Discussion Paper No. 959

COMPETITION IN A MARKET FOR INFORMED EXPERTS' SERVICES

by

Asher Wolinsky

Department of Economics
Northwestern University

September 1991

I thank Kyle Bagwell and Mike Riordan for useful comments. I am particularly grateful to Michael Lynch who referred me to the work that was done in the FTC on these subjects.

COMPETITION IN A MARKET FOR INFORMED EXPERTS' SERVICES

ABSTRACT

Many important services share the feature that the seller is also the expert who determines how much of the service is needed. Even when the outcomes of such service are observable, it might be difficult for the customer to determine what the expert actually did and whether it was needed. This paper presents a simple model of a market of this type and investigates how the information asymmetries characteristic of such markets might affect their organization.

The main insights of this paper are as follows. The asymmetry of information special to these markets may induce vertical specialization. When experts are liable to make diagnosis errors, there is a negative search externality present in such markets which tends to raise prices. The search-cum-diagnosis costs and the accuracy of diagnoses play a clear role in the determination of the market's form of organization: when the former are low and the latter is high, the market is more likely to be organized in a way whereby experts provide binding estimates in advance and consumers search; otherwise the more likely organization is that customers are billed after the service was performed and experts are disciplined by reputation.

Competition in a Market for Informed Experts Services

1. Introduction

Medical and legal services as well as other less glorified repair services share the feature that the seller is also the expert who determines how much of the service is needed. Even if the success of performing a service of this type is easy to determine and observe, it is often difficult for the customer to determine the extent of the service that the expert indeed performed and whether it was needed. This information asymmetry creates obvious incentives for opportunistic behavior by the sellers[1], and these in turn presumably prompt the emergence of the mechanisms which discipline sellers' behavior. The purpose of this paper is to present a simple model and use it to investigate how the information asymmetries characteristic of such markets may affect their organization.

Markets of this type involve a number of informational problems. Besides the problem of honest diagnosis just mentioned, the customer may be unable to evaluate the results of the service even after its completion, and, for a given diagnosis, the service may have quality dimensions affected by the seller's unobservable ability or effort. We shall focus here on the first problem, which seems most special to such markets, and abstract from the others. Thus, the main feature of the markets we consider is that, while customers can determine whether their problem was fixed, they cannot determine how much treatment they received. The car repair industry seems to fit this description.

The manner in which experts' incentives to misrepresent minor treatments as major ones translate into behavior will depend of course on the market organization and customers' behavior. We shall discuss two alternative settings in which market forces mitigate these incentives to some extent. The search setting is such that experts diagnose customers and provide them with binding estimates, so that customers can meaningfully search, prior to making a decision. In the reputation setting customers do not get binding estimates prior to their treatment. Instead, they keep some kind of record of experts' performance and punish experts who seem to overcharge by withdrawing their business. Note, for example, that the car repair market features some mixture of both systems.

The model we adopt assumes only two possible types of problem--a major, high cost, problem (denoted H) or a minor, low cost, one (denoted L). In the first parts of this paper we consider a market which operates along the lines of the search setting. Customers can visit experts who diagnose them and offer a level of treatment, and in the process they incur a fixed search-cum-diagnosis cost per each expert they sample. The main result here is that, when this cost is not too high, the market equilibrium (which is "essentially" unique when there are sufficiently many experts) involves complete specialization. Some experts provide only the minor treatment; others provide only the major treatment; the competition drives the prices of both services to their respective costs; customers start by sampling a minor treatment expert and if rejected continue to a major treatment expert. The experts' incentives to cheat are removed because experts who specialize in the minor treatment have clearly nothing to gain from misrepresenting--they anyway lose the business of those customers they diagnose as requiring the major treatment. The equilibrium specialization imposes costs in that customers who require the major treatment will end up going to two experts and being diagnosed twice. Thus, while there is no "cheating" in equilibrium, the asymmetry of information is translated into expert specialization and hence into the costs of excessive search and diagnosis.

We also consider the case in which experts are liable to make errors in their diagnosis. This feature adds another source of asymmetric information, and may induce customers to try to take advantage of experts' errors. The main difference between the specialization equilibrium discussed above and its counterpart here is that, when the search-cum-diagnosis cost is sufficiently low, customers will search more vigorously to take advantage of experts' diagnosis errors. Viewed from the side of the experts, this creates a "winner's curse" which translates into higher equilibrium prices for the minor treatment. Viewed from the side of the customers, this is a search externality inflicted by customers who require the major treatment on those who require only the minor treatment and end up paying higher prices. The added information imperfection exacerbates the welfare costs, since customers search on the average longer.

Finally, we consider an alternative arrangement in which experts' behavior is disciplined by reputation. Here, experts do not guarantee the cost in advance--customers entrust their problem to an expert who first fixes it and then presents them with a bill. We capture this by a minimal modification of the model: each customer requires the service exactly twice. In equilibrium the price of the minor treatment will embody a sufficient mark-up which, together with the customers' punishing behavior, will deter experts from cheating. The main question is what determines which form of market organization--the search or the reputation settings--is more likely to arise. In the context of the simple model considered here it turns out that, other things equal, lower search-cum-diagnosis costs and greater precision of the diagnosis will tend to favor the search scenario. That is, with low cost and high precision, the reputation equilibrium will be destroyed by experts who specialize in the minor treatment and offer to commit to the cost of treatment in advance. This suggests that in a market, such as the car repair market, in which both the reputation and search systems coexist, the former is more likely to cater to customers with higher search costs and to deal with problems which are harder to diagnose in advance.

The theoretical literature on markets of this type is not very large. Arrow (1963) discusses informational aspects of the market for medical services. His discussion is much broader and does not address in detail the alternative ways whereby competition and monitoring devices determine behavior in such market. Darby and Karni (1973) term such goods "credence goods" and discuss how reputation combined with market conditions and technological factors affect the amount of fraud. Plott and Wilde (1980) characterize customers' optimal search in such a market, but do not attempt an equilibrium analysis; Plott and Wilde (1991) report laboratory experiments[2] designed to mimic such a market . Pitchik and Schotter (1987, 1989) are closer to the present paper both in terms of the model and its focus. Their first paper considers a mixed strategy equilibrium in the expert-customer game in which experts randomize between reporting truthfully or not and customers randomize between accepting and rejecting a major treatment recommendation (see section 4 below). Their second paper models in greater detail the process of search by customers. The present model differs from theirs in its competitive interaction through explicit price competition and in the search and reputation mechanisms considered.

In a less direct way this paper is also related to the literature on product quality provision under conditions of asymmetric information (see, e.g., Klein and Leffler (1981), Wolinsky(1983)). There too a better informed seller faces less informed buyers and some of the main questions concern the manner in which competition and monitoring devices such as search or reputation interact to determine the prices and the range of products. The special nature of the services considered here separates this paper from that literature. On the demand side the customers in the present model do not know what service they are getting even after they have received it--their uncertainty is not about the benefit they will derive from the purchase, but rather about the benefit they may obtain elsewhere. On the supply side there is a role for economies of scope--the search-cum-diagnosis cost makes it more efficient to have one expert provide a range of services.

2. The model

Each customer in this market has a problem which can be of one of two types: major or minor. A customer knows that he has a problem, but does not know how serious it is. An expert can diagnose and fix the problem. The cost of fixing the major problem is H (for "high") and the cost of fixing the minor one is L (for "low"), $L<H$. The existence of a problem is both observable and verifiable, but the type of treatment (H or L) is not observable to customers. Therefore, agreements that condition payments on the resolution of the problem, are in principle enforceable, while payments conditional on the treatment type are not. This last feature will make cheating possible: an expert can try to misrepresent a minor treatment as a major one.

The interaction is modelled as follows. Time is divided into discrete periods and has no beginning or end (i.e., time goes from $-\infty$ to ∞). At the beginning of each period a new batch of customers enters the market. Their measure is M and fractions w and (1-w) have the H and L problems respectively. They join the pool of customers who were left from previous periods. Then the N experts present in the market simultaneously make their three decisions for this period: (i) the menu of treatments they provide (whether both L and H or just L); (ii) their prices; (iii) their recommendation policies (what treatment to recommend as a function of diagnosis). An expert's decision to offer both treatments at the prices P_L and P_H respectively will be denoted (P_L, P_H); the decision to provide only service L at the price P_L will be denoted (P_L, ∞). (It is thus assumed that, if an expert decides to have the ability to perform the H service, this will automatically include the ability to perform the L service.) Next, customers observe the menu and price decisions of experts, but not their recommendation policies, and each customer chooses an expert to visit. An expert diagnoses each visiting customer and recommends either the H-treatment (the major) or the L-treatment (the minor) according to his policy. If the customer accepts this recommendation, the expert will fix the problem for the price he had quoted for the recommended level of treatment and the customer will leave the market. If the customer does not accept, he will proceed to the next period.

