FREE COMPETITION AND THE OPTIMAL AMOUNT OF FRAUD*

MICHAEL R. DARBY and *EDI KARNI*
University of California, Los Angeles *University of Tel Aviv*

THIS paper explores the reasons for and determinants of the provision by a firm of false information to a consumer so as to induce purchases which would not be made if the consumer possessed full information about the qualities of his purchase. Market responses to potential for fraud are analyzed in detail. It is shown that fraud and related practices follow from significant costs both in the determination of quality of a particular good or service and in the effective vertical integration of seller and buyer through some exchange of property rights.

Much of our discussion focuses on the key problem of the joint provision of diagnosis and services—such as the choice and execution of an automobile repair or taxicab route—but the model developed will be seen to have general applicability whenever the seller provides information which influences purchases, as through advertising or salesmen's promises.

In the context of the repair problem, we explore the reasons for and the determinants of the provision of repair services in amounts greater than would be economically efficient, given the price of the services and their marginal product in terms of the service flow from the repaired commodity.[1] The possibility of this situation is suggested by the observation that in a considerable number of cases involving medical, automotive, and other repair services, contrary to the basic assumption of conventional demand theory, the consumer is unaware of the ability of the repair service to satisfy a given want.[2]

* This paper was prepared while the authors were at the Ohio State University. The authors would like to acknowledge helpful comments on earlier versions of this paper by Gary Becker, Gregg Lewis, Marc Nerlove, George Stigler, Lester Telser, and other members of the Applications of Economics Workshop at the University of Chicago and by members of the Econometrics Workshop at the Ohio State University.

[1] This is related to, but not identical with, the provisions of repair services in greater amounts than would be the case if the repairer and owners of the object were the same, as will be shown *infra*.

[2] This is the other side of the "information is costly" coin, which implies that an individual renting specialized knowledge can evaluate only the results and not the procedure. Reports of "overselling" are a staple for the news media. Recent examples are the report that Dr. John Knowles, outgoing director of Massachusetts General Hospital,

Hence, the contention that competitive markets are sufficient to prevent fraud by, at least, established firms, because of the effect on future sales of the eventual discovery of fraud, does not hold in this case. The provision of joint diagnosis and repair implies that some fraud can be successful because of the high, if not prohibitive, costs of discovery of the fraud. If a part is changed prematurely, say, then the new part will operate in the same manner as if the replaced part were really faulty. Similarly for a premature removal of an appendix. Under such circumstances the consumer is sold services which, if he were adequately informed, would not have been purchased. Alternative market arrangements replace perfect competition so as to economize in the gathering of informaton about the quality of services rendered. Proposals for governmental intervention in markets which do not completely eliminate the regular practice of fraud are considered. Much the same costs and rewards are present for governmental investigators as for private evaluators of services, and thus there is no apparent reason why governmental intervention should cause an increase in welfare sufficient to pay its cost.

Section I analyzes the behavior of a consumer who is unfamiliar with the exact qualities of a particular purchase. Section II discusses the behavior of a seller of a product with qualities measurable only at considerable expense and the reasons that he would, at times, engage in the provision of fraudulent information about the qualities of its product. Section III discusses the market equilibrium amount of fraud and the market response to fraud through the internalization of benefits of honesty by branding and in the form of quasi-property right which we term the client relationship. Section IV analyzes governmental intervention in such markets. The conclusions are summarized in Section V.

I. Consumer Analysis

In a recent article, Nelson has distinguished between the "search" and "experience" qualities of a good, where search qualities are those that can be ascertained in the search process prior to purchase and experience qualities are those that can be discovered only after purchase as the product is used.[3] We find that it is important to distinguish a third class of properties which we term "credence" qualities. Credence qualities are those which, although

"says up to 40 percent of the nation's doctors are 'making a killing' in overcharging patients, performing unnecessary surgery and in other 'gouging' practices." Columbus Dispatch, January 14, 1972. Similar charges against auto and TV repair industries in California are reported in The Wall Street Journal, January 11, 1972.

[3] Nelson suggests as an example of each the style of a dress and the taste of a can of tuna fish, respectively. Phillip Nelson, Information and Consumer Behavior, 78 J. Pol. Econ. 311 (1970).

worthwhile, cannot be evaluated in normal use. Instead the assessment of their value requires additional costly information. An example would be the claimed advantages of the removal of an appendix, which will be correct or not according to whether the organ is diseased. The purchaser will have no different experience after the operation whether or not the organ was diseased. A similar example would apply to replacement of a television tube, certain automobile repairs and the like. The line between experience and credence qualities of a good may not be always sharp, particularly if the quality will be discerned in use, but only after the lapse of a considerable period of time.

We distinguish then three types of qualities associated with a particular purchase: search qualities which are known before purchase, experience qualities which are known costlessly only after purchase, and credence qualities which are expensive to judge even after purchase. In this section we will first examine the case of a good characterized by only the first and third types of qualities. The model is then easily generalized to apply to goods with any mixture of qualities.

Credence qualities arise whenever a good is utilized either in combination with other goods of uncertain properties to produce measurable output or in a production process in which output, at least in a subjective sense, is stochastic, or where both occur. Much attention has been focused in public debate upon the credence qualities involved in the provision of repair services, which will serve as a focus for our discussion.[4] A consumer cannot fully evaluate the repair of a malfunctioning durable good or human being, since he is unfamiliar with the intricacies and peculiarities of the particular machine. Due to lack of knowledge of the inputs of repair services required to maintain given flows of ultimate services from a commodity, the consumer must purchase both information and repair services. If there were no additional costs involved in buying diagnosis and the actual repair service from different sources, then the consumer will generally do so, thereby avoiding the temptation to fraud at no additional cost.[5] If, on the other hand, the cost function is inseparable in the sense that it is cheaper to provide information and service jointly than separately, then the individual tends to buy both information and repair

