The Economics of Breakdowns, Checkups, and Cures

Curtis R. Taylor
Texas A&M University

A market in which the owner of a durable good, X, contracts with an expert for diagnostic and treatment services is studied. Good X may be in one of three states: "health," "disease," or "failure." Only experts can determine whether X is healthy or diseased and perform treatment. The owner cannot tell whether recommended treatment is really needed. This creates an information-based demand for health insurance by risk-neutral consumers. Imperfections in the market for spot insurance may give rise to free diagnostic checks, strategic procrastination, and long-term health maintenance agreements.

I. Introduction

This paper investigates a long-standing dilemma in the economics of information. Specifically, how should an individual view the diagnosis and recommendations of an expert who has a vested interest in providing treatment? Consumers confront this predicament in many important markets, including medical and dental care, automobile and equipment repair, household pest control, and personal financial planning, among others. The services supplied in these markets are credence goods because a consumer typically never discovers the veracity of a diagnosis or even whether the treatment authorized was

I am grateful to Ray Battalio, Andreas Hornstein, Nancy Lutz, Bill Neilson, Jamie Thompson, Steve Wiggins, and participants in several seminars. I am especially indebted to Donald Deere for his excellent comments, to Canice Prendergast for his many insights, and to Sam Kortum for his continuous time assistance.

1 See, e.g., the classic article by Darby and Karni (1973) and the experimental work of Plott and Wilde (1982).
really performed.\textsuperscript{2} The effects of contracting for these expert services in a competitive market setting are studied below, and some rather intriguing results are obtained.

The papers most closely related to this are Demski and Sappington (1987) and Wolinsky (1993).\textsuperscript{3} Demski and Sappington also consider contracting with an expert, but they focus on the problem of inducing the expert to acquire costly expertise. Specifically, they consider the case in which a manager must be given incentives to gather information prior to undertaking a project. This incentive problem does not arise here because it is assumed below that a diagnostic check is necessary prior to any type of treatment. In the context of this paper, Demski and Sappington consider the case in which it is possible to perform "blind" treatment. For example, a service technician might recommend replacing an entire subsystem rather than exert effort to discover whether replacing a smaller component would suffice. This problem is important in some situations, but it has limited applicability in the setting considered here, since a diagnostic check is necessary prior to most types of treatment.

Wolinsky's recent paper is more closely related to this one. He also considers markets for informed expert services, although his analysis differs from that presented below in several key respects. First, Wolinsky does not consider the insurance solution at the heart of this study. Instead, he emphasizes the roles of consumer search and reputation in disciplining experts. In order to focus on these features, it is necessary to assume that experts price diagnostic services at cost and that they cannot commit to treatment prices before performing diagnostic checks. The contracts considered here are somewhat more general because both of these assumptions are relaxed. Also, while Wolinsky focuses primarily on issues of market organization, this paper concentrates on individual incentives, especially concerning dynamic behavior. Because of the difference in emphasis, most of the findings in this paper and Wolinsky's are complementary.

In Section II, a model of stochastic breakdowns is presented. A durable good \( X \) can be in one of three states: health, disease, or failure. It is possible to "cure" \( X \) when it is diseased through treatment, but determining whether \( X \) is healthy or diseased requires a costly diagnostic check. The first-best policy for maintaining \( X \) con-

\textsuperscript{2} Nelson (1970) first made the distinction between a search good for which quality is evident prior to purchase and an experience good for which quality is known only after consumption. The notion of a credence good for which quality may never be known is due to Darby and Karni (1973).

\textsuperscript{3} Other related papers are Milgrom and Roberts (1986), Milgrom (1988), Dana and Spier (1993), and Dybvig and Lutz (1993). See also the recent paper on this topic by Emons (1994).
sists in choosing a checkup frequency and a service plan that balances the cost of maintenance against the cost of treatment.

The incentive problems associated with hiring an expert to perform checkups and treatment are detailed in Section III. The main problem for the owner of X is that he never receives a signal that discriminates between the situations in which X was actually treated for disease and was never diseased in the first place. This means that an expert who discovers X to be healthy has an incentive to submit a false diagnosis of disease and to recommend unnecessary treatment. The expert may also have incentives to shirk on performance of diagnostic checks and treatment, but these incentives—unlike the incentive to recommend unnecessary treatment—can be overcome with incentive compatible service warranties or legal liability.

Section IV studies short-term contracts in which the strategies of the owner and the expert firms in the industry do not depend on the history of prior actions or outcomes. In this setting, a firm never diagnoses X as healthy in equilibrium. While this is somewhat extreme, it captures the spirit of the incentive problem under study and agrees with considerable anecdotal evidence. The fact that firms always recommend treatment does not, however, automatically create inefficiency or even information rents. If firms post diagnostic and treatment prices ex ante (before learning the condition of X), then they bid away all expected rents and essentially offer fair insurance against the presence of disease. Unfortunately, this insurance solution breaks down because ex ante pricing provides no incentive for the owner to perform the efficient level of maintenance.