A customer incurs a cost k per each expert that he samples, independently of whether or not he chooses to be treated by this expert. This cost accounts for the time and effort incurred in going to an expert. Also, if experts incur costs in performing the diagnosis, it is assumed that k already includes diagnosis fees which exactly cover these costs (this assumption is discussed in section 4).

The utility to a customer who visited n experts and ended up being treated for the price P is B-P-nk, where B is the value of having the problem fixed. Customers are maximizers of expected utility. The reservation value B is assumed large enough (say, well above H+k) so as to assure that participation is always desirable for customers. In a given period, an expert's profit is the sum of revenues minus costs over the customers he treated in that period. Experts are assumed myopic: in each period they seek to maximize their expected profit in that period.

Although this model has a time dimension, there will be no strategic links between periods. As implied by the assumption on experts' myopia, experts will behave as in the static model. The reason we introduce the time dimension is to assure that the distribution of customers in this market will be consistent with the fact that as customers with different problems may have different search durations[3].

Finally, we think of the number of experts, N, as a finite large number, and assume that experts are anonymous in the sense that, if a customer decides to sample an expert who chose (P_L, P_H), he will sample one of those experts with equal probability. We shall think of the customer population as a "continuum without aggregate uncertainty," in the sense that probabilities will be identified as exact fractions of the population[4].

A more formal description of the interaction is as follows. A <u>strategy</u> for an expert is a triple $\{(P_L, P_H), x\}$, where P_H can assume the value ∞ as well, and the recommendation policy $x \in [0,1]$ is the probability that this expert will recommend the H-treatment when the diagnosis indicates that only the L-treatment is needed. Since experts will never have an incentive to misrepresent an H treatment as an L treatment, we just assume this possibility away and let the recommendation policy be captured by a single probability.

A customer's information in each period includes the experts' observable choices in that period--menus of treatments and prices--and the updated probability of the type of problem he has. This probability is computed from the prior probability and the past search experience. A customer's strategy prescribes, for each possible information, whether or not to accept the last recommendation and, if not, where to continue the search. Customers will form beliefs over the unobservable experts' recommendation policies.

A steady state is such that the distribution of experts' and customers' strategies are constant from period to period, and the flow of entering customers, M, is exactly matched by the flow of departing customers.

The equilibrium notion, to be called market equilibrium (ME), is a steady state in which experts' strategies $\{(P_L^i, P_H^i), x^i\}_{i=1,...N}$, and customers' strategies and beliefs satisfy:

(i) For each i, $\{(P_L^i, P_H^i), x^i\}$ maximizes expert i's one period's profit, given $\{(P_L^j, P_H^j), x^j\}_{j \neq i}$ and the customers' strategies.

(ii) Each customer's strategy is optimal, for any possible information, given his beliefs.

(iii) In equilibrium the customers' beliefs are confirmed.

Before proceeding, let us remove some potential ambiguities by emphasizing the degree of commitment expressed in experts' decisions (we shall return to these points in Section 4 below). First, experts' decision on the menu of treatments in the beginning of a period is taken as a commitment for the duration of that period, and it is assumed observable by customers. One may think, for example, that there are two alternative technologies with costs (L,H) and (L,H') respectively, where H' is sufficiently larger than H, say H'>H+k. Customers are informed about the technology choice and the commitment to providing only the L service is identified with the choice of the technology (L,H'). Second, the assumption that prices are chosen in the beginning of each period means, in particular, that they cannot be renegotiated after the customer has already come to visit. Third, note that there are two possible scenarios regarding the extent to which experts are committed to provide the service. In one scenario the price announcement means very firm commitment to provide the service--if, for example, an expert announced $P_H<H$ he still must treat for that price a customer who has the H-problem. In an alternative scenario, after the diagnosis but before they agreed, the expert is not yet committed to treat this customer and can reject him. In the latter case the expert would rather reject a customer than treat him at a price below cost. We find the second interpretation more appealing and hence will adopt it, but we shall argue in section 4 that similar results hold for the alternative "firm commitment" interpretation. It follows from this assumption that from now on we can restrict the experts' strategies to always satisfy $P_L \geq L$ and $P_H \geq H$. This restriction will have the same effect as if we left the prices unrestricted and added another action--rejection--to those that the expert can take after diagnosing a customer.

3. The equilibria: specialization as a result of asymmetric information

The market equilibrium is shaped mainly by the following two factors. First, the asymmetry of information gives experts incentives to cheat by misrepresenting a minor treatment as a major one and, in turn, may induce customers to search. Second, the search-cum-diagnosis cost, k, induces customers to economize on the number of visits. The following analysis presents the ME which balances these factors through specialization. If the cost k is not too high, at this ME some experts provide only the L-treatment and others effectively provide only the H-treatment. By providing only the L-treatment, experts create a credible commitment not to cheat, since they simply lose the customers to whom they do not recommend the L-treatment. Customers start their search with such an expert and reach other experts only if they have the major problem. The Bertrand style competition among experts of each type drives their prices to L and H respectively.

Proposition 1:

(i) If $k \leq (1-w)(H-L)/w$, there exists a specialization ME in which some experts (at least two) employ the strategy $\{(L,\infty),0\}$, while others (at least two) employ the strategy $\{(H,H),x\}$. Customers sample first an expert who offers (L,∞) and, if recommended the costly treatment go to an expert who offers (H,H).

(ii) If $k > (1-w)(H-L)/w$, there exists a no-specialization ME in which all experts offer the schedule (H,H). Customers are served by the first expert they sample.

Proof: (i) Consider the proposed ME configuration. Customers' beliefs are that the (L,∞) experts recommend truthfully. First, observe that, given their beliefs, the customers' behavior is indeed optimal. The two relevant alternatives are the described behavior and going directly to a (H,H) expert. The expected cost of the former is $k+(1-w)L+w(k+H)$, while the expected cost of the latter is $k+H$. But since $k<(1-w)(H-L)/w$, the former cost is lower, establishing the optimality of customers' strategies. Second, observe that an expert of either price group cannot gain from deviating only to a different recommendation policy. Third, we have to complete the description of the equilibrium beliefs and strategies such that no arbitrary expert's deviation, in both prices and recommendation, yields the deviant expert positive profit. Consider an arbitrary deviation $\{(P_L,P_H), x\}$. If $P_H \geq H+k$ this offer is clearly not more attractive than (L,∞) and if $P_L \geq H$ it is not more attractive than (H,H), regardless of the customers' beliefs about x. In the remaining cases, $P_L \leq H$ and $H \leq P_H < H+k$, the customers' belief is that x=1, so this offer is again not more attractive than (H,H). Thus, in all these cases, not going to this expert is best response for all customers. Therefore, the deviation is not profitable.

(ii) Consider the proposed equilibrium configuration. By repeating the same arguments of part (i), we conclude that the only deviant offers, which might conceivably upset this ME are of the form (P_L,∞), where $L \leq P_L < H$. The expected cost incurred by a customer who visits the deviant expert is, at least, $k+(1-w)P_L+w(k+H)$. But the assumption $k>(1-w)(H-L)/w$ and $P_L \geq L$ imply that $k+(1-w)P_L+w(k+H)>k+H$, which is the customer's cost of following the ME. Hence, no profitable deviation is possible. QED

Thus, three magnitudes determine whether or not this ME involves specialization: the search-cum-diagnosis cost k, the cost differential H-L, and the prior w. Specialization is more likely the smaller is k, the larger is H-L and the smaller is w. This is because a large H-L and a small w increase the expected gains from customers' search, while a small k reduces the cost.

The ME of proposition 1 is obviously not unique. For example, the schedule (H,H) posted by the H treatment experts is not pinned down--the same behavior will constitute an equilibrium, if all or some of the H treatment experts employ the strategy $\{(P_L,H),1\}$ with $P_L \le H$. Also, the number of experts who make each of the offers is left undetermined. These variations, however, preserve the fundamental features of the ME of proposition 1. That is, in the range of (i) experts specialize and customers with the L and H problems are treated at the prices L and H respectively; in the range of (ii) each expert provides both services and all customers pay H.

More importantly, however, since the beliefs of customers off the equilibrium path are unrestricted, there are many more ME which significantly differ from the one described above. The following proposition argues that, if customers' beliefs satisfy a certain mild restriction and if there are "many" experts, then all ME have the essential features of the ME outlined in proposition 1.

Restriction on beliefs: if an expert offers (Q,∞) where $Q>L$, then customers believe that $x=0$.

Note that, for such an expert, if $Q>L$ and $x>0$, then $\{(Q,\infty),x\}$ is dominated by $\{(Q,\infty),0\}$, since he makes positive profit only on customers to whom he recommends the L treatment. Thus, this restriction on beliefs is implied by any refinement which assigns probability zero beliefs to dominated actions.