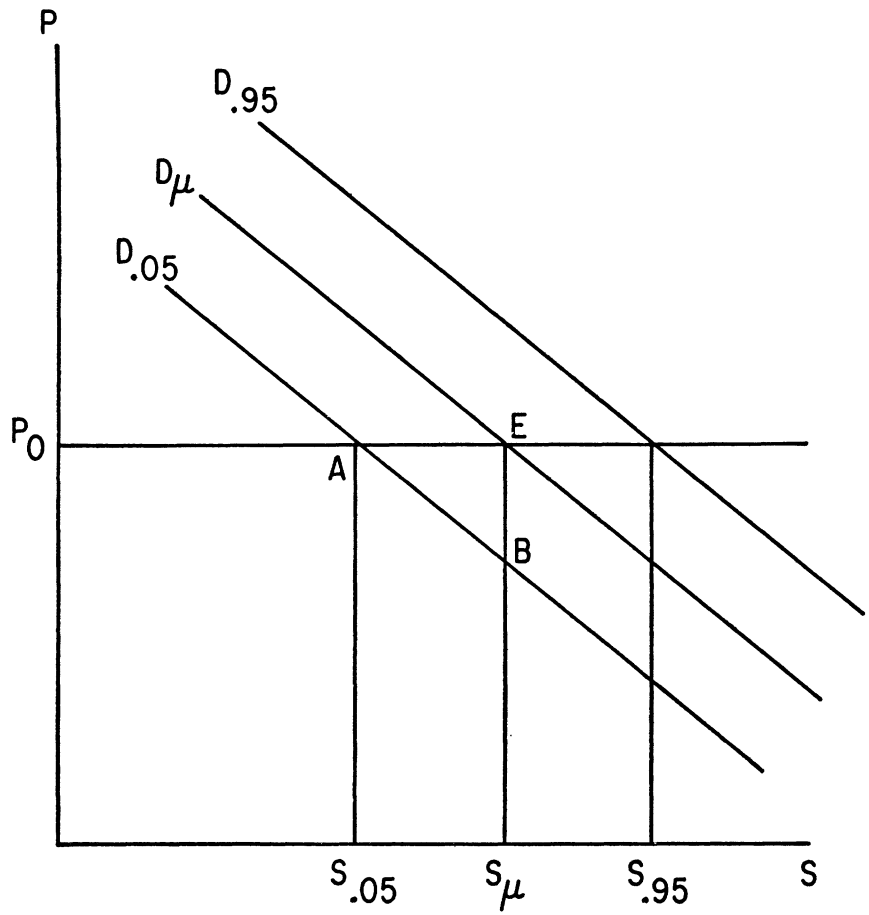
[4] The model is really applicable to a large class of purchases: Consider the route taken by a taxicab in Paris according to whether the trip is requested by a Parisian or an American tourist. Similarly the claimed life of a "one-hoss shay" type durable cannot be experienced for some time, and considerable purchases may be made prior to discovery of misinformation.

[5] An example is the separation of prescription and preparation of drugs. The costs of separation of diagnosis and repair, in general, involve not only possibly increased monetary costs, but also may involve additional time or other transactions costs, which must be compared with any expected reduction in fraud.

services from the same source.[6] Even in such a case, if the information were costlessly checkable, or if the purchaser possessed the ability to judge and associate the quality of output with the price of that output, he could not be defrauded for long and indeed the competitive market will largely eliminate continuous attempts at fraud. If, however, information is not costless, and output is a stochastic function of service input, then the possibility of the existence of fraud has significant scope of interest. A most transparent case of such a relationship is between repair services and the flow of services from sophisticated durable goods, such as from automobiles, electronic equipment and human beings. This point is most easily illustrated by considering the derived demand for repair services on the part of an expert buyer. By an expert buyer we mean someone who can adequately assess the actual production function relating the ultimate service flow to the repair inputs. The derived demand schedule of the expert buyer is given by the value of marginal product of the repair services, which in a stochastic world depends on the particular durable good or human being under consideration. Assume that the probability distribution has finite mean and variance. Then Figure I illustrates three derived demand curves for repair services. The curve denoted D_μ is a demand curve for repair services on the part of an expert buyer who owns a durable good which happens to be an "average unit" in the sense that the flow of services from this unit coincide with the mean of the distribution of flow of services for this particular good. The derived demand for repair services for a unit which perform above the average, is to the left of D_μ. One such curve for a unit which performs better than 95 per cent of the units is denoted $D_{.05}$. Similarly, the derived demand for services for a unit which performs worse than 95 per cent of the units in the market is denoted $D_{.95}$. Given the price of repair service, P_0, the expert buyer buys $S_{.05}$, S_μ, $S_{.95}$ or any other amount which depends on the particular unit which he possesses. The non-expert buyer, who by definition is unable to adequately assess the production function, can be led to believe that he owns a unit whose performance level is inferior to its actual level of performance. For example, a consumer who owns a unit which performs better than 95 per cent of the units in the market, and thus should behave according to the demand curve, $D_{.05}$, might be misled by a supplier of information to believe that the actual performance of his unit is equal to that of the average unit. Hence, at P_0 the non-expert consumer buys S_μ of services. He pays $(S_{.05}\text{-}S_\mu)$ P_0 for marginal services worth only the area $S_{.05}ABS_\mu$. This consumer, according to our definition,

[6] For example, it is easier to repair any damage while the transmission or belly is open to see what is wrong, than to put everything back together and go elsewhere to repeat the process for the actual repair.

FIGURE I



has been defrauded. The value of the loss from this fraud is given by the area ABE.

In case of complete ignorance on the part of the consumer, the amount of fraud is limited only by the price differential between a new unit and the present one. In less extreme cases, it seems plausible to assume that the non-expert buyer substitutes information about repair expenditures for exact technical knowledge,[7] and as a result possesses some notion with respect to a

[7] A related hypothesis has been used in the literature to explain an upward sloping demand curve when an inexpert buyer uses price as an indicator of quality.

subjective probability distribution of repair costs. It can be shown (see the Appendix) that for sufficiently subjectively improbable price quotations, the consumer will choose to take his future business elsewhere because of the likelihood of fraud. For repair quotations which exceed a higher subjective limit he will pay the price for diagnosis but no repair, a sunk cost, and try at another shop for a different diagnosis at a lower cost. At any point in time, the individual's decision is a function of predetermined variables[8] and the quoted price. Different individuals will have different decision regions with respect to the quoted price, reflecting differences in subjective probability distributions, preferences, and wealth.

