEDIS measures.[19]

Table 5 analyses the determinants of cost and quality in 1997. Column (1) of Table 5 presents a simple regression of *Total Medical Expenditures (pmpm)* on *HEDIS Quality Points*. We find a statistically significant and negative relationship between costs and quality, i.e. the panels with the best quality scores have the lowest average costs. In columns (2) and (3) we find that this negative correlation is statistically significant for stop-loss expenditures but not for other medical expenditures.

Under the incentive contract in place in 1997, payouts for quality were influenced by the same "free-riding" problems as payouts for cost-containment. In addition, panels were not eligible for quality bonuses unless they also kept costs under target levels. For both these reasons, we would expect quality measures to deteriorate as the number of physicians in the panel increased. Improvements in quality are reflected in an increase in the *Hedis Quality Points* measure. Thus, we can state the quality analogy to hypotheses 1 as follows: *Hedis Quality Points* should decrease as the number of physicians in the panel increases (conditional on member-months).

---

[19] For two of the non-HEDIS measures (membership retention and extended hours), the indicators were so unreliable that all panels were simply given the maximum number of points at the end of the year. The third non-HEDIS measure (patient satisfaction) gave maximum points to all but two panels. The fourth measure (whether the panel was accepting new HMO enrollees), has little connection to quality of care. The fifth measure (focused office reviews) was based on criteria that were not clearly explained to physicians.

Column (4) of Table 5 analyzes the determinants of the *HEDIS Quality Points* earned by a panel. Consistent with moral hazard in teams, we find that the number of quality points earned by a panel falls as the number of physicians in the panel grows. The coefficient is statistically and economically significant. Increasing the number of physicians from 10-12 reduces *HEDIS Quality Points* by 1.19 or 7% of the mean. The implied elasticity is 0.41.

Column (5) of Table 5 re-estimates the quality equation in column (2), but introduces additional variables that measure characteristics of panels, their physicians and enrollees. Introducing these variables has little effect on the magnitude and statistical significance of the coefficient on *Log Physicians in POD*.

The results in columns (4) and (5) are consistent with the notion that the HMOs incentive system is moving physicians to offer both lower cost care and care with higher measured quality.

Conclusion

The evidence we have analyzed suggests that the HMO's incentive system had a marked effect on costs and quality in the years 1994-1997. There is some reason to question, however, whether these results can persist over long periods of time.

The possible fragility of the incentive effects can be traced to the system's reliance on "target" expenditures. Target expenditure levels are ratcheted downwards each year as the previous year's cost-containment efforts are built into next year's forecasts. If panels make the easy cost-cutting moves first, then the rate of improvement may slow from one year to the next. If this slowed rate of improvement is not factored into forecasted expenditures, many panels may find themselves too far "under water" to hope to make the current year's incentive target. The resulting discouragement may undermine physician confidence in the fairness and generosity of

the HMO's system. Similar problems will result if medical costs increase in ways not built into the "target" estimates.

It may be possible to investigate the "fragility" of the incentive system by examining the evolution of medical costs over the year. Specifically, the behavior of panels and physicians should change over the course of the year as it becomes clearer whether they are likely to meet their target. If there is a marked "discouragement" effect for groups that have little hope of meeting their target this may indicate the importance of the fragility effect discussed above. We hope to undertake such a study using data from 1999 and 2000.

Our analysis suggests that the key to the HMO's incentive system lies in the operation of the physician panels. To date, however, we know very little about what actually happens within these panels.

Our surmise is that panels do well on cost-containment when: (1) physicians have a good understanding of the incentive system and (2) communicate with each other directly. We have, however, no direct evidence that about these points. We intend to investigate these issues further by surveying physicians in the HMO's network.

Perhaps the most intriguing results of the analysis are those relating to quality. What explains our finding that low cost panels are also high in measured quality? The HEDIS quality measures emphasize preventive care and reward those panels that are best at educating and tracking their enrollees. These same organizational capacities may also help panels control costs. Alternatively, it may be that low cost panels have high measured quality because they are especially committed to winning incentive payouts. These panels, in other words, may find ways to make sure their flu immunization and mammography rates are high without otherwise improving the true quality of care they deliver.