The simplest way to consider here the effect of many experts is to think of the expert population as a continuum. We shall therefore present proposition 2 in these terms. But in the appendix we bring its somewhat messier counterpart, proposition 2^*, for the case of finite number of experts. Now the experts' strategies are a measurable function from $[0,1]$ which takes as values triples of the form $\{(P_L,P_H),x\}$. In a ME the experts' profits are the same for all offers in the support. We shall think of a deviation to a price schedule outside the support as the introduction of a small mass of experts with the new schedule. We shall say that a deviation is profitable, if the limit of the per expert profit in the deviating mass, as this mass approaches zero, is greater than the common per expert profit in the support.

Proposition 2: With continuum of experts and the above restriction on beliefs, all ME are of the form described in proposition 1 (up to irrelevant differences).

The proof is brought in the appendix. Its main point is that, for an expert to provide both services at different prices, the price of the L service has to incorporate sufficient mark-up over cost and at least some of the customers must be planning to reject an H recommendation. But these conditions together with assumption of many experts (continuum) assure that there is room for profitable appearance by an expert who specializes in L. In the appendix we also bring the statement and proof of Proposition 2^*, the finite N version of proposition 2. It argues that, if N is sufficiently large, then any ME which satisfies the restriction on beliefs is very close to the ME described in proposition 1.

4. Discussion

Specialization

The main point of Propositions 1 and 2 is that the information asymmetry in a market of this type may create or reinforce an inducement for vertical specialization. While this is obviously an abstraction and it is probably impossible to point out an industry that features such perfect specialization, it identifies an element which may be present in the conduct of such industries. The car repair industry, for example, features some degree of vertical specialization of this type: there are shops that specialize in the lower end of repairs and such shops seem more often to compete by providing estimates in advance.

It may seem that the information problem in such markets may be alleviated by separation of the diagnosis and treatment. That is, either due to market forces or due to regulation, the function of diagnosis may be performed by a separate set of experts who do not provide the treatment itself. Indeed, FTC(1979) considered in some detail this idea and the policy questions of how to encourage it. However, as that report notes, such arrangements may not be sustainable since they would presumably raise new problems regarding the proper incentives for the diagnostician. The present paper shows that, in an environment in which such separation is impossible, the equilibrium vertical specialization essentially mimics its effect.

Configurations with fraud

To sharpen the understanding of the ME described above, let us compare it with another candidate for an equilibrium. Consider the following symmetric configuration. All experts announce the schedules $(P_L, P_H)=(L+e, H)$ and use recommendation policy $x \in (0,1)$; customers always accept an L recommendation; in their first visit they accept an H recommendation with probability $y \in (0,1)$, but accept it with certainty in later visits. The probabilities x and y and the mark-up e satisfy

$$e = \frac{(H-L)[y+x(1-y)]}{1+x(1-y)} \tag{1}$$

$$H = k + \frac{(1-w)x(1-x)(L+e)+[w+(1-w)x^2]H}{w+(1-w)x.} \tag{2}$$

Equation (1) assures that an expert who faces a customer is just indifferent between reporting truthfully and earning e or cheating and earning H-L with probability $[y+x(1-y)]/[1+x(1-y)]$. This probability takes into account that a fraction $1/[1+x(1-y)]$ of the L customers are visiting for the first time and hence will accept an H-recommendation with probability y, while the remaining fraction $x(1-y)/[1+x(1-y)]$ have visited already another expert before and will accept such recommendation with certainty. Equation (2) assures that a customer who has received an H-recommendation in his first visit is just indifferent between accepting it or not. The RHS captures the expected value of continuing the search: the conditional probability that this customer has the L-problem is $(1-w)x/[w+(1-w)x]$ which multiplied by $(1-x)$ gives the probability that the next expert will offer him the L-treatment; the complementary probability, that the next expert will offer him the H-treatment, is then $[w+(1-w)x^2]/[w+(1-w)x]$. Now, if the prices are exogenously fixed this configuration is an equilibrium, since (1) and (2) assure that both the expert's strategy, which reduces to the choice of x, and the customer's strategy are best responses. This is essentially the type of equilibrium considered by Pitchik and Schotter. It exactly balances the incentives to cheat and the incentives to search: there is sufficient amount of fraud to induce some customers to seek second opinions, and customers seek second opinions sufficiently vigorously to prevent experts from cheating all the time. When prices are flexible, this candidate for an equilibrium is destroyed, since the mark-up e which is necessary to support this configuration as an equilibrium invites price undercutting.

Diagnosis costs

Recall that the cost k represents both the customer's search cost and the cost of diagnosis and is assumed to be born by customers. A more complete model would distinguish between a direct customer search cost, k_1, a direct diagnosis cost, k_2, where $k_1+k_2=k$, and a diagnosis fee, d, which experts determine as part of the competition. Our assumptions concerning k amount in fact to assuming that the fee d exactly covers the cost k_2. By going over the proof of proposition 1, it is quite straightforward to verify that the specialization configuration with $d=k_2$ will still be an equilibrium in a wider model in which the diagnosis fee is determined endogenously. The addition of a decision variable makes the uniqueness more difficult to establish, and although the same factors seem to be at work in the wider model, that point might require more work.

Customers' information about prices

Recall that the interaction is modelled as if customers learn about experts' prices before they go to be diagnosed, and that these prices are not renegotiated later. Contrast this with an alternative model in which customers learn about prices only when they actually visit an expert. It follows from a result of Diamond(1971) that, in search models of the latter type, the unique equilibrium outcome is the no-trade outcome in which all experts charge prohibitive prices and no customer bothers to go to an expert. This result follows from the observation that, due to the search cost, at any configuration with trade it would pay at least some experts to raise their prices by less than k. Now this devastating effect of the search cost, which is not unique to this set up, would destroy price competition in any model, if we allow prices to be "renegotiated" after the buyer incurred some cost. Thus, the reason we formulated the model as we did was to avoid this modeling problem which is tangential to the issues we are considering. One may think that the search cost associated with finding out prices is negligible, and the search costs that matter are the diagnosis costs.

Consumers' information about service menus

It was also assumed that experts can commit to specialization in the L-treatment, at least for the duration of a period, and that customers actually observe experts' menu decisions before they go on their search. As we mentioned before, this can be thought of as actual choice of technology. Now, while in the absence of this assumption the notion of experts specializing in the L-treatment is less meaningful, it is quite straightforward to verify that the specialization configuration could still be supported as an equilibrium . This is due to the large freedom in the choice of equilibrium beliefs, although in that environment the equilibrium will most probably not be unique.

Commitment to serve at prices below cost

Recall that the analysis has been done under the assumption that experts cannot commit to provide service at a price below cost. In particular, an expert cannot credibly commit to serve all customers for $P_L=P_H=(1-w)L+wH$, regardless of their problem, since then he would presumably turn away the customers he would diagnose as H. If we assumed instead that experts could credibly make such a commitment, the analysis of the previous section remains similar, but some of the magnitudes change. The differences are that now the no-specialization price schedule would be $P_L=P_H=(1-w)L+wH$ and the range of k for which the specialization ME prevails will be smaller, $k<(1-w)(H-L)$. To see this, observe that the expected cost of going directly to a non specialized expert who charges $P_L=P_H=(1-w)L+wH$ is $k+(1-w)L+wH$, whereas the expected cost of sampling first an (L,∞) expert and only then continuing, if necessary, to the $P_L=P_H=(1-w)L+wH$ expert is $k+(1-w)L+w[k+(1-w)L+wH]$. For $k<(1-w)(H-L)$ the former cost exceeds the latter. This means that the deviation to $P_L=P_H=(1-w)L+wH$ will not be profitable in the specialization ME since it will attract only H customers. It also means that the non-specialization configuration in which all experts charge $P_L=P_H=(1-w)L+wH$ is not an ME since it can be upset by the offer $(L+\varepsilon,\infty)$, where ε is sufficiently small.

Welfare

Since in this model the demands are inelastic and all customers are served, the only magnitude which affects the total surplus (the sum of customers' and experts' surplus) and which varies in the model is the aggregate search-cum-diagnosis costs. This cost is minimized when each customer visits only one expert, in which case it is exactly k per customer. In the specialization ME this cost is $(1+w)k$ per customer, since customers who truly need the major treatment end up visiting two experts and incurring these costs twice. Thus, the welfare cost associated with the information asymmetry is wk per customer.

Note, however, that although specialization has here an adverse effect on total surplus, customers' surplus taken alone is higher under the specialization ME than it would be under a non-specialization configuration in which the prices are (H,H). Therefore, this analysis does not provide a serious rationale for limiting the form of the competition.