The model generalizes easily to take account of experience qualities of goods. Here too the higher the price, the less likely is any consumer to believe it worthwhile to purchase a particular good. However with respect to experience qualities after the sale is made, the consumer is able to judge what is received. The greater the discrepancy between promised and actual experience qualities, the less likely the customer is to do further business with the same firm and the greater the likelihood that it will be worthwhile to recover damages or file charges for fraud.

II. The Supply Side

Section I demonstrated that incomplete information on the part of the individual consumer creates favorable conditions for fraud. This section develops a simple model to investigate the motivations of a competitive firm to engage in fraudulent practices, and to determine the optimal amount of fraud for such a firm. For this purpose we first assume that a repair firm faces a stochastic demand for its services, and analyze the incentive to engage in fraud under two exclusive and exhaustive states of the demand for repair services. The first state obtains during time intervals when the length of the queue of customers waiting for service is zero. The second state obtains during periods when the length of the queue of customers waiting for service is positive. The specific nature of the fraudulent practices is shown to vary accordingly. Later in this section, fraud with respect to experience qualities will be examined.

Consider a competitive firm producing repair services subject to a given probability distribution of service time, and random flow of customers. Assume that the firm is in a long-run equilibrium position. Assume further that the probability of the event "no customers waiting for service in the time interval $t + \Delta t$" is positive. Then this probability equals the average fraction of total working hours during which the service post is idle. Since the accumulation of

[8] Such as prior experience with the particular durable and repair shop, the fixed cost of diagnosis, his wealth and his attitude toward risk.

inventories is ruled out by the assumption that the firm produces services, some of the variable costs, and in particular labor costs, become fixed cost during those time intervals. This reduces the marginal costs of operation during time intervals when no customer is waiting for service. Hence, there is an incentive to prolong the service term of customers receiving service by selling repair services in amount greater than would have been purchased by an expert buyer, even at increasing costs. These increasing costs indeed exist in the form of increasing the risk of losing the "good will" of the customer as a result of trying to convince him to buy more services. In more formal terms, increasing the amount of services prescribed on the basis of the diagnosis, increases the probability of entering the customer's critical regions for going elsewhere now and in the future. Such events result in losing the present value of present and/or future sales to this particular customer, and other customers to whom he may pass the information. Taking this consideration into account, the firm will carry fraud up to the point where the expected marginal profit is zero.

The discussion can be greatly simplified by the use of simple mathematics and the following notation:

π = the expected present value of total profit from a customer.[9]
P = the price of the repair services.
S = the total amount of services prescribed.
V = the anticipated present value of future profits from services to a given customer.
C(S) = the total variable costs of services prescribed.[10]

The expected profit from a particular customer is equal to the profit if the diagnosis is accepted times the probability it will be accepted plus the present value from future sales to the customer times the probability that the customer will return. This is given algebraically as,

$$\pi = [P \cdot S - C(S)][1 - F(S)] + V \cdot [1 - H(S)], \qquad (1)$$

where F(S) is the cumulative probability function that the customer will refuse the present services and H(S) is the probability that the customer will take his future business elsewhere.[11]

[9] The problem is simplified by exclusion of the fixed fee for and costs of diagnosis.

[10] These costs are those related to the actual amount of services, such as parts. Costs such as labor which are fixed for the period are excluded. There may, however, be a positive shadow price or scarcity cost applied to the fixed factors due to having to give up some fraud to other customers. This factor will be discussed further in Section III, *infra*. For the moment, we assume a situation which short-run marginal costs are less than long-run marginal costs.

[11] The firm may use different subjective F(S) and H(S) functions for different customers according to the firm's estimate of a particular customer's knowledgability and the like.

In equilibrium, the returns to prescribing additional services (the price less the marginal cost all multiplied by the probability of making the sale) must be equal to the costs in terms of increased likelihood of losing the value of both the present and future sale. This can be proven by differentiating the profit function with respect to the decision variable S, and setting the result equal to zero:[12]

$$\frac{d\pi}{dS} =$$

$$[P - C'(S)][1 - F(S)] - [P \cdot S - C(S)]\, F'(S) - V \cdot H'(S) = 0 \tag{2}$$

or

$$[P - C'(S)][1 - F(S)] = [P \cdot S - C(S)]\, F'(S) + V \cdot H'(S). \tag{3}$$

In order for fraud to occur, it must be true that at the amount of services which would be purchased by an expert buyer, S*, the marginal return to selling additional services is greater than the marginal cost in terms of the expected loss of business; that is

$$[P - C'(S^*)][1 - F(S^*)] > [P \cdot S^* - C(S^*)]\, F'(S^*) + V \cdot H'(S^*). \tag{4}$$

This can clearly be met where there are significant costs of diagnosis. In this model there are circumstances in which the firm would sell less than the optimal amount of services or charge less than their cost in order to retain the customer's good will.

An important implication of this analysis is that the higher the anticipated present value of future profits from a customer, the smaller is the tendency to defraud him.[13] Therefore tourists and casual customers are likely to be

---

The distributions F and H will be affected by the optimum amount of service for the particular durable to the extent the customer is likely to be able to judge. These distributions are assumed to be smooth in the relevant region so that F'(S) and H'(S) exist.

[12] The second order condition is

$$\frac{d^2\pi}{dS^2} = -C''(S)[1 - F(S)] - 2[P - C'(S)]F'(S) - [P \cdot S - C(S)]F''(S) - V \cdot H''(S) < 0.$$

[13] To prove this proposition, differentiate the equilibrium condition (3) with respect to the parameter V to obtain

$$\frac{dS}{dV} = \frac{H'(S)}{-C''(S)(1 - F(S)) - 2[P - C'(S)]F'(S) - [P \cdot S - C(S)]\, F''(S) - V \cdot H''(S)}$$

The numerator is positive, while the demoninator is negative by the second order condition. Hence, $dS/dV < 0$.