Distinguishing between these two possibilities requires that we compare cost outcomes to quality measures that are *not* included in the incentive formula as well as to those that are included in the formula. More generally, the quality measurements available in 1997 were both limited and crude. The results so far suggest that further analysis of the determinants of measured quality may be worthwhile.

BIBLIOGRAPHY

Blair, Douglas A. 1998. "The 'PIP' Regulations in Perspective: Analysis and Comparisons With Other Federal Regulations Governing Physician Incentive Plans," *Memphis State University Law Review,* 137.

Dudley, R.A., R.H. Miller, T.Y. Korenbrot, and H.S. Luft. 1998. "The Impact of Financial Incentives on Quality of Health Care," *Milbank Quarterly,* 76: 649-86.

Encinosa, W., Gaynor, M. and Rebitzer, J. 1999. "The Sociology of Groups and the Economics of Incentives: Theory and Evidence On Incentives in Medical Groups,".

Epstein, Arnold M., Begg, Colin B., and Barbara J. McNeil. 1986. "The Use of Ambulatory Testing in Prepaid and Fee-for-Service Group Practices," *New England Journal of Medicine*, April 24, 314 (17):1089-1094.

Freudenheim, Milt. 1999. "Humana Sued in Federal Court Over Incentives for Doctors," *New York Times*, October 5, Section C, p. 7, column 1.

Gaynor, Martin, and T. Mark. 1999. "Physician Contracting with Health Plans: A Survey of the Literature," unpublished manuscript, Carnegie Mellon University.

Gaynor, Martin, and Gertler, Paul. 1995. "Moral Hazard and Risk Spreading in Medical Partnerships," *Rand Journal of Economics,* 26:4 p. 591-613.

Gaynor, Martin, and Pauly, Mark V. 1990. "Compensation and Productive Efficiency in Medical Partnerships: Evidence from Medical Group Practice," *Journal of Political Economy*, 98:3 p. 544-573

Gold, M. 1999. "Financial Incentives in Managed Care: Current Realities and Challenges for Physicians," *Journal of General Internal Medicine* (forthcoming).

Gold, M.R., R. Hurley, T. Lake, et al. 1995a. "A National Survey of the Arrangements Managed Care Plans Make with Physicians," *New England Journal of Medicine,* 333:1678-83.

Hadley, J. and J. Mitchell. 1997. "Effects of HMO Market Penetration on Physicians' Work Effort and Satisfaction," *Health Affairs*, 16(6), 99-111.

Tversky, Amos and Kahneman, Daniel. 1991 "Loss Aversion in Riskless Choice: A Reference Dependent Model" Quarterly Journal of Economics 106**:** 4 p. 1039-1061.

Gold, Marsha, Lyle Nelson, Timothy Lake, Robert Hurley, and Robert Berenson. 1995. "Behind the Curve: A Critical Assessment of How Little Is Known About Arrangements Between Managed Care Plans and Physicians," *Medical Care Research and Review*, September.

Hillman, Allan L., Mark V. Pauly, Keith Kerman, and Caroline Rohr Martinek. 1991. "HMO Managers' Views on Financial Incentives and Quality," *Health Affairs*, Winter, p. 208-219.

Hillman A.L., Pauly M.V., Kerstein J.J. 1989. "How do Financial Incentives Affect Physicians' Clinical Decisions and the Financial Performance of Health Maintenance Organizations," *New England Journal of Medicine*, 321, 2, 86-92.

Hillman, Alan L., Welch, Pete and Pauly, Mark V. 1992. "Contractual Arrangements Between HMOs and Primary Care Physicians: Three Tiered HMOs and Risk Pools," *Medical Care*, 30:2, 136-148.

Kandel, Eugene and Lazear, Edward P. 1992 . "Peer Pressure and Partnerships," *Journal of Political Economy* 100: 4, p. 801-817.

Knez, Marc and Simester, Duncan. "Firm-Wide Incentives and Mutual Monitoring At Continental Airlines" June 1998, Mimeo,

Milgrom, John and Roberts, Paul 1992 Economics, Organizations and Management

Miller, R.H., and H.S. Luft. 1994. "Managed Care Performance Since 1980: A Literature Analysis." *Journal of the American Medical Association,* 271:1512-19.

Prendergast, Canice. 1999. "The Provision of Incentives in Firms," *Journal of Economic Literature*, 37:1, p. 7-63.