Equilibrium rents

In the non-specialization ME the price H is charged for both treatments. Although this price exceeds the expected cost $(1-w)L+wH$, it cannot be undercut since, by assumption, the experts cannot commit to providing the treatment below cost. That is, an expert who announces the schedule (P,P), where $(1-w)L+wH<P<H$, will make profit by treating both problems, but he will make even higher profit by rejecting the customers who need major treatments. Thus, in the non-specialization ME there is rent accruing to experts. The literature on the provision of product quality discusses how advertising and other expenditures might dissipate such rents (see, e.g., Klein and Leffler(1981)). Since this is besides the point of this paper, we shall not discuss here the various possibilities of rent dissipation. However, the fact that such rent accrues to experts only in the non-specialization case may suggest that we can expect to observe more rent dissipation activities, such as advertising, when the conditions preclude specialization.

5. Diagnosis Errors

So far we have assumed that experts diagnose the problem perfectly. The possibility of experts' errors, introduced in this section, may affect behavior in this market, via the incentives it creates for customers to try to take advantage of experts' errors.

Let us modify the model in a minimal way that allows to capture this aspect. Suppose that the diagnosis yields one of two signals "H" or "L". Let $\text{Prob}\{\text{"H"}|H\}=h=1-\text{Prob}\{\text{"L"}|H\}$. Assume that $h>1/2$ and that $\text{Prob}\{\text{"L"}|L\}=1$. Thus, the signal "H" has the interpretation that the problem is more likely to be H, while "L" has the interpretation that the problem is more likely to be L. The assumption that $\text{Prob}\{\text{"L"}|L\}=1$ implies that an "H" signal reveals with certainty that the problem is major, but an "L" signal can result in either case. It follows that when an expert has to recommend a treatment, he is not always certain about the cost. The expert will discover the true cost only later, in the process of fixing the problem. However, the nature of the expert's commitment remains as before: if a customer accepts an expert's recommendation, this expert is obligated to fix the problem regardless of whether he estimated the cost correctly[5].

As before, a strategy for an expert is a triple $\{(P_L,P_H),x\}$. But now the recommendation policy $x\in[0,1]$ is the probability that this expert will recommend the H-treatment when the diagnosis yields the signal "L". A customer's strategy and the notion of market equilibrium (ME) are just as before.

The analysis in this section is similar to the previous analysis. The main difference is that here the cost an expert expects to incur in treating a customer does not depend only on the signal the diagnosis of this customer generates. It also depends on the prior probability the expert has over this customer's problems. This prior probability may in turn depend on the endogenously determined steady state population distribution. Let $E(\pi)$ denote the expected cost of treating a customer conditional on the diagnosis "L" and provided that he is drawn randomly from the population in which L and H occur with the probabilities $(1-\pi)$ and π respectively. That is,

$$(3) \qquad E(\pi) = \frac{(1-\pi)L+\pi(1-h)H}{(1-\pi)+\pi(1-h)}$$

For $g\in[0,1]$, let

$$(4) \qquad \pi(g) = \frac{w(1-hg)}{(1-w)(1-h)+w(1-hg)}$$

and let

$$(5) \qquad E_g = E(\pi(g)) = \frac{(1-w)L+w(1-hg)H}{(1-w)+w(1-hg)}$$

Note that

$$(6) \qquad E_1 = \frac{(1-w)L+w(1-h)H}{(1-w)+w(1-h)} \qquad \text{and} \qquad E_0 = (1-w)L+wH.$$

The following proposition characterizes the counterpart of the ME described in proposition 1.

<u>Proposition 3</u>: There exists an ME with the following features.

(i) If $k>(1-w)(H-L)/wh$, then all experts offer (H,H); all customers are treated by the first expert they visit.

(ii) If $(1-w)(H-L)/wh>k>(1-h)(1-w)(H-L)/h[1-w+w(1-h)]$, then some experts (at least two) employ the strategy $\{(E_1,\infty),0\}$, while others (at least two) employ the strategy $\{(H,H),x\}$; customers sample first an expert who offers (E_1,∞) and, if recommended the costly treatment, go to an expert who offers (H,H);

(iii) If $(1-h)(1-w)(H-L)/h[1-w+w(1-h)]>k>(1-h)(1-w)(H-L)/h$, then there is some $g\in(0,1)$ such that at least two experts employ the strategy $\{(E_g,\infty),0\}$, while the rest (at least two) employ the strategy $\{(H,H),x\}$; some customers (a fraction g) go only once to an expert who offers (E_g,∞) offer and then continue to an (H,H) expert, while others (fraction 1-g) keep visiting the (E_g,∞) experts until one recommends the L treatment.

(iv) If $k<(1-h)(1-w)(H-L)/h$, then all experts employ the strategy $\{[(1-w)L+wH,\infty],0\}$; customers keep visiting experts until one recommends the L treatment.

The proof is relegated to the appendix. Parts (i) and (ii) are the direct counterparts of the two parts of proposition 1. If k is high, customers do not search at all. If k is lower, there is a specialization ME: customers try once an expert who specializes in the L-treatment; if they are diagnosed as H, they go to an expert who specializes in the H treatment. Note that the uncertainty of the experts is reflected in the price, E_1. Observe from (6) that E_1 is equal to the expected cost of treating a customer, conditional on diagnosis "L" and the information that this is this customer's first visit. In each period of the steady state that corresponds to part (i) the only customers in the market are the M that entered in the beginning of that period. In each period of the steady state of (ii) there are M(1+wh) customers--the new M plus the Mwh who generated the "H" signal in the previous period.

Parts (iii) and (iv) have no parallel in proposition 1, since they stem from the experts' imperfect diagnostic ability. In part (iv) k is low enough so that customers try to take advantage of experts' imperfect diagnostic ability by sampling experts repeatedly until some expert is fooled. The steady state customer population has M+Mwh/(1-h) customers per period. Those with the L problem leave the market after one period, while the H customers stay on the average for 1/(1-h) periods. Consequently, the relative share of the H customers in the population is not w anymore but

$$\frac{Mw + Mwh/(1-h)}{M + Mwh/(1-h)} = \frac{w}{w + (1-w)(1-h)} > w.$$

The price (1-w)L+wH announced by all experts is equal to the expected cost conditional on the diagnosis "L", calculated using the above base probabilities. It is therefore higher than the E_1 price of part (ii), since E_1 is the expected cost conditional on "L" relative to base probability w of the H customer. This is some sort of experts' winners' curse: the expected cost conditional on "L" reflects the fact that H customers keep searching until they are diagnosed as "L".

The case of part (iii) is intermediate between (ii) and (iv). Customers who find out their problem is H are exactly indifferent between trying to take advantage of experts' errors as in (iv) or going to an (H,H) expert as in (ii). In the ME, fractions 1-g and g adopt the former and latter behaviors respectively. Consequently, the steady state customer population has $M+Mw(1-g)h/(1-h)$ customers per period, the probability that a randomly drawn customer has the H problem is $\pi(g)$ given by (4) and the expected cost conditional on the signal "L" is E_g given by (5). Note from (5) and (6) that $E_1<E_g<E_0$, so that the lower price in this case is intermediate between the prices in cases (ii) and (iv).

Over the range of k described in parts (i) and (ii) the behavior of prices, total surplus and customers' surplus are similar to the behavior of these magnitudes in the case with no diagnostic errors. The special effect of these expert errors is over the range described in (iii). There, the effect of k on the prices and the welfare measures seems perverse: lower levels of k are associated with higher prices, lower total surplus and lower customers' surplus. The cause is the negative search externality exerted by searching customers who need the H treatment. To see this, fix the prices for a moment. With lower k, customers who require the H service will have an incentive to continue searching until meeting an expert who errs and offers the lower price. This implies that on the average the low price experts will treat more H customers and make a loss, i.e., they experience a winners' curse. Therefore the equilibrium low price will have to be higher. This search externality then induces more search as well as higher prices.

6. Reputation

In the search scenario considered so far experts commit to the cost of the treatment in adva and customers are therefore able to solicit and compare a number of bids before they decide. While this arrangement may conform with the practice in certain repair services, it is clearly not the only form of organization in such markets. Under an alternative arrangement, experts do not guarantee the cost in advance: customers entrust their problem to an expert who first fixes it and then presents them with a bill. What may keep experts in this alternative system from always overcharging is their need to maintain reputation. That is, if customers require this service repeatedly and tend to withdraw their business from experts who seem to recommend the H treatment too frequently, then experts may be deterred from cheating.