defrauded to a greater extent than regular and steady customers. A case of particular interest is perfect competition in which case $V = 0$. Here the necessary condition for maximum profit requires that the ratio of marginal to total return is equated to minus the ratio of the marginal to total probability of making the sale:

$$\frac{P - C'(S)}{PS - C(S)} = \frac{F'(S)}{1 - F(S)} \tag{5}$$

Another implication pertains to the effects of changes in the customers ability to assess the services prescribed on the optimum amount of fraud for the firm. The following analysis is somewhat complicated, but the conclusion is that improved knowledge on the part of the consumer under observed conditions, will reduce the firm's optimal level of fraud. Changes in ability to assess the services may be described as shifts in the cumulative probability functions, $F(S)$ and $H(S)$. In particular, if S^* is the level of services which would be purchased by an expert buyer, then for levels of service S higher than S^* improved knowledge would increase the values of the cumulative probability functions $F(S)$ and $H(S)$. Similarly for values of S less than S^* $F(S)$ and $H(S)$ would decrease under improved knowledge. Thus the $F(S)$ and $H(S)$ functions should converge to a vertical line between 0 and 1 at $S = S^*$ as knowledge improves. These intuitive ideas can be introduced by rewriting equation (1) as

$$\pi = [P \cdot S - C(S)][1 - F(\theta)] + V \cdot [1 - H(\theta)] \tag{6}$$

where $\theta = \theta(S,b) = S^* + (S - S^*)b$ and $0 \leqslant b \leqslant \infty$. Here b is a shift parameter representing the state of knowledge. An increase in b increases the values of $F(\theta)$ and $H(\theta)$ for a given $S > S^*$ and decreases these values for a given $S < S^*$ as required. If the fixed functions $F(\cdot)$ and $H(\cdot)$ are defined for an initial state of knowledge represented by $b = 1$, then it is straightforward to examine marginal changes in knowledge. The first and second order conditions derived from equation (6) for $b = 1$ can be shown to be identically equal to those given in equation (2) and footnote 12 by use of the conditions that $\theta = \theta(S,1) = S$ and $\frac{\partial \theta}{\partial S} = 1$. If we differentiate the first order condition with respect to b, we obtain the effect on the equilibrium value of S due to marginal change in knowledge as represented by b:

$$\frac{dS}{db} = \frac{[P - C'(S)]F_b + [PS - C(S)]F_{Sb} + VH_{Sb}}{-C''(S)[1 - F(S)] - 2[P - C'(S)]F'(S) - [PS - C(S)]F''(S) - VH''(S)} \tag{7}$$

where

$$F_b = (S - S^*)F'(S)$$

$$F_{Sb} = (S - S^*)F''(S) + F'(S)$$

$$H_{Sb} = (S - S^*)H''(S) + H'(S)$$

The denominator of equation (7) is negative by the second order condition. The sign of the numerator is ambiguous even for the case of interest in which the initial equilibrium $S > S^*$. Although $F'(S)$ and $H'(S)$ will be positive, $F''(S)$ and $H''(S)$ might be sufficiently negative that the sign of the numerator could be negative. For a normal cumulative probability distribution function, however, $F''(S)$ and $H''(S)$ are negative only for $F(S) > .5$ and $H(S) > .5$ respectively. This means that more than half of the time customers must refuse the initial offer or refuse to return in order for even $F''(S)$ or $H''(S)$ to be negative. Since this is not generally the case, it seems very likely that the numerator will be positive so that $\frac{dS}{db}$ would be negative. Therefore, it is reasonable, though not logically necessary, that improved knowledge on the part of the consumer results in a reduction in the optimal amount of fraud for an individual firm.

An alternative case occurs when the length of the queue of customers waiting for service is positive. Here there is no longer incentive to oversell in order to utilize idle resources, and the firm will try to minimize the indirect cost from lost goodwill by reducing the amount of services to the optimal amount or even less. Fraud may still take place, however, but it will be of a somewhat different nature. Two forms of fraudulent practices may appear: (1) Services reported but not provided—the customer may agree to extra prescribed services which are then not provided. This may increase the probability of loss of goodwill to the extent that the customer can eventually detect that the claimed services were not provided. (2) Resale of used parts—if good parts from a durable good can be resold, then this will increase the return to prescribing replacement of a supposedly broken part which is in fact in good condition. These alternative forms of fraud can be analyzed with only trivial changes in the basic model.

All of these types of fraud may occur simultaneously, but charging for services not provided increases the risk of detection significantly. Otherwise the discussion is similar to the first case.

To the extent that goods are sold as bearing more costly experience qualities than they in fact possess, fraud is accomplished analogous to charging for services not provided.[14] However, the fraud is sure to be detected so that lack

[14] It would be possible to consider a more complicated model in which the entrepreneur

of customer repeats and bad word-of-mouth advertising is much more likely. The associated costs vary with the frequency of purchase and the rate of depreciation of market information. Also to the extent that fraud is significant enough to justify the costs of prosecution the term should also take account of the expected costs of prosecution.[15]

III. The Equilibrium Amount of Fraud and the Client Relationship

We have shown in the preceding sections that consumer ignorance and additional cost of separate diagnosis and repair provide motivation for a service firm to defraud its customers. We now turn to an analysis of the market's response to the possibilities of fraud in terms of competing arrangements for the servicing of durable goods. This analysis will give some insight into the effect of fraud—in our strict sense—on the total cost of consuming the services of a durable good, the meaning of negative fraud, and the motives for a client relationship.

In the strictest competitive framework for the model, the present value of future relationship with the customer would be zero, since there is no reason for the customer to return in the future in any case.[16] Fraud would thus be positive and, to the extent that consumers gather their observations on repair costs from their neighbors, limited only by the eventual high return on technical knowledge, which will induce acquisition of such knowledge or of paying the extra cost involved in separation of diagnosis and repair. But this is a rather anemic case, which does not take account of the ability of the market to develop competing arrangements which alter or remove the incentive to engage in fraudulent practices. The question of market equilibrium therefore focuses on what arrangement of property rights minimizes the total cost of obtaining the desired service flows from a durable good. The particular set of arrangements will be shown to vary with the particular case.

The central problem in fraud is the existence of credence qualities due to the joint production of directly estimable output by cooperating factors of uncertain quality and quantity. For example, the consumer can accurately evaluate the transportation services obtained by the use of an automobile but there is no way to know the marginal product of, say, automobile repair services, unless the quality of the automobile itself, of the fuel inputs, and of

---

is treated as maximizing discounted expected cash flow with respect to price, quality, and sales effort, but this is an unnecessary complication for the current analysis.

[15] Note that these costs can be reduced by flight of the seller, which should thus be more common with respect to experience than with credence qualities and with greater than with lesser frauds.