Rubin, Alissa J. and Henry Weinstein. 1999. "As Smoke Clears, Lawyers Take on Health Insurers," *Los Angels Times*, October 9, Part A, Page 1.

Simon, C.J. W.D. White, S. Gamliel, and P.R. Kletke. 1997. "The Provision of Primary Care: Does Managed Care Make a Difference?" *Health Affairs*, 16, 6, 89-98.

Weitzman, Martin L. and Douglas L. Kruse. 1990. "Profit Sharing and Productivity" in *Paying for Productivity: A Look at the Evidence*, Alan S. Blinder, ed., The Brookings Institution: Washington, DC.

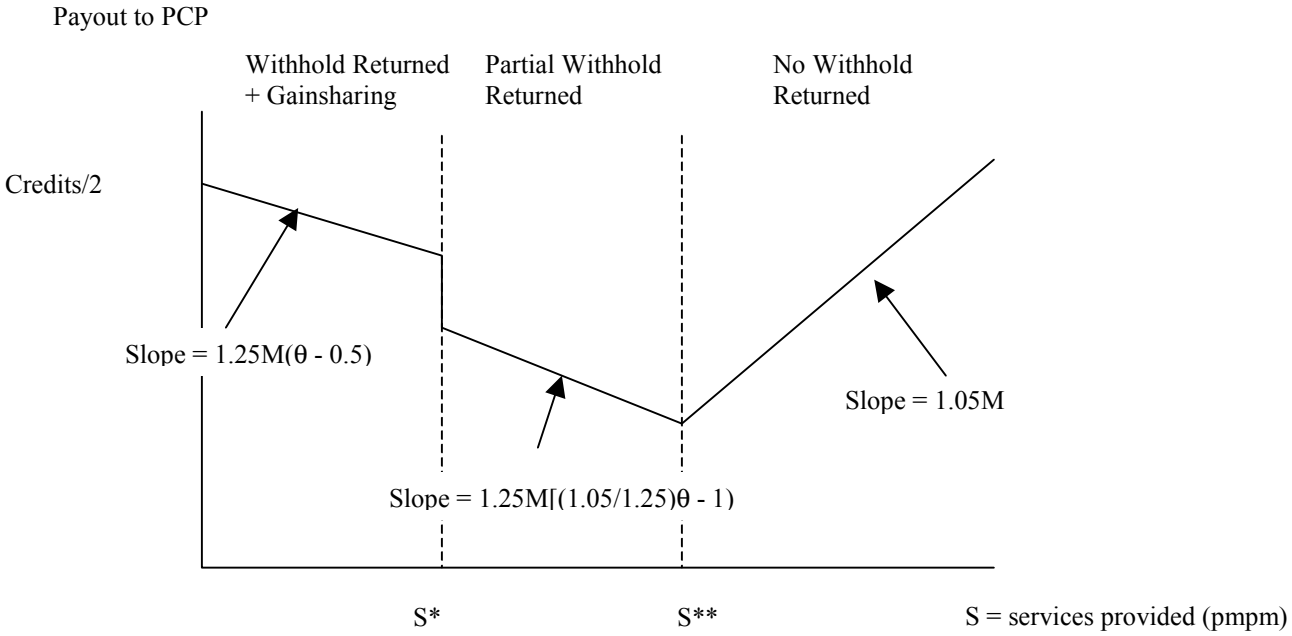

Figure 1
The HMO Payment Schedule 1994-1996

Notation: S is the quantity of medical services provided. For simplicity think if S as a single service. θ = fraction of Debits accruing as income to PCPs. Debits are total medical services (per member per month) charged against credits in enrollees' patient care account. Debits = PMS where M = Medicare price for S and P = HMO's premium over M. To the left of S*, P = 1.25 and to the right of S*, P = 1.05. S* is the value of S such that Credits = 1.25MS where Credits is the number of credits (per member per month) assigned to the patient population. S** is the value of S where the deficit in the patient care account equals the total of withheld income or 0.2(1.25MSθ)

Withhold is returned and gain-sharing is paid out when debits- stop loss expenses < credits. For simplicity, we assume no stop loss expenses, i.e. that expenses did not exceed $15,000 for an enrollee in the year. The region of the figure to the left of S* physician income is then:
PCP Income = 1.25MSθ + 0.5[Credits -1.25MS] = 1.25MS(θ-0.5) +0.5Credits

In the region between S* and S**, the withhold is partially returned. Specifically, for every dollar of deficit, physicians looses a dollar of withhold. Physician income is
PCP Income = 1.05MSθ + [0.2(1.25M(S*)θ)-(1.25MS-Credits)]
= 1.25MS[(1.05/1.25)θ -1] + .25MS*θ + Credits

In the region to the right of S**
PCP Income = 1.05MSθ

As a matter of algebra $0<\theta<1$. In the data, we find average θ is roughly 0.2.