Let us modify the above model in a minimal way that will facilitate discussion of the alternative market organization as well. As before, there are N experts in this market. But the demand side now has a simple overlapping generations structure. In each period a new batch of M customers enters the market. Each customer lives for two periods and experiences the problem exactly twice--once in each period. The customer's two problems are independent, i.e., each problem is L or H with probabilities (1-w) and w respectively. In the beginning of each period experts announce their price schedules. Then each customer chooses an expert. Finally, after treating a customer, an expert can claim the treatment was H or L and accordingly charge the price he had announced for that treatment.

Before we proceed, two remarks are in order. First, note that the interpretation of the time period in this model is different than before--it corresponds to a more significant interval of time. Still we continue to assume for now that a customer can sample only one expert per period, but we shall relax this assumption shortly. Second, recall that in the previous sections we restricted attention to prices (P_L, P_H) such that $P_L \geq L$ and $P_H \geq H$, and justified this by the assumption that experts can reject customers whose treatment would be unprofitable. Here however $P_H < H$ may still be relevant, since reputation effects may induce the expert not to reject a customer even if treating him may involve a loss. We shall therefore not impose the restriction $P_H \geq H$ and include explicitly the rejection action for experts. It will be evident that, since the expert can always misreport the L problem as an H problem, there is no such problem with P_L so that we can continue to assume $P_L \geq L$.

An _action_ for an expert in a given period is a four-tuple $\{r, x, (P_L, P_H)\}$, where $P_L \geq L$, $x \in [0,1]$ is the charging policy -- the probability with which customers with the L problem will be charged for the H problem, and $r \in [0,1]$ is the rejection policy -- the probability with which an expert will turn away customers with the H problem. A _truthful_ action is such that $x=r=0$.

In his each period the _information_ of a customer consists of the prices chosen by experts in that period; in the second period it also includes the results of his first period visit. A customer's _strategy_ prescribes choice of an expert in each period as a function of the information. A customer has _beliefs_ over experts' recommendation and rejection policies, as a function of his information.

The payoff to a customer is the sum of his two periods payoffs. If the customer's first and second period problems were fixed at the prices p and q respectively, his payoff is $(B-p)+(B-q)$. The requirement that the reservation value B is sufficiently large will be given here the specific meaning that (1-

w)(B-L)<B-H, i.e., a customer prefers an expert who fixes both problems at the price H to one who fixes only the L problem at the price L. At period t, the payoff to an expert is the average per period profit from that period on, i.e., $\lim\sup(\sum_{i=t}^{T} a_i)/(T-t)$.

A symmetric reputation equilibrium (SRE) consists of an action $\{r,x,(P_L,P_H)\}$ for all experts and all periods, a strategy for all customers and customers' beliefs such that:

(i) Given customers' strategy and that all experts use $\{r,x,(P_L,P_H)\}$ in perpetuity, no expert can make a sequence of deviations that will increase his payoff.

(ii) Customers' strategy is best response, given their beliefs.

(iii) Customers' beliefs are correct on the equilibrium path.

Proposition 4: (i) The SRE in which customers' expected costs are minimal is such that $P_L=P_H=H-(1-w)(H-L)/(1+2w)$ and r=0. Customers' strategy is to return to the same expert with the second problem if he treated them in the first visit, but to go to another expert if the first expert rejected them.

(ii) For w≤2/3 the above SRE is also the one in which P_L is minimal. For w>2/3 the SRE in which P_L is minimal is such that $P_L=L+2(H-L)/(3-w)$, $P_H=H$ and x=0. Customers' strategy is to return to the same expert with the second problem if and only if he treated them in the first visit.

The proof is relegated to the appendix. In both of the above equilibria the inducement for experts to be truthful is the prospect of repeat business. In the SRE of (i) this prospect offsets the temptation to reject a customer who has the H problem rather than treat him at a price below cost; in the SRE of (ii) it offsets the temptation to obtain the immediate gain of P_H-P_L=(1-

w)(H-L)/(3-w) by charging the customer for H when his problem is indeed L. The mark-up, embodied in the price P_L in each of these cases mitigates the incentive to cheat by increasing the profitability of the repeat business; in case (ii) it also reduces the immediate gain from overcharging.

Our next step is to combine this reputation model with the previous analysis to discuss what determines whether the market is more likely to be organized along the lines of the search or the reputation settings. To do this we have to allow customers to visit more than one expert within each period of the reputation model. Assume, therefore, that each period of the reputation model is divided into a number of search periods which correspond to the periods in the previous sections. There is no technical or conceptual problem in assuming this, especially since, as we have noted, it is only natural to think of the period of the reputation setting as substantially longer than the period of the search setting. In this setting, consider any one of the SRE described in proposition 4. Consider also a deviation by an expert who announces the schedule (Q,∞), $L<Q<P_L$, in the same meaning it had in the search setting considered before. That is, this expert specializes in the L-service and announces an alternative arrangement whereby he will diagnose customers in advance and in case he recommends the L-treatment the price will be Q. Now, since the equilibrium P_L embodies a mark-up, such a deviation with $L<Q<P_L$ could be profitable, provided the number of experts N is sufficiently large. To see this, suppose first that, as in sections 2-3, an expert can diagnose a problem with certainty. The expected cost incurred by a customer who goes first to this deviant expert is $k+(1-w)Q+w(k+P_H)$, while the expected cost for a customer who goes directly to one of the others is $k+(1-w)P_L+wP_H$. Now, the latter cost exceeds the former if and only if $k<(1-w)(P_L-Q)/w$. For example, in the SRE of part(i) of the proposition where $P_L=H-(1-w)(H-L)/(1+2w)$, there exists $Q>L$ such that (Q,∞) upsets the reputation equilibrium, if and only if, $k<3(1-w)(H-L)/(1+2w)$. And, if N is large enough such deviation would be profitable. Recall that for $w\leq 2/3$ the considered SRE has the lowest P_L over all SRE's. (For other values of w we can use of course the SRE of part (ii) to find a similar condition on k.) Thus, the implication is that, if the search-cum- diagnosis cost is sufficiently low, other things equal, the reputation arrangement cannot be sustained as an equilibrium and will presumably be replaced by the search arrangement. Notice that, in the setting that permits search which we are discussing now, the specialization equilibrium of proposition 1 is indeed an equilibrium.

If experts are liable to make diagnosis errors, the analysis is similar. As above, consider the reputation setting with one of the equilibria of proposition 4, and a deviant expert who announces the schedule (Q,∞), i.e., offers to commit in advance to the cost Q if he gets the signal "L". For such deviation to be profitable, it is necessary that Q covers the expected cost conditional on the signal "L", i.e., $Q>E_1=[(1-w)L+w(1-h)H]/[(1-w)+w(1-h)]$. Now, the expected cost incurred by a customer who goes first to this expert is $k+[(1-w)+w(1-h)]Q+wh(k+P_H)$, while the expected cost for a customer who goes directly to one of the others is $k+(1-w)P_L+wP_H$. Thus, for the SRE of part (i) of the proposition, there exists a $Q>E_1$ such that the latter exceeds the former if and only if $k<(1-w)(H-L)(2+h)/h(1+2w)$. The conclusion is that the reputation arrangement of charging after the treatment is more likely to be sustained the larger is k and the larger is h, i.e., the lower is the probability of diagnosis error, 1-h.

The reputation setting obviously minimizes the total search-cum-diagnosis costs, since customers go with each problem only to one expert. This is in contrast to the search setting which at equilibrium necessarily involves higher costs. On the other hand the reputation setting gives rise to higher prices, and is hence more profitable to experts. Now, the higher profitability together with the preceding observations imply that there is some scope for collusion on the conduct. If a trade association succeeds in ruling out the type of behavior that upsets the reputation equilibrium, experts will profit at the expense of customers, even when price competition is totally free.

In the car repair market both forms of arrangement--the search and the reputation--coexist. Some repair shops would agree to diagnose and provide binding estimates prior to the treatment, at least for a certain range of problems, while other shops would insist on the other method. The distinction is of course not clear and many shops presumably work in both methods. This diversity is not necessarily inconsistent with the above analysis, since in practice customers are heterogenous with respect to their search costs and different problems may differ with the respect to the cost and accuracy of a diagnosis. The above analysis suggests that perhaps the search system serves problems that can be more easily and more accurately diagnosed and is patronized by customers with lower search costs, while the reputation system serves problems which are harder to diagnose and is patronized by those with the higher costs.

The simple models that we consider expose the effect of only two elements, k and h. However, the manner the reputation effect works and its effect on prices depend of course on many other factors that were ignored here. The frequency of service, the speed with which bad reputation spreads, the expertise of customers and the nature of the legal system may all have an effect and hence a role in the determination of the form of organization.

7. Concluding Remarks

The important subset of markets considered in this paper are not discussed extensively in the theoretical literature although their special features sufficiently distinguish them and merit some special attention. Markets in this category are also sometimes subject to some forms of licensing and regulation, so a closer understanding is immediately useful for public policy.