[16] The current discussion abstracts from penalties for being detected in fraud. The differential effect of these penalties will be discussed *infra*.

the care in driving are known and the production function relating these inputs to the output is nonstochastic. Otherwise the value of repair services rendered is not known and must be estimated. The situation is perfectly symmetric with respect to the other inputs, except that information may be known about some of their values by direct observation such as if the car is owner-driven. If there were only one variable input, then a contract could be made with the producer of that input to provide a given level of the easily mensurable output even though the quantity of the input is costly to measure.[17] Fraud would then be simple breach of contract, and could be dealt with in those terms. Some output may be mensurable in subjective but not objective terms, where quality is a matter of taste, and there might be some room for fraud or misrepresentation with respect to this type of experience quality, but this has also been discussed before.[18] The current wave of consumerism has been concerned in large part with what we have called fraud with respect to credence qualities,[19] and the rest of this paper will deal with that problem.

Market arrangements which improve the estimation or monitoring of various inputs serve several purposes, only one of which is the elimination of fraud. In addition, improved estimation rewards more accurately higher quality or competence of inputs so that it is worthwhile to bear more costs to improve quality which is a move toward efficiency.[20] Estimation or monitoring can be effectively improved by arrangements which increase the ability of the consumer to judge the services rendered, which increase the goodwill value of the consumer, or which otherwise shift the losses from inferior services to the provider.

There are some users, notably business firms, which have a sufficiently large demand for repair services that they are able to reach the neighborhood of the flat portion of the average cost curve so that it is efficient for the firm to operate its own repair shop, thus eliminating fraud through internalizing its costs.[21] We are interested in those cases where this is not an attractive

[17] There would still be an insurance problem if the production function is stochastic, but this is discussed *infra*.

[18] See, for example, Phillip Nelson, *supra* note 3.

[19] There has also been an important strain which denies consumers' valuations entirely, but that is a different problem.

[20] Alchian and Demsetz have considered this effect for a firm's employees. Armen Alchian & Harold Demsetz, Production, Information Costs and Economic Organization, 62 Amer. Econ. Rev. 777 (1972).

[21] We abstract here from the firm's problems of estimation and monitoring and its effect on output. See Michael R. Darby, Entrepreneurship, the Firm, and the General Market Information Problem, (Univ. of Calif., L.A., Dep't of Econ., Disc. Paper No. 33, 1973).

alternative and market purchases of goods and services with credence qualities are made, but the "fleet" case of owner-repaired durables makes a good control with which to compare the operation of the market repair shops of our example.

Generally there are daily variations in the derived demand for services for each fleet car, say, according to whether there have been breakdowns, the length of time since previous preventive maintenance operations and the like. The shop will schedule maintenance in such a way that there will be a shadow price, corresponding to P_0 in Figure I, at which the quantities of service "demanded" for each automobile sum to the total amount of services available. If this shadow price were regularly greater than the long run marginal cost, the firm would expand the shop and vice versa in the case of persistently low shadow prices. Thus, the optimal situation would involve fluctuating prices for service according to the stochatic demand faced by a service shop during a specific time interval.[22]

Similarly, in the model of Sections I and II, the amount of services provided by a competitive repair shop to each durable good presented for service on a particular "day" will vary inversely with the strength of the summed demand of all durable goods presented. That is, when demand is low, the shop will be kept busy performing "early" preventive maintenance and the like. When the shop is busy, the shadow price of the facilities will be increased to the point that the effective marginal cost is greater than the price and the firm will undersell, a kind of "negative fraud." Unlike the fleet case, however, the shop owner must service only the cars actually presented to him rather than those of his customers most suitable for service given the current shadow price. Further, since the costs and benefits of the services are not fully internalized, the allocation of the services performed will not only be determined on the basis of the derived demand curve for each particular car, but also according to the value of V, and to the different estimates of the F and H functions associated with each customer. Thus we see that repair services are allocated in a loosely analogous way to the case of the fleet. The amount of services performed per durable good varies according to the relative strength of total demand.[23] But factors other than the derived demand curves affect the allocation of services among the individual durable goods.

A contractual transfer of property rights in the durable good to the servic-

[22] If there were no costs of information, then the fluctuation in shadow prices could be reduced by transfers of cars from busy to relatively idle service shops.

[23] Note that the variation in the shadow price will be greater than for the fleet case because of inability to call in those durables for which service would be optimal with respect to the current shadow price. In some cases customers are called in when demand is low, such as for dental prophylaxis.

ing firm internalizes the cost of overselling repair services. This eliminates fraud directly. The major examples of this technique are service contracts, leasing arrangements, and extensive warranties.[24] Each of these methods combines an insurance component and a service component. There is insurance because the price paid for service does not vary with the amount of service provided. Therein lies the major flaw with these methods—they essentially eliminate quality and fraud problems in repair services, but do so at the cost of eliminating self-monitoring of care by the user. In insurance terminology, moral hazard arises because the marginal return to breakdown prevention on part of the user is zero. Therefore we observe these techniques used only where user effects on repair requirements are minor or easily detectable and contractually excludable. Where there is a difference in the importance of user care, split warranties or the like are often observed with different terms for different types of repairs, such as a five year warranty on motor or sealed system and a one year warranty on other repairs. There are other difficulties with these arrangements in terms of firm size, taxation, and the like, but the essential difficulty appears to lie in the fact that they merely shift, not eliminate, the monitoring problem.

A less costly market solution in some instances is an informal service contract called the client relationship. A client relationship is an implicit understanding that the customer will return for future services so long as he does not detect fraud or low quality services. The client will be better able to judge the quality and costs of services over an extended period of repeated sampling. This reduces fraud (and increases quality) both by improving the customer's judgement and by establishing the customer's "goodwill" as a quasi-property right of the firm, part of its organizational capital as an ongoing business. Thus there is a demand side motivation, besides any possible technical or locational factors, for the establishment of client relationships. The client relationship is, in effect, an informal service contract in which repair services are provided over time at a fixed price per unit but allocated according to the stochastic demand to the shop.[25] The consumer still bears the marginal costs of increased repairs due to any misuse on his part. The firm will undersell its clients when demand is high and oversell them when demand is low, with respect to the announced price by varying the timing of services. This response to fluctuating demand is directly analogous to Oi's

[24] For example, Sears, Roebuck and Co. sells renewable service contracts on a variety of home appliances. The "Kaiser Plan" for medical care is essentially such a service contract for humans. It is possible to lease automobiles and other durable goods from a firm which handles the servicing of the goods. Warranties up to five years have been tried for automobiles.