Table 1
Determinants of Medical Utilization Costs
1994-1997

	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)	Fixed Effect (7)	Fixed Effect (8)	Fixed Effect (9)
Dependent Variable	Total Medical Expenditures (pmpm)	Expenditures Net Stop Loss	Fraction Withhold Returned	Total Medical Expenditures (pmpm)	Expenditures Net Stop Loss	Fraction Withhold Returned	Total Medical Expenditures (pmpm)	Expenditures Net Stop Loss	Fraction Withhold Returned
Log Physicians in POD	36.168	31.911	-0.269	27.346	24.494	-0.272	46.102	46.39	-0.314
	(2.955)	(3.544)	(3.157)	(2.290)	(2.937)	(3.420)	(3.017)	(3.569)	(2.519)
Log Physicians in POD*1997	-29.669	-29.808	0.12	-26.879	-27.531	0.149	-27.862	-23.459	0.117
	(2.507)	(3.208)	(1.245)	(2.129)	(2.697)	(1.628)	(1.928)	(1.908)	(0.995)
Log Member-months	-10.335	-10.375	0.096	-6.545	-6.631	0.041	-3.872	-9.11	-0.058
	(2.632)	(3.334)	(3.542)	(1.485)	(2.185)	(1.260)	(0.682)	(1.886)	(1.250)
Log Membermonths*1997	2.921	3.037	0.015	3.323	3.199	0.001	2.142	1.361	0.042
	(0.638)	(0.761)	(0.341)	(0.655)	(0.725)	(0.018)	(0.395)	(0.295)	(0.945)
Target Expenses (pmpm)	0.99	0.823	-0.001	0.756	0.66	0.001	0.6	0.536	0.001
	(6.938)	(6.074)	(0.904)	(3.629)	(3.184)	(1.177)	(5.829)	(6.121)	(1.405)
Years POD is in Sample	9.111	6.615	-0.025	6.443	4.962	-0.013	28.607	23.679	-0.549
Since 1994	(2.434)	(2.028)	(0.534)	(1.409)	(1.283)	(0.254)	(0.978)	(0.952)	(2.301)
Year is 1995	-14.516	-4.418	-0.011	-12.198	-5.741	-0.013	-30.865	-19.133	0.592
	(1.651)	(0.667)	(0.148)	(1.356)	(0.895)	(0.166)	(1.015)	(0.740)	(2.387)
Year Is 1996	-51.834	-29.038	0.051	-47.079	-28.777	0.048	-88.932	-60.931	1.204
	(4.481)	(3.201)	(0.492)	(3.754)	(2.740)	(0.445)	(1.524)	(1.227)	(2.530)
Year Is 1997	-8.909	4.868	-0.4	-17.176	-4.726	-0.283	-71.918	-50.294	1.085
	(0.244)	(0.149)	(1.117)	(0.437)	(0.130)	(0.821)	(0.948)	(0.779)	(1.754)
Single Specialty POD				-21.586	-21.54	0.031			
				(2.341)	(2.808)	(0.423)			
POD In Network 1				-13.416	-11.637	0.528			
				(1.564)	(1.671)	(5.439)			
POD In Network 2				-31.11	-31.212	0.598			
				(2.635)	(3.315)	(3.991)			
POD In Network 3				29.138	30.59	-0.23			
				(1.298)	(1.879)	(1.943)			
POD In Network 4				-7.263	-11.801	0.266			
				(0.585)	(1.730)	(2.233)			
POD In Network 5				1.946	1.535	0.191			
				(0.280)	(0.229)	(1.243)			
Fraction POD Family Practice				6.798	-1.368	0.213			