The main insights of this paper are into (i) the possibility that the asymmetry of information special to these markets may induce vertical specialization; (ii) the negative search externality present in such markets and its effect on prices; and (iii) the role of basic parameters, the search-cum-diagnosis costs and the accuracy of diagnoses, in the determination of the market's form of organization--search or reputation.

Appendix

Proposition 2: With large numbers and the above restriction on beliefs, all ME are of the form described in proposition 1 (up to irrelevant differences).

Proof: Consider a particular ME. Let $\{(P_L, P_H), x\}$ denote the ME strategy of an expert, who is visited by some customers on their first visit. Let

$$y = \text{prob}\{\text{a customer who visits this expert agrees to the H treatment} \mid \text{this customer's problem is minor}\}$$

Step 1: If $x<1$ and $P_L<P_H$, then $P_L=L$, $P_H \geq H+k$, $x=0$ and $y=0$.

Proof: First note that, if $y=1$ then $x=1$, contrary to $x<1$. Therefore, $y<1$. That is, there are some customers who are indifferent or strictly prefer to leave after an H recommendation. Notice that the customers for whom this is the first visit must be among them, since the probability of having the L problem is higher for them than for others. If $P_L>L$, these customers would strictly prefer to sample first an expert with price schedule (Q,∞), where $L<Q<P_L$, since by the restriction on the beliefs they would believe that he recommends truthfully. Therefore, (Q,∞) will attract all first period customers, so it is a profitable deviation. Similarly, if $x>0$ there exists some $Q>L$ such that all first visit customers will strictly prefer to sample first an expert with (Q,∞), which makes this a profitable deviation. Hence, $x=0$.

The optimality of $x<1$ implies that $P_L^i - L \geq (P_H^i - L)y$, i.e., the expected profit of recommending L to a customer with the L problem (the LHS) is at least as high as the expected profit of recommending H (the RHS). Thus, $y=0$

Since $x=0$, a customer's posterior probability of having the H problem, given that this expert recommended H is 1. Hence, the expected cost this customer will bear by going elsewhere is at least $H+k$. Now, the optimality of

$y=0$ implies that P_H is not smaller than this cost, for otherwise the customer would prefer not to leave (i.e., y would be equal 1). Hence, $P_H \geq H+k$. ∎

<u>Step 2</u>: If $x=1$ or $P_L=P_H$, then $P_H=H$.

<u>Proof</u>: In either case this expert's customers will pay P_H^i. If $P_H>H$, all this expert's customers and hence all customers on their first visit will be attracted by an offer (Q,Q) such that $P_H>Q>H$, which makes it a profitable deviation. Therefore, $P_H=H$. ∎

<u>Step 3</u>: If $k<(1-w)(H-L)/w$, all experts visited by first period customers are of the type described in step 1.

If $k>(1-w)(H-L)/w$ all experts visited by first period customers are of the type described in step 2.

<u>Proof</u>: Suppose $k<(1-w)(H-L)/w$. If all experts are of the type described in step 2, then there is $\varepsilon>0$ such that the deviation $(L+\varepsilon,\infty)$ will attract all first visit customers, since the expected cost of first going there is lower than the expected cost of going directly to an expert who charges H. That is, for k in this range and sufficiently small ε,

$$k+(1-w)(L+\varepsilon)+w(H+k)<k+H.$$

Similarly, it may not be that some experts who get first visit customers are as in step 1 and some are as in step 2, since for these values of k all first visit customers will prefer to go to the experts described in step 1. This again follows from the last inequality.

Suppose $k>(1-w)(H-L)/w$. It may not be that all or some experts who get first visit customers are as in step 1, since for k in this range the expected cost of going first to an (H,H) expert is lower than going first to an (L,∞) expert. I.e., $k+H<k+(1-w)L+w(H+k)$. Thus, if there exists an (H,H) expert, customers will not go to an (L,∞) expert; if there is no (H,H) expert, then

(H,H) will be a profitable deviation as it would attract all first period customers. ∎

Step 4: If $k<(1-w)(H-L)/w$, there are experts who serve only second visit customers and charge $P_H=H$.

Proof: It follows from steps 3 and 1 that, in this case, all customers with the H problem, go on to a second expert after the first expert recommended to them the H treatment. Since $x=0$, these customers know they have the H problem. Hence, they will prefer an expert with $P_H<H+k$ to an expert with $P_H \geq H+k$. By step 3 all experts who serve first visit customers have $P_H \geq H+k$. Therefore, it must be that all second visit customers go to other experts with $P_H<H+k$, for otherwise a deviation with say $P_H<H+k$ can profitably attract all of them. Now, it may not be that an expert who serves the second visit customers has $P_H>H$, since then a deviation (Q,Q) such that $H<Q<P_H$ would attract all second visit customers and hence would be profitable. ∎

Observe that steps 1-4 establish the proposition. For $k<(1-w)(H-L)/w$, steps 1,3 and 4 imply that all ME have the essential features of the ME described in part (i) of proposition 1; for $k>(1-w)(H-L)/w$, steps 2 and 3 imply that all ME have the essential features of the ME described in part (ii) of proposition 1. QED

Proposition 2*: For any $\varepsilon>0$, there exists an $N(\varepsilon)$ such that if $N>N(\varepsilon)$ then any ME that satisfies the assumption on beliefs is ε close to the ME described in proposition 1 in the following sense.

(i) For k> the ME is exactly as in proposition 1

(ii) For $k<(1-w)(H-L)/w - \varepsilon$, any such ME has the following features. All customers go first to experts with $P_L \leq L+\varepsilon$. Those with the L problem are recommended the L treatment with probability of more than $1-\varepsilon$. Those who get the H recommendation continue with probability of more than $1-\epsilon$ to another expert who treats only second visit customers and charges $P_H \le H+\epsilon$ to all.

Proof of proposition 2*:

Suppose that the number of experts N is large enough so that, in any ME, there is some expert whose profit is less than δ (say, $N>BM/\delta$). Define the following six magnitudes: (i) $a(\delta)=\delta/(1-w)M$; (ii) $c(\delta)=a(\delta)/k$;
(iii) $b(\delta)=(H-L)(1-w)c(\delta)/[(1-w)c(\delta)+w]$; (iv) $d(\delta)=a(\delta)/(H-L)$;
(v) $e(\delta)=[c(\delta)(1-w)+w(1-w)(1-c(\delta))a(\delta)](H-L)/w[w+(1-w)c(\delta)]$;
(vi) $f(\delta)=\delta/w(1-d(\delta))M$.

We shall explain these magnitudes later as we shall use them. The important fact to notice now is that all six magnitudes approach 0 when δ approaches 0.

Consider a particular ME. Let $\{(P_L^i,P_H^i),x^i\}$ denote the ME strategy of expert i, who is visited by some customers on their first visit. Let
y^i = prob{a customer who visits expert i agrees to the H treatment|
this customer's problem is minor}

Step 1: If $x^i<1$ and $P_L^i<P_H^i$, then $P_L^i \le L+a(\delta)$, $P_H^i \ge H+k-b(\delta)$, $x^i \le c(\delta)$ and $y^i \le d(\delta)$.

Proof: First note that, if $y^i=1$ then $x^i=1$, contrary to $x^i<1$. Therefore, $y_i<1$. That is, there are some customers who are indifferent or strictly prefer to leave after an H recommendation. Notice that the customers for whom this is the first visit must be among them, since the probability of having the L problem is higher for them than for others. If $P_L^i>L+a(\delta)$, these customers would strictly prefer to sample first an expert with price schedule $(L+a(\delta),\infty)$, since by the restriction on beliefs they would believe that he recommends truthfully. Therefore, this offer will attract all first period customers. By the choice of $a(\delta)$, the expected profit of an expert who makes this offer is at least δ, so there must be an expert for whom $\{(L+a(\delta),\infty),0\}$

is a profitable deviation. Similarly, by the choice of $c(\delta)$, if $x^i > c(\delta)$, all first visit customers will strictly prefer to sample first an expert whose price schedule is $(L+a(\delta),\infty)$.

The optimality of $x^i<1$ implies that $P^i_L - L \geq (P^i_H - L)y^i$, i.e., the expected profit of recommending L to a customer with the L problem (the LHS) is at least as high as the expected profit of recommending H (the RHS). Thus, $y^i \leq a(\delta)/(H-L)=d(\delta)$.

A customer's posterior probability of having the H problem, given that expert i recommended H is $w/[(1-w)x^i+w]$. Hence, the expected cost this customer will bear by going elsewhere is at least

$k+L(1-w)x^i/[(1-w)x^i+w]+Hw/[(1-w)x^i+w] \geq H+k-(H-L)(1-w)c(\delta)/[(1-w)c(\delta)+w]=H+k-b(\delta)$

Now, the optimality of $y^i<1$ implies that P^i_H is not smaller than this cost, for otherwise the customer would prefer not to leave (i.e., y^i would be equal 1). Hence, $P^i_H \geq H+k-b(\delta)$.