[25] The risk as to the quality of the particular durable also remains with the consumer.

theory of labor as a quasi-fixed factor.[26] This sort of "fraud" is largely to the benefit of the client who pays a lower bill over time due to the reduction in the announced price resulting from a reduction in idle time. The closer the client relationship, the closer the result approaches the fleet case. Any *net* overselling is clearly not in the client's interest, but there may be no alternative with lower total costs.

An alternative, possibly complementary market arrangement effectively converts the two credence goods—the durable good and its servicing—into a single experience good, by combining the provision of sales and service into a single market entity. When a consumer is considering the purchase of a durable good he essentially compares the discounted value of the expected future service stream with the discounted value of the price and repair costs. Clearly higher repair costs due to fraud reduce the price a seller can charge for his product. Thus a seller which established a non-fraudulent service department for use by its customers can establish a reputation which will allow sales at a higher price.[27] Whether the manufacturer or a dealer will be the operative agency would appear to depend on the optimum size of service departments, and we observe different situations in different industries.[28] The firm has an incentive to correctly monitor its outputs of product and services precisely because the individual is able to monitor his own inputs into the joint production of the consumption flow.

Franchise operations or national firms can extend the advantages of client relationships and internalization through joint sales and service, but only at the cost of increased monitoring costs within the firm.[29] As a result, a larger fraction of the customers of a national chain store will be newcomers to the area in which they are shopping than will be the case for a local store. Increased mobility will increase the relative importance of national chains and conversely.

We can now sharpen and generalize the interpretation of our basic model.

[26] See Walter Y. Oi, Labor as a Quasi-Fixed Factor, 70 J. Pol. Econ. 538 (1962).

[27] Note that this is an extension of Ronald Coase's discussion of social costs, with fraud effectively eliminated by a market transfer of the social cost of fraud to the seller where it can be internalized and acted upon. See R. H. Coase, The Problem of Social Cost, 3 J. Law & Econ. 1 (1960).

[28] It is clear that this internalization of the losses from fraud requires strong branding, which may be specific to the manufacturer or the dealer or both. Indeed this is a case in which branding and the resultant (short-run) monopoly power is the market means of *reducing* total costs to consumers. Abuse of this power is prevented by the effect upon the value of the brand reputation and hence future sales. Ford Motor Company has instituted an explicit program of gathering consumer complaints in order to monitor its authorized repair shops.

[29] Darby develops the idea of increasing monitoring difficulties as the essential basis of entrepreneurship as a limit to firm size. Michael R. Darby, *supra* note 21.

Diagnosis refers simply to exercise of expertise in the purchasing or monitoring of goods or services laden with credence or experience qualities. This expertise may be purchased in the market or be provided directly by the demander of the goods or services. Which will be the case is determined by whether it is cheaper to judge the expertise of the diagnosis or carry out the monitoring function by oneself. For frequently purchased experience goods, a consumer might purchase the goods without attached information, but carry out his own search instead.[30] For infrequently purchased experience goods and credence goods, a consumer might well find it cheaper to hire expertise in purchasing instead of investing his own time in the process. It is a worthwhile convenience to be able to place greater credit in a salesman's promise that a particular product has desirable qualities. A firm which, generally, sells honest advice will over time build up a reputation among its clients which enables it to charge a higher price for a given good than competitors whose promises carry less weight. Building up a reputation or satisfied clientele is in the nature of a capital investment, because losses must be borne because of competition of shops with either similar prices and quality and better reputations or similar reputations but lower prices and quality. But once the reputation is established the firm can charge the higher prices consistent with its reputation for quality diagnosis.[31] In general, the quality of this diagnosis will be less than if it could be judged without error by the consumer, but that is simply not a viable alternative because of the real costs of gathering information. If there are information costs in charging for advice about the quality of television sets, say, then we have an instance of the classic case for retail price maintenance as developed by Telser.[32]

Firms specializing in selling relatively accurate quality evaluation will be more heavily patronized by high income groups for which time spent gathering information is more costly. To the extent that the level of quality demanded increases with income, these stores will tend to sell both high quality information and high quality goods, but the two types of quality are conceptually quite different.

To the extent that a positive amount of fraud exists, additional economic resources will be diverted to the credence good industry which represents a misallocation of resources relative to the costless information case. Long term equilibrium without fraud implies that no firm is making net profits. The

[30] See Phillip Nelson, *supra* note 3.

[31] This provides an explanation of the information costs which Demsetz has suggested are neglected in the standard model of monopolistic competition. Harold Demsetz, The Inconsistencies in Monopolistic Competition: A Reply, 80 J. Pol. Econ. 592 (1972).

[32] See Lester G. Telser, Why Should Manufacturers Want Fair Trade?, 3 J. Law & Econ. 86 (1960).

discovery of fraud would increase profits and draw resources into the industry until equilibrium is reestablished. This causes less efficient production in which the value of the marginal social product is less than the marginal cost of the resources involved so long as real resources are used in commission of the fraud.

A central difficulty with the definition of fraud is the crucial importance of intent. A particular incorrect service prescription may equally well be due to fraudulent intent or simple incompetence. The harm done to the consumer is not affected by whether a mistake is due to fraud or ignorance. Therefore, the question arises whether there is a meaningful distinction between fraud and incompetence, or whether this is simply a case of fancy terminology. The distinction between the two concepts derives from the social (and private) costs of reducing fraud and of reducing incompetence. In the more general context the distinction is between the costs of reducing purposeful misrepresentation and of increasing the quality of information on the part of the seller. To increase the level of competence or quality of information involves investment of real resources both in labor and goods. On the other hand, a reduction in fraud or misrepresentation requires no additional resources but only a decision to stop. A universal adherence to honesty would thus involve a clear social gain, while an increase in competence yields social gains only to the extent that the gains are greater than the increase in resource costs.