Table 1
Determinants of Medical Utilization Costs
1994-1997

Dependent Variable	OLS (1) Total Medical Expenditures (pmpm)	OLS (2) Expenditures Net Stop Loss	OLS (3) Fraction Withhold Returned	OLS (4) Total Medical Expenditures (pmpm)	OLS (5) Expenditures Net Stop Loss	OLS (6) Fraction Withhold Returned	Fixed Effect (7) Total Medical Expenditures (pmpm)	Fixed Effect (8) Expenditures Net Stop Loss	Fixed Effect (9) Fraction Withhold Returned
				(0.386)	(0.093)	(1.148)			
Fraction POD Internal Medicine				26.15	15.269	-0.071			
				(1.564)	(1.070)	(0.397)			
Fraction POD Pediatrics				1.452	1	0.12			
				(0.093)	(0.081)	(0.629)			
POD Only OB/GYN				13.946	13.407	0.027			
				(0.952)	(1.147)	(0.152)			
% Medicaid > 10 in 1997				8.559	11.312	0.008			
				(1.142)	(1.790)	(0.083)			
Constant	33.042	37.122	0.374	41.181	43.924	0.483	-22.186	8.612	1.827
	(1.072)	(1.398)	(1.944)	(1.065)	(1.283)	(1.796)	(0.577)	(0.263)	(5.801)
Observations	353	353	349	340	340	339	353	353	349
Number of PODs	118	118	116	109	109	109	118	118	116
Adjusted R-squared	0.457	0.464	0.066	0.454	0.488	0.154	-0.058	-0.039	-0.384
Absolute value of robust t-statistics in ( ) for OLS estimates.									
F test that Log Physicians in POD*1997 and Log Membermonths*1997 are jointly zero	3.56	5.47	1.22	2.55	3.83	1.52	2.3	2.38	2.34
p value	0.0317	0.0054	0.299	0.0825	0.0247	0.223	0.1021	0.0952	0.0986

All equations estimated using data from 1994-1997.
There are 120 panels of physicians (or PODs) in the data. The sample is an unbalanced time-series/cross-section with 1 to 4 years of observations per panel. The median POD was in the sample for 3 years.

Table 2
Selction Vs. Incentives:
Do High Cost Physicians Exit Shrinking Panels and Do High Cost Physicians Enter Growing Panels?

Panel A

Physician Exits Panel After Year t		Number of Physicians in POD Declines from Year t to Year t+1: no	yes	All
yes	Relative Costs Year t	1.02	1.02	1.02
	Number Physicians in Cell	257	41	298
no	Relative Costs Year t	0.99	1.003	0.995
	Number of Physicians in Cell	678	338	1016
All	Relative Costs Year t	0.998	1.005	1.000
	Number of Physicians in Cell	935	379	1314

Panel B

Physician Enters in Year t		Number of Physicians in POD Increases from Year t-1 to Year t: no	yes	All
no	Relative Costs Year t	1.00	0.98	0.99
	Number Physicians in Cell	678	338	1016
yes	Relative Costs Year t	1.06	1.01	1.01
	Number of Physicians in Cell	45	701	746
All	Relative Costs Year t	1.00	1.00	1.00
	Number of Physicians in Cell	723	1039	1762

Relative Costs is the average of a physician's total medical expenses in year t to the panel's total medical expenses in year t.
None of the cross-cell difference in the panel are statistically significant.

In panel A, t refers to years 1994 and 1995. In panel B, t refers to years 1995 and 1996.

In panel A, of the physicians in the sample, 298 exited the panel after years 94 or 95. The relative costs of these exiters was 2% above the mean of their POD for both shrinking and growing PODs.

In panel B, of the physicians in the sample, 746 entered the panel in years 95 or 96. The relative costs of these entrants was 1% above the mean of their POD. This 1% differential also held for growing PODs