Step 2: If $x^i=1$ or $P^i_L=P^i_H$, then $P^i_H=H$.

Proof: In either case this expert's customers will pay P^i_H. If $P^i_H>H$, all this expert's customers and hence all customers on their first visit will be attracted by an offer (Q,Q) such that $P^i_H>Q>H$. Since fraction (1-w) of them have the L problem, there must be an expert for whom (Q,Q) is a profitable deviation. Therefore, $P^i_H=H$.

Step 3: If $k<(1-w)(H-L)/w - e(\delta)$, all experts visited by first period customers are of the type described in step 1.

If $k>(1-w)(H-L)/w$ all experts visited by first period customers are of the type described in step 2.

Proof: Suppose $k<(1-w)(H-L)/w - e(\delta)$. If all experts are of the type described in step 2, then the deviation $(L+a(\delta),\infty)$ will attract all first visit customers, since the expected cost of first going there is lower than the expected cost of going directly to an expert who charges H. That is, as can be verified by direct computation, $e(\delta)$ is such that for k in this range

$$k+(1-w)(L+a(\delta))+w(H+k)<k+H$$

Similarly, it may not be that some experts who get first visit customers are as in step 1 and some are as in step 2, since all first visit customers will prefer to go to the firms described in step 1. This is because, it again follows from the choice of $e(\delta)$ that, for k in this range

$$k+(1-w)(1-x^{i})(L+a(\delta))+[w+(1-w)x^{i}(H+k)<k+H$$

Suppose $k>(1-w)(H-L)/w$. It may not be that all or some experts who get first visit customers are as in step 1, since for k in this range the expected cost of going first to an (H,H) expert is even lower than going first to an $\{(L,\infty),0\}$ expert. I.e., $k+H<k+(1-w)L+w(H+k)$. Thus, if there exists an (H,H) expert, customers will not go to an expert of the type described in step 1; if there is no (H,H) expert, then (H,H) will be a profitable deviation as it would attract all first period customers.

Step 4: If $k<(1-w)(H-L)/w - e(\delta)$, there are experts who serve only second visit customers and charge $P_H^i \leq H+f(\delta)$

Proof: It follows from steps 3 and 1 that in this case, at least a fraction $1-d(\delta)$ of all customers with the H problem, go on to a second expert after the first expert recommended to them the H treatment. These customers are almost certain to have the H problem. Hence, they will prefer an expert with say $P_H<H+k/2$ to an expert with $P_H \geq H+k-b(\delta)$. By steps 3 and 1 all experts who serve first visit customers have $P_H \geq H+k-b(\delta)$. Therefore, it must be that all second

visit customers go to other experts with $P_H<H+k-b(\delta)$, for otherwise a deviant expert with say $P_H<H+k/2$ can profitably attract all of them. Furthermore, by step 1, the second visit customers will agree to an H recommendation from an expert with $P_H<H+k-b(\delta)$. Therefore, the experts who serve the second visit customers end up charging P_H regardless of their customers' problem. Now, it may not be that an expert who serves the second visit customers has $P_H>H+f(\delta)$, since then by deviating to an offer with $P_H=H+f(\delta)$ an expert would attract all second visit customers. By the choice of $f(\delta)$ this would give such an expert profit of at least δ and hence would be profitable for some expert.

<u>Step 5</u>: For any $\varepsilon>0$, choose δ such $\max\{a(\delta),b(\delta),c(\delta),d(\delta),e(\delta),f(\delta)\}<\varepsilon$. This is always possible since all these magnitudes approach 0 as δ approaches. Let $N(\varepsilon)$ be such that there must be an expert whose profit is below δ, say $N(\varepsilon)>2BM/\delta$. The proposition now follows from steps 2-5 after substituting ε for each of the magnitudes $a(\delta)...f(\delta)$. QED

<u>Proof of proposition 3</u>:

Part (i). Consider an expert's deviation $\{(P_L,P_H),x\}$ and let us complete the ME description as follows. Customers' beliefs are that if (a) $P_H\leq H+k$ or $P_L\leq E_1$, then x=1, and if (b) $P_L>E_1$ and $P_H>H+k$, then x=0. The customers' strategy is not to go to the deviant expert.

Note that in both (a) and (b) the customers' strategy is optimal given the beliefs. In the cases described in (a) a customer who visits this expert will pay at least H. In the cases described in (b) the expected cost incurred by a customer who visits this expert is at least $k+[(1-w)+w(1-h)]P_L+wh(k+H)$. This is because with probability $[(1-w)+w(1-h)]$ the diagnosis of this customer would yield the signal "L", while with probability wh it would yield "H". But $P_L>E_1$ and $k>(1-w)(H-L)/wh$ imply that

$k+[(1-w)+w(1-h)]P_L+wh(k+H)>k+[(1-w)+w(1-h)]E_1+wh(H+k)=$

$k+(1-w)L+w(1-h)H+wh(H+k)=k+(1-w)L+wH+whk>k+(1-w)L+wH+(1-w)(H-L)=k+H$

That is, this expected cost is higher than the cost of going directly to an (H,H) expert. Hence, this offer does not upset the ME.

Part (ii). First observe that the postulated customers' behavior is optimal. The expected customer's cost in this case is $k+[(1-w)+w(1-h)]E_1+wh(H+k)$. This is because with probability $[(1-w)+w(1-h)]$ the diagnosis of this customer would yield the signal "L" so the price is E_1, while with probability wh it would yield "H" so this customer will make another visit and pay H. The two alternatives we have to consider are going directly to an (H,H) expert and sampling (E_1,∞) experts more than once. In the first alternative the customer's cost is k+H. But, the condition $k<(1-w)(H-L)/wh$ implies

$$k+[(1-w)+w(1-h)]E_1+wh(H+k)<k+H,$$

so that going directly to an (H,H) expert is not a profitable deviation. In the second alternative, the cost depends on the number of times the customer plans to sample (E_1,∞) experts. Notice, however, that once a customer produced the signal "H", his type is H with certainty. Therefore, if it pays such a customer to sample an (E_1,∞) expert again, it pays him to sample such experts perpetually. The expected cost associated with this alternative is $E_1+k(1+wh/(1-h))$. Now, this together with the condition $k>(1-h)(1-w)(H-L)/h[1-w+w(1-h)]$ imply $k+[(1-w)+w(1-h)]E_1+wh(H+k)<E_1+k(1+wh/(1-h))$, so that sampling (E_1,∞) more than once is not a profitable customer's deviation.

Next consider an expert's deviation $\{(P_L,P_H),x\}$, and let us complete the ME description as follows. Customers' beliefs are that x=0 if $P_L\geq H$ and $P_H>H+k$, and that x=1 otherwise. Customers' strategy is not to go to a deviant expert. Note that the customers' strategy is optimal, given their beliefs, since in all contingencies the price this expert recommends is at least H. Therefore, no such $\{(P_L, P_H), x\}$ deviation upsets the ME.

Part (iv). Consider an expert's deviation $\{(P_L, P_H), x\}$, and let us complete the ME description as follows. Customers' beliefs are that if $P_L<(1-w)L+wH$ or $P_H\leq(1-w)L+wH+k/(1-h)$ then $x=1$; otherwise any arbitrary x. Customers' strategy is not to go to a deviant expert.

Let us verify that the customers' strategy is optimal, given their beliefs. First note that, if $P_L>(1-w)L+wH$ and $P_H>(1-w)L+wH+k/(1-h)$, a customer cannot possibly gain from going to this expert, regardless of x. This is because, after an H recommendation this customer would like to continue elsewhere, while P_L is higher than the ME price. Second, note that in all other contingencies customers believe in $x=1$. Therefore, the only way a customer could possibly gain from visiting such a deviant expert is if he intends to accept the H recommendation. That is, if the cost P_H+k is lower than the expected cost associated with sampling of other experts. Now, even for a customer who has already been diagnosed as H, the latter cost is $(1-w)L+wH+k/(1-h)$, where $(1-w)L+wH$ is the prevailing price and $k/(1-h)$ is the expected search cost until this customer generates the signal "L". But the condition $k<(1-h)(1-w)(H-L)/h$ implies $(1-w)L+wH+k/(1-h)<H+k\leq P_H+k$. Thus, visiting the deviant expert is not profitable for a customer who has been diagnosed as H, and hence it is not profitable for any customer. Therefore, the customers' strategy of not going to this expert is optimal and so $\{(P_L, P_H), x\}$ does not upset the ME.