IV. Governmental Intervention

A question of some current interest is whether government intervention in the markets for credence goods can improve the allocation of resources.[33] In order to improve the allocation, the benefits from improved monitoring must exceed any additional costs. Since the market would force the institution of such beneficial measures if they are privately available, any governmental intervention must be justified on the basis of reduced transactions costs or the government's greater ability to enforce contracts by jailing.[34]

So far as the demand for a shift of the unwritten contract involved in any sale is concerned, except as altered by a written contract, from no guarantee to guarantee of fitness, we are unable to draw any conclusions. If the parties can make contractual exceptions and sell on an "as is" or limited

[33] The reader will find an illuminating discussion of many of the same and related issues in Milton Friedman, Capitalism and Freedom ch. 9, "Occupational Licensure" (1962).

[34] See Cotton M. Lindsay, Impurities in the Theory of Public Expenditure, (Univ. of Calif., L.A., Dep't of Econ., Disc. Paper No. 20, 1972) for an analysis of this point.

guarantee basis, the only difference between the two states is which sales must bear the cost of written contracts.[35]

There are however proposals that the government set and enforce standards of quality for various credence goods. In essence, this is a proposal that the government provide a substitute for consumer monitoring either of credence goods or of market monitors of credence goods or of both. The basic argument may be summarized as follows: The evaluation of the quality of credence goods produced at a certain repair shop, say, has value to all of its potential customers but is produced under conditions of a large fixed cost of production of information and a relatively fixed cost of its dissemination. Once an evaluation is produced, it is argued, no one can charge more than the cost of dissemination for a report, and the costs of contracting to pay the fixed costs of production of information are very large relative to the benefits. If the costs of contracting are lowered by use of the government, then it may be optimal for the government to produce evaluations.

One technique of government information production is to rate the output of firms on the basis of an appropriate sample. This sort of rating could also be produced in the market, but requires consumer monitoring of the rating firm in turn. If the government produces ratings, the consumer must instead monitor the quality of the government ratings. In view of the relatively lower impact on any individual's wealth of a fall in the value of the government rating bureau's reputation, one would expect bribery for a better rating to be more widespread than would be the case for private rating firms. Thus failure of incentives suggests that the government is likely to have no advantage over private rating firms retained by the producers of credence qualities.

An alternative frequently suggested is to set quality standards and impose fines for failure to comply. This is rather similar to increasing the goodwill value of a firm's customers by adding a term in the penalty and the probability of receiving the penalty to equation (1). For appropriate levels of penalty and probability, it will be in the firm's interest to provide services of any given quality. If only one standard is set, then no variation in quality is allowed to reflect differences in costs of production or identification of different qualities. Thus any such program should allow producers to choose the degree to which they will guarantee their diagnosis, say, by choice and posting of quality class A, B, or C, or the like. The firm would then be subject to a penalty if it were to be found to fail to meet the corresponding

[35] We suspect, however, that the demand for implicit guarantees is really a demand for higher quality credence goods as discussed immediately *infra*.

standards in an appropriate sampling. This would still allow fraud if a firm could diagnose at quality level A, but chooses to do so only at the level of C which it posts, but the higher price received by class A shops will make this contrary to self interest. But this same effect could be had in the market since a firm can offer a reward ("Triple your money back!") to anyone finding the firm failing to meet appropriate levels of quality under the same ground rules.[36] A form of governmental intervention which combines both methods is for the government to prescribe standards for self-rated classes A, B, and C which are then enforced by private individuals filing remunerative malpractice suits if the claimed standards are not met. This "bounty hunter" system effectively turns expert buyers into monitors for the less expert customers.

There seem to be two arguments which might justify these forms of governmental intervention. The first is that it is cheaper to communicate the fact that one is subject to uniform government standards than individual rules, but this is a fairly weak argument since uniform private standards are established wherever the cost is worth the trouble. The second argument is of more interest. If a reward is offered to anyone discovering the shop violating uniform private standard A, the reward and probability of loss must be such that their expectation just covers the cost of an investigation. If instead governmental enforcement is utilized, then a smaller probability of loss and a higher fine could be established.[37] As a result, less resources could be devoted to production of costly evaluations of the quality of services provided. This follows from the fact that production of information about any given shop by costly tests is a natural monopoly.

In this type of enforcement scheme the probability of conviction for any offence is the probability of being investigated times the probability of being convicted if investigated. Most crimes involve the sequence of offence, report, investigation, possible conviction. This scheme—as well as such crimes as prostitution and gambling—involves the sequence of investigation, possible discovery of offence, and conviction. Here the investigation involves the question "Does X commit this crime?" instead of "Who has committed this crime?" But how are investigations to be allocated? Since the investigators can designate the potential investigatees, one likely rule-of-thumb is to

[36] For publicly owned firms, this guarantee is implicit in the possibility of a "consumer advocate" selling the firm's stock short before announcing discovery of fraud. As a result private provision of fraud detection could be greater than the social optimum if the private returns from the short sale are greater than the social returns in resource savings. This point is discussed in Jack Hirshleifer, The Private and Social Value of Information and the Reward to Inventive Activity, 61 Amer. Econ. Rev. 561 (1971).

[37] This is a direct application of the model in Gary S. Becker, Crime and Punishment: An Economic Approach, 76 J. Pol. Econ. 169 (1968).

investigate those individuals who refuse to pay not to be investigated.[38] The payments can, of course, be in money or kind, direct or indirect. Thus the appropriate fine will be lower because the probability of receiving it is greater for the relevant group actually subject to investigation.[39] If the salary of a building inspector, vice squad member, or the like is not lower than for comparable positions by the value of the bribes, then competition among potential holders of these positions will transfer the value of the bribe to the person or organization allocating the positions.[40] There is some evidence that such positions in many cities are considered rewards for political services.

But to return to the effect of this sort of regulation on fraud, or more generally, the quality of services provided, we see two cases according to whether or not a shop risks investigations. If it chooses to be investigated, then the marginal conditions apply as supposed in the scheme's justification, at least to those customers that the firm believes might be policemen, but these firms might as well have offered a reward to everyone since they are intensively investigated in either case. The second group bears a marginal cost of fraud—as opposed to effectively a lump sum licensing fee—only so far as the enforcers are able to price discriminate in bribes charged according to the amount of fraud practiced, which is to say according to the monitoring ability of the enforcers.

Therefore it is not at all clear that proposed schemes of governmental intervention can improve on market methods of monitoring the provision of credence goods, and this is primarily due to the fact that governmental operatives respond to much the same forces as are present in the market.