Table 3
Determinants of Medical Utilization Costs by Category
1994-1996

Dependent Variable	OLS (1) Total Medical Expenditures (pmpm)	OLS (2) Expenditures Not Paid to Self (pmpm)	OLS (3) Expenditures Paid to Self (pmpm)	OLS (4) Total In-patient (pmpm)	OLS (5) Total Out-Patient (pmpm)
Log Physicians in POD	26.227	24.062	2.165	5.055	9.512
	(2.110)	(1.932)	(1.364)	(0.674)	(2.020)
Log Member-months	-5.818	-5.31	-0.508	-1.691	-3.543
	(1.191)	(1.097)	(1.072)	(0.461)	(2.198)
Target Expenses (pmpm)	0.627	0.55	0.077	0.224	0.31
	(2.156)	(2.334)	(1.274)	(3.331)	(1.852)
Years POD is in Sample	11.319	9.64	1.679	5.162	5.583
Since 1994	(1.250)	(1.144)	(1.420)	(1.356)	(1.156)
Year is 1995	-16.521	-14.995	-1.527	-14.621	-0.905
	(1.512)	(1.409)	(1.361)	(2.353)	(0.201)
Year Is 1996	-54.104	-50.957	-3.147	-30.488	-9.794
	(2.918)	(2.910)	(1.264)	(4.439)	(1.030)
Single Specialty POD	-37.389	-30.139	-7.25	-5.325	-14.649
	(2.355)	(2.129)	(2.873)	(0.840)	(1.823)
POD In Network 1	-17.127	-14.158	-2.969	-7.119	-1.808
	(1.697)	(1.495)	(1.251)	(1.100)	(0.451)
POD In Network 2	-41.323	-35.253	-6.07	-7.924	-13.631
	(2.508)	(2.236)	(3.037)	(0.827)	(1.504)
POD In Network 3	35.323	38.888	-3.566	-0.186	26.937
	(1.118)	(1.315)	(0.934)	(0.008)	(2.432)
POD In Network 4	-8.68	-5.651	-3.029	5.253	-3.263
	(0.506)	(0.336)	(2.116)	(0.399)	(0.641)
POD In Network 5	4.086	-1.338	5.424	-8.356	6.642
	(0.406)	(0.150)	(1.641)	(1.145)	(1.009)
Fraction POD Family Practice	14.118	29.622	-15.504	5.794	18.153
	(0.449)	(1.075)	(1.938)	(0.494)	(1.535)
Fraction POD Internal Medicine	35.361	51.305	-15.944	13.718	25.028
	(1.156)	(1.923)	(2.038)	(0.995)	(2.153)
Fraction POD Pediatrics	7.923	18.158	-10.235	0.756	9.8
	(0.240)	(0.631)	(1.275)	(0.052)	(0.860)
POD Only OB/GYN	24.719	31.365	-6.646	11.907	15.182
	(0.788)	(1.155)	(0.814)	(0.891)	(1.355)
% Medicaid > 10 in 1997	9.924	8.017	1.907	0.381	1.767

Table 3
Determinants of Medical Utilization Costs by Category
1994-1996

	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)
Dependent Variable	Total Medical Expenditures (pmpm)	Expenditures Not Paid to Self (pmpm)	Expenditures Paid to Self (pmpm)	Total In-patient (pmpm)	Total Out-Patient (pmpm)
	(0.930)	(0.801)	(1.246)	(0.066)	(0.514)
Constant	45.739	22.574	23.165	16.444	2.677
	(0.785)	(0.443)	(2.135)	(0.820)	(0.098)
Observations	236	236	236	236	236
Number of PODs	99	99	99	99	99
Adjusted R-squared	0.42	0.414	0.393	0.176	0.472

Absolute value of robust t-statistics in ( ) for OLS estimates.

All equations estimated using data from 1994-1997.
There are 120 panels of physicians (or PODs) in the data. The sample is an unbalanced time-series/cross-section with 1 to 4 years of observations per panel. The median POD was in the sample for 3 years.