Part (iii) is just a combination of (ii) and (iv) and we leave its proof to the reader. QED

Proof of proposition 4: Note that, because of the symmetry, on the equilibrium path of a SRE a customer's strategy can be described by the probability this customer returns to his first expert as a function of the result of that visit. That is, the probability he returns after an L treatment, after an H treatment and after a rejection. It is easy to see that, if there is a SRE in which the first or third probabilities are different from 1 and 0 respectively, then there is a SRE where customers' expected cost is at least as low and these probabilities are 1 and 0 respectively. This is because putting these probabilities at 0 and 1 respectively maximizes the expert's reward of not misrepresenting and the punishment for rejection. We shall therefore consider only customer strategies in which these probabilities are 1 and 0 respectively. Let y denote the probability that a customer returns to his first expert after an H treatment.

For the expert side it is enough to consider a SRE's in which the experts' action is truthful, i.e., $\{0,0,(P_L,P_H)\}$. This is because it is immediate that, if there is a SRE in which $x>0$ or $r>0$, then there is a SRE with $x=0$ and $r=0$ in which customers' expected cost is at least as low.

Consider, therefore, a SRE in which the experts' action is $\{0,0,(P_L,P_H)\}$ and the customers return to an expert who treated them for H with probability y. Let $m=M/N$, i.e., m is the number of customers per expert in each generation. Let $\Pi=[(1-w)(P_L-L)+w(P_H-H)]$. Note that at this equilibrium an expert's expected profit per period is $2m\Pi$. Now, if an expert switched to $x=1$ for one period, his profit in that period would be $2m[P_H-(1-w)L-wH]$, in the next period it would be $[m+ym+(1-y)wm][(1-w)(P_L-L)+w(P_H-H)]$, and after that it will return to the SRE level. The number of customers who visit an expert in the period that follows a deviation, $[m+ym+(1-y)wm]$, consists of the m new arrivals, the ym customers who returned to this expert after having been charged P_H last period, and the (1-y)wm who represent this expert's share in the old customers who switch an expert in this period. The optimality of x=0 therefore implies that

(7) $$4m\Pi \geq 2m[P_H-(1-w)L-wH] + [m+ym+(1-y)wm]\Pi$$

Similarly, if an expert switched to r=1 for one period, his profit in that period would be $2(1-w)m(P_L-L)$ and in the next period it would be $[m+(1-w)m+(1-y)wm][(1-w)(P_L-L)+w(P_H-H)]$. The optimality of r=0 therefore implies that

(8) $$4m\Pi \geq 2(1-w)m(P_L-L) + [m+(1-w)m+(1-y)wm]\Pi$$

Thus, any SRE has to satisfy (7) and (8).

Next, let us solve the problems: (i) find P_L, P_H and y to $\min\{(1-w)P_L+wP_H\}$ s.t. (7) and (8); and (ii) find P_L, P_H and y to $\min P_L$ s.t. (7) and (8). First, observe that in the solution to any of these problems it may not be that both (7) and (8) are non-binding. Second, observe that if either (7) or (8) is non-binding, then y=0 or 1. This is because, if 0<y<1 the binding constraint can be relaxed by changing y and a lower value of the objective $(1-w)P_L+wP_H$ or P_L can be achieved. Third, if both (7) and (8) bind, they can be solved to yield

(9) $$P_L = (a+by)/(c+dy)$$

(10) $$P_H = (3-w)P_H/(2-3w) + e$$

where a,b,c,d,e are constants which depend on w, L and H. But by substituting from (9) and (10), we get that both $(1-w)P_L+wP_H$ and P_L are either monotonically increasing or a monotonically decreasing functions of y. Thus, if both (7) and (8) are binding, then at the solutions of either problem y=0 or 1. It follows that either y=1, in which case (8) is binding and the solution is $P_L=P_H=H-(1-w)(H-L)/(1+2w)$, or $y=0$, in which case (7) is binding and the solution is $P_L=L+2(H-L)/(3-w)$, $P_H=H$. By direct comparison it is now easy to verify that $P_L=P_H=H-(1-w)(H-L)/(1+2w)$ and $y=1$ is the solution to minimization problem (i) and, for $w \leq 2/3$, it is also the solution to minimization problem (ii) above; $P_L=L+2(H-L)/(3-w)$, $P_H=H$ and $y=0$ is the solution to minimization problem (ii) above, for $w>2/3$.

Note that if there is a SRE with $P_L=P_H=H-(1-w)(H-L)/(1+2w)$, it is the one that minimizes $(1-w)P_L+wP_H$ and also the one that minimizes P_L for $w \leq 2/3$, since any truthful SRE has to satisfy (7) and (8). And, similarly, if there is a SRE with $P_L=L+2(H-L)/(3-w)$ and $P_H=H$ it is the one that minimizes P_L, for $w>2/3$. Therefore, it remains only to establish that there are such SRE. For the configuration with $P_L=P_H=H-(1-w)(H-L)/(1+2w)$ condition (8) already assures that experts have no incentive to deviate and reject a customer who has the H problem; for the configuration with $P_L=L+2(H-L)/(3-w)$, $P_H=H$ condition (7) assures that experts have no incentive to charge for H when the problem is L. To complete the proof it remains to show that, in each of these cases, other possible deviations are also unprofitable. For both cases, let us complete the description of customers' beliefs, following an expert's deviation to some schedule $(Q_L,Q_H) \neq (P_L,P_H)$, in a way that makes it optimal for customers not to visit such an expert. If $Q_H \geq H$, let customers believe that this expert has $x=1$ and $r=0$; if $Q_H<H$, let customers believe that $r=1$. With these beliefs, it is best response for customers to avoid the deviant expert, in each of the two cases discussed here. This establishes that both the configuration in which all experts charge $P_L=P_H=H-(1-w)(H-L)/(1+2w)$ and treat all customers, and the configuration in which all experts have $P_L=L+2(H-L)/(3-w)$ and $P_H=H$ and charge truthfully can be supported as SRE. QED

Footnotes

1. Although evidence on seller honesty is naturally difficult to obtain, regulatory agencies have made some attempts to study this aspect. A field study of the optometry industry conducted by the FTC(1980) documented consistent tendency by optometrists to prescribe unnecessary treatment; a survey of 62 automobile repair shops conducted by the Department of Transportation found that 53 cents of each service dollar were charged for needless repairs (see New York Times, 1979).

2. It is interesting to note that this experimental study was sponsored by the FTC. The difficulty of gathering direct evidence on behavior in such industries prompted the FTC to sponsor such experimental work as well as a field study of the optometry industry as reported in FTC (1980).

3. Different search durations of L and H customers imply that the distribution of problems in the population may differ from what it is in the entering batch, and the dynamic setting will allow us to capture this. While this feature is not too important in the first parts in which the expert diagnosis is certain, it plays a more important role when the diagnosis is uncertain.

4. E.g., we shall assume that exactly fraction w of the entering batch have the H problem, and if customers are exactly indifferent among some 10 experts, then exactly 1/10 of them will go to each etc.

5. There is no contradiction between an expert's decision to provide only the L service and this expert's commitment to fix the problem even if it turns out to be H. If, however, we adopt the interpretation of alternative technologies described at the end of section 2, the actual computations will slightly change as they will have to incorporate the cost H', but the basic argument will remain the same.

References

Arrow, K. (1963), "Uncertainty and the Economics of Medical Care," AER, 53, 941-973.

Darby, M. and E. Karni (1973), "Free Competition and the Optimal Amount of Fraud," Journal of Law and Economics, 16, 67-88.

Diamond, P. (1971), "A Model of Price Adjustment," JET, 3, 156-168.

FTC (1979), Consumer Information Remedies--Policy Review Session.

FTC (1980), Staff Report on Effects of Restrictions on Advertising and Commercial practice in the Professions: The Case of Optometry.

Klein, B. and K. Leffler (1981), "The Role of Market Forces in Assuring Contractual Performance," JPE, 89, 615-641.

"Faulty Car Repairs Found Widespread," New York Times (May 8, 1979)

Pitchik, C. and A. Schotter (1987), "Honesty in a model of Strategic Information Transmission," AER, 77, 1032-1036.

Pitchik, C. and A. Schotter (1989), "Information Transmission in a regulated Market," AER, 77, 1032-1036.

Plott, C. and L. Wilde (1980), "Toward A Theory of Professional Diagnosis and Service: Consumer Behavior," Social Science Working Paper 352, California Institute of Technology.

Plott, C. and L. Wilde (1991), "Professional Diagnosis vs. Self-Diagnosis: An Experimental Examination of Some Special Features of Markets with Uncertainty," in V. Smith (ed) Research in Experimental Economics, vol. II, JAI Press.

Wolinsky, A. (1983), "Prices as Signals of Product Quality," RES, 50, 647-658.