V. Summary

This paper has discussed the existence of credence qualities of goods which are costly to evaluate. As a result monitoring becomes valuable and "perfect competition" fails as a viable form of market organization. In part, the monitoring of credence qualities is carried out within a firm which, by integration, sells output requiring less monitoring. Also some firms specialize in the evaluation of the output of other firms, and this evaluation is a key function of a sales organization. When goods bearing credence qualities are sold in the market, branding, and the client relationship are tools used in monitoring the qualities provided. At the current state of

[38] Press reports indicate the use of this rule-of-thumb, for example, by vice squad policemen in allocating their resources in the investigation of prostitution and gambling.

[39] To the extent that the value of bribes are received in part by political organizations (see *infra*), the level of the fine might be influenced by derived demand conditions.

[40] To the extent that person or organization can price discriminate among potential entrants by monitoring.

knowledge, no strong case can be made for governmental intervention even in markets where deliberate deception is a regular practice. This is because governmental evaluators will be subject to much the same costs and temptations as are present for private evaluators.

APPENDIX

This appendix develops a model to explain the process underlying the consumer's decision to accept or decline a service offer and to return to or avoid the firm in his future dealings.

Let $\Phi(S)$ be the total cost to the consumer of purchasing a diagnosis and also S units of repair services. The cost of diagnosis alone if the consumer refuses any repairs is therefore $\Phi(O)$ and this amount is assumed to be fixed in advance.

Denote the subjective probability density function of a particular individual by $p\{\Phi(S)\}$. Upon detecting a decline in the flow of services from a durable good, the individual takes it to a service shop. At this point there is certain amount of uncertainty with respect to the repair costs, accordingly the expected utility of the individual based upon his prior information concerning repair costs, is given by,

$$EU(W) = \int_0^R U(W - \Phi(S))p\{\Phi(S)\}dS \tag{7}$$

where $U(W)$ denotes the maximum utility at a given wealth level W, and R is the level of service corresponding to the replacement cost of the durable good.

If the repair shop prescribes services in the amount $\bar{S}$, the individual can either accept the offer and pay $\Phi(\bar{S})$ or else reject it, pay the diagnosis fee of $\Phi(O)$, and try another shop. In utility terms, the individual can obtain a utility of $U(W - \Phi(\bar{S}))$ with certainty by accepting the offer $\Phi(\bar{S})$, or else $EU(W - \Phi(O))$ by turning down the offer and taking his business elsewhere.[41]

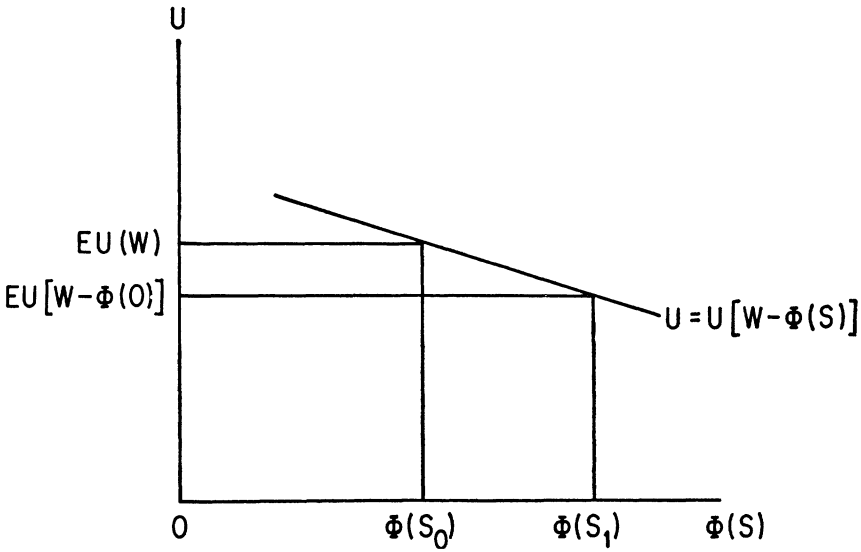
Taking the present and future repairs into account it is possible to derive a decision rule: (a) If $U(W - \Phi(\bar{S})) < EU(W - \Phi(O))$, then the consumer believes that he is better off taking this and all future business elsewhere since he suspects that he is being defrauded. (b) If $EU(W - \Phi(O)) < U(W - \Phi(\bar{S})) < EU(W)$ then the consumer believes that he is being defrauded, but having to pay the fixed costs $\Phi(O)$, he is better off accepting this repair, but he will take all future business elsewhere.[42] (c) If $U(W - \Phi(\bar{S})) > EU(W)$, then the consumer finds the quoted price acceptable, and will continue to patronize the firm.

These decision rules can be illustrated by reference to Figure II. The regions 0 to $\Phi(S_0)$, $\Phi(S_0)$ to $\Phi(S_1)$, and $\Phi(S_1)$ up correspond to decision rules (c), (b),

[41] In fact, the quoted price for service should affect the probability density function $P\{\phi(S)\}$, as there is some chance that a large quoted price is correct and not fraudulent. In the main, it will be sufficient to think of the p.d.f. as conditional on the quotation without further complication of the notation.

[42] In the present model a single detection of fraud is sufficient to prevent the customer from returning. A somewhat richer model is outlined *infra,* however.

FIGURE II



and (a) respectively. Note that whether $U(E[W - \Phi(S)])$ is greater or less than $EU(W)$ depends on whether the individual is a risk avoider or risk preferrer in the relevant range and hence the curvature of the $U = U[W - \Phi(S)]$ line. Thus each individual will have his own decision regions with respect to a quoted $\Phi(S)$ determined by his subjective probability distribution, his preferences, and his wealth. The repair shop entrepreneur must form a subjective conception of what these regions are; these conceptions are formalized as $F(S)$ and $H(S)$ in Section II.

A somewhat richer model of individual behavior is obtained if, at the cost allowing the firm some degree of monopoly power, it is assumed that a loss results to the individual if he decides to use another firm. In this case, it is possible for a customer to refuse an offer at one time and yet return in the future, adding yet a fourth decision region. It would also be possible to explicitly consider reduction in good will value due to lowered regions for detection of fraud on future purchases from higher prescribed services in the present. As neither of these qualifications leads to any significantly different implications, we retain the simpler model for the current paper.