Table 4
Physician Fixed Effects Estimates of Utilization Costs

Dependent Variable	(1) Total Medical Expenditures (pmpm)	(2) Total In-patient (pmpm)	(3) Total Out-Patient (pmpm)	(4) Total Medical Expenditures (pmpm)	(5) Total In-patient (pmpm)	(6) Total Out-Patient (pmpm)
Log Member-months	-10.821	-0.646	-3.13	-8.555	0.416	-2.842
For Individual Physician	(2.086)	(0.183)	(2.741)	(1.655)	(0.118)	(2.482)
Log Member-Months for Others in POD	-3.18	8.008	-4.207	-33.336	-6.124	-8.043
	(0.507)	(1.878)	(3.050)	(3.562)	(0.957)	(3.880)
Log Number of Other Physicians in POD				72.403	33.93	9.209
				(4.309)	(2.955)	(2.474)
Target	-0.483	-0.371	0.053	-0.543	-0.399	0.046
	(1.709)	(1.929)	(0.855)	(1.935)	(2.080)	(0.734)
Year is 1995	-3.313	-14.625	3.932	-0.887	-13.488	4.24
	(0.333)	(2.159)	(1.793)	(0.090)	(1.995)	(1.935)
Year is 1996	-42.524	-42.333	5.399	-40.215	-41.251	5.692
	(2.089)	(3.058)	(1.205)	(1.992)	(2.990)	(1.273)
Constant	203.931	-6.922	74.436	290.185	33.499	85.407
	(4.043)	(0.202)	(6.704)	(5.387)	(0.910)	(7.159)
Fixed Physician Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2059	2059	2059	2059	2059	2059
Number of Physicians in Sample	1042	1042	1042	1042	1042	1042
R^2	0.0204	0.012	0.0291	0.038	0.0207	0.035

Absolute value of t-statistics in ( )
F test that both Log Member-months variables are jointly zero
R^2 refers to the fraction of "within" physician variation explained by the equation.

Table 5
Determinants of HEDIS Quality Measures
1997

Dependent Variable	(1) Total Medical Expenditures (pmpm)	(2) Stop Loss	(3) Expenditures Net Stop Loss	(4) HEDIS Quality Points	(5) HEDIS Quality Points
HEDIS Quality Points	-0.805				
	(3.744)				
HEDIS Quality Points		-0.666			
		(3.343)			
HEDIS Quality Points			-0.138		
			(1.532)		
Log Physicians in POD				-6.628	-5.663
				(2.857)	(2.157)
Log Member-months				3.403	1.575
				(4.099)	(1.175)
Target Expenses (pmpm)				-0.095	-0.03
				(1.779)	(0.555)
Years POD is in Sample				4.529	4.425
Since 1994				(3.487)	(2.889)
Single Specialty POD					-1.999
					(0.705)
POD In Network 1					-1.655
					(0.330)
POD In Network 2					12.455
					(3.007)
POD In Network 3					-9.259
					(2.143)
POD In Network 4					9.001
					(2.465)
POD In Network 5					9.316
					(2.522)
Fraction POD Family Practice					-2.255
					(0.406)
Fraction POD Internal Medicine					1.836
					(0.373)
Fraction POD Pediatrics					-2.052
					(0.369)
POD Only OB/GYN					7.003
					(1.092)
% Medicaid > 10 in 1997					-5.581
					(2.120)
Constant	79.875	67.769	12.106	-7.409	3.155
	(15.141)	(13.349)	(5.444)	(1.215)	(0.312)
Observations	107	107	107	107	104
Adjusted R-squared	0.101	0.096	0.009	0.287	0.309

Absolute value of robust t-statistics in ( ) for OLS estimates.

Data Appendix 1
Variable Names and Descriptions
(Tables 1, 3, 5)

Variable Name	Mean	Std. Dev.	Min	Max	Notes
Dependent Variables					
Total Medical Expenditures (pmpm)	89.56	65.60	8.82	583.09	The sum of expenditures due to in-patient and out-patient procedures, professional services and "other" expenditures, minus the amount paid by other insurers.
Total In Patient Expenditures (pmpm)[#]	28.76	33.46	0.00	263.99	sub-category of total medical expenses available for 1994-1996
Total Out Patient Expenditures (pmpm)[#]	35.39	31.43	5.27	306.48	sub-category of total medical expenses available for 1994-1996
Fraction Withhold Returned	0.50	0.49	0.00	1.00	Fraction of withheld fees returned to POD at end of year
Expenditures Paid to Self (pmpm)[#]	14.49	11.28	0.00	107.52	Utilization Expenses that Accrue as Income to the referring PCP's
Expenditures Not Paid to Self (pmpm)[#]	84.83	67.75	8.82	475.57	Total Medical Expenses - Expenditures Paid to Self
Hedis Quality Points[@]	16.08	13.52	0.00	50.00	
Other Variables					
Log Physicians in POD	2.16	0.53	1.10	3.56	Log of the number of physicians in POD.
Log Member-months	8.48	1.69	2.71	11.27	Log of the number of HMO membermonths at the panel
Single Specialty POD	0.60	0.49	0.00	1.00	POD physicians all in one specialty: either internal medicine, family practice, pediatrics or ob/gyn
Target Expenses (pmpm)	79.42	36.51	14.92	265.79	Level of per member per month expenditures at which withhold is returned and below which bonuses are paid.
Stop Loss Expenses (pmpm)					Medical expenses in excess of $15,000 for patients having more than $15,000 in expenses divided by the number of patient member months in the panel.
% Medicaid > 10 in 1997	0.30	0.46	0.00	1.00	Dummy variable = 1 when % medicaid > 10% of panel's patient population. HMO began accepting Medicaid in Feb. 1995, but it remained very small (and wasn't counted separately) until 1997.
Year is 1995	0.25	0.43	0.00	1.00	Dummy variable = 1 when year is 1995
Year Is 1996	0.26	0.44	0.00	1.00	Dummy variable = 1 when year is 1996
Year Is 1997	0.31	0.46	0.00	1.00	Dummy variable = 1 when year is 1997
Years POD is in Sample Since 94	2.27	1.04	1.00	4.00	
POD In Network 1	0.03	0.18	0.00	1.00	Dummy variable indicating panel is part of larger network
POD In Network 2	0.03	0.18	0.00	1.00	Dummy variable indicating panel is part of larger network
POD In Network 3	0.06	0.24	0.00	1.00	Dummy variable indicating panel is part of larger network

Data Appendix 1
Variable Names and Descriptions
(Tables 1, 3, 5)

Variable Name	Mean	Std. Dev.	Min	Max	Notes
POD In Network 4	0.17	0.38	0.00	1.00	Dummy variable indicating panel is part of larger network
POD In Network 5	0.08	0.27	0.00	1.00	Dummy variable indicating panel is part of larger network
Fraction POD Internal Medicine	0.30	0.41	0.00	1.00	
Fraction POD Pediatrics	0.26	0.39	0.00	1.00	
Fraction POD Family Practice	0.25	0.38	0.00	1.00	
POD Only OB/GYN	0.11	0.32	0.00	1.00	

Data available for 1997 only
@ Data available for 1994-96 only

Data Appendix 2

Quality Incentives 1997

Quality incentives were the sum of points awarded for two types of quality measures: quality of care/preventative care measures; and other measures of patient satisfaction. The first category was referred to within the HMO as HEDIS measures. These measures varied by specialty.

HEDIS Quality Measures (50 points maximum)

Pediatrics:

- One point given for each percentage point above 70% compliance with age specific immunizations for children aged birth to 2 years.
- One point given for each percentage point above 60% compliance with annual well child examination for children aged 3-5 years.

Internal Medicine:

- One point given for each percentage point above 60% of a mammogram in the past 24 months in women age 50-65.
- One point given for each percentage point above 50% of patients with high risk medical conditions or age 65+ having received an annual influenza vaccine.

Family Practice:

- One point given for each percentage point above 70% compliance with age specific immunizations for children aged birth to two years.
- One point given for each percentage point above 60% of a mammogram in the past 24 months in women age 50-65.

OB/GYN

- One point given for each percentage point above 60% of a mammogram in the past 24 months in women age 50-65.
- One point given for each percentage point above 65% of women 18-65 with a Pap test in the past three years.

Non- HEDIS Quality Measures (50 points maximum + Medicaid Supplement)

- Panels given 10 points and then points were deducted for families requesting a change in PCP for perception of poor quality of care or poor provider/patient relationships. Deductions varied with size of panel. In 1997, problems with indicator led to all panels being assigned 10 points. This indicator was improved and continues to be used in the HMOs incentive system.
- Points awarded based on patient satisfaction results from the annual patient satisfaction survey. This measure proved to be uninformative – all but 2 panels in 1997 got the full 10 points for this measure.
- Points awarded to PCPs offering regular office hours outside of M-F, 8AM to 5PM. In 1997 problems with indicator led to all panels being assigned 10 points.
- Fraction of physicians open for accepting new patients from the HMO times 10.
- Points awarded on the basis of a focused office review, a kind of "white glove" inspection of office facilities including checks of medical records.
- 10 quality points awarded to PCPs with more than 10% of their enrollees in Medicaid.