One way to induce the owner to do efficient maintenance is the common practice of ex post contracting. Each firm still sets its diagnostic price before learning the condition of X, but it commits to a treatment bid only after performing an inspection that reveals the owner’s level of care and whether X is diseased. This provides incentives for the owner to do maintenance but gives the firm performing the checkup private information and ex post information rents. These rents are dissipated ex ante (at least partially) because diagnostic checks are provided free of charge in equilibrium. The owner reduces any residual rents by delaying checkups. Procrastination raises the probability that X actually needs treatment and thereby

---

4 See, e.g., Patterson (1992). Unnecessary car repairs were recommended by Sears mechanics in 90 percent of the test cases run by state officials throughout California.

5 The demand for insurance by risk-neutral consumers also arises in models of warranting product quality, as in Cooper and Ross (1985). The present paper identifies a different information-based demand for insurance, however. Specifically, consumers demand insurance from experts ex ante to counter the informational advantage experts enjoy ex post.
lowers the value of the firm’s private information. Of course, it also forfeits the benefits of early detection.

Multperiod contracts are considered in Section V. When the owner and a firm interact repeatedly, it is sometimes possible to support an equilibrium in which the firm provides fair insurance against disease and the owner does efficient maintenance. These long-term agreements naturally interpreted as health maintenance or extended service plans. Even when efficient long-term contracts are not supportable in equilibrium, it may be possible to support health maintenance agreements that are more efficient than short-term contracts by implementing a checkup frequency that is higher than first-best. A brief conclusion is provided in Section VI, and the Appendix contains several results. Some straightforward proofs have been abbreviated or omitted to save space.

II. The Model and the First-Best Policy

Let $X$ be a durable good subject to random failure. For instance, $X$ might be the physical health of an individual, the brakes on an automobile, or the financial status of a household. Good $X$ begins in a state of health 0. The amount of time spent in this state is governed by an exponential distribution with hazard rate $\lambda$. When $X$ leaves state 0, it enters the “disease” state, 1. The amount of time spent in this state is also exponentially distributed with hazard rate $\lambda$. Finally, when $X$ transits state 1, it enters a failure or death state, 2, where it remains forever.

Good $X$ can be treated for disease at cost $C$. Treatment immediately returns $X$ from state 1 to state 0. The problem is that a diagnostic check costing $D$ is necessary in order to ascertain whether $X$ is healthy or diseased. It is assumed that a diagnostic check is required before treatment can be performed. In other words, it is not possible to forgo the payment of $D$ and conduct “blind” treatment. Of course, $C$ is paid only if the diagnostic check reveals $X$ to be diseased.

If $X$ fails, its owner suffers a terminal cost $J$, which includes any decrease in his utility from the loss of services provided by $X$. For

\footnote{Grossman (1972) provides the first formal attempt to model the physical health of an individual as a capital good.}

\footnote{Here, “time” is a measure of usage that may be distinct from chronological age. For instance, if $X$ is an automobile, then “time” corresponds to mileage.}

\footnote{Distinct hazard rates $\lambda_1 \neq \lambda_2$ can be accommodated, but the algebra is substantially more complicated, and the results are unchanged.}

\footnote{The “two-horse-shay” stochastic process governing the health of $X$ may be interpreted as sequential component failure. Components fail according to a Poisson process, with two failures corresponding to a system crash. See Gale and Rosenthal (1994) for a model with a similar information structure but a very different application.}
instance, \( J \) could be the costs incurred when a tooth is pulled, a car engine is replaced, or an individual declares personal bankruptcy. Finally, as long as \( X \) survives, the owner bears a flow cost of \( K \) per unit of time, which represents the effort he exerts in maintaining \( X \). For instance, if \( X \) is physical health, \( K \) represents the cost per unit of time of following a proper diet and exercise regimen.

It is assumed that the owner is risk-neutral and that he discounts at rate \( r \). His objective is to minimize the expected present value of all maintenance, diagnostic, treatment, and terminal costs. He accomplishes this by selecting the length of time between checkups (the checkup interval) \( T \in [0, \infty) \) and a service plan \( (K, C) \in \{(K_L, C_H), (K_H, C_L)\} \), where \( K_L < K_H \) and \( C_L < C_H \). In other words, the owner may choose to exert either a low or a high level of effort in maintaining \( X \). If he opts for low maintenance, then \( X \) will be more expensive to treat when it becomes diseased.

Before the owner's objective function can be expressed and the cost-minimizing policy characterized, a few definitions are needed. Let \( F(S) \) be the probability that \( X \) is diseased, given that it has not failed, where \( S \) is the amount of time since \( X \) was last known to be healthy, that is, since its last checkup or since the beginning of its life. Some straightforward computation reveals \( F(S) = \lambda S / (1 + \lambda S) \). Let the probability that \( X \) fails by time \( S \) be denoted by

\[
G(S) = 1 - (1 + \lambda S) e^{-\lambda S}.
\]

Also, denote the hazard rate of failure at \( S \) by \( H(S) = \lambda F(S) \). That is, the instantaneous failure rate of \( X \) equals \( \lambda \), the hazard rate of moving from the disease to the death state, times \( F(S) \), the probability of being in the disease state. Finally, Let \( J(T, K, C) \) denote the expected present value of all future costs when the owner adheres to a checkup interval of length \( T \) and a service plan \( (K, C) \). Then \( J(T, K, C) \) is defined by the following recursive equation:

\[
J(T, K, C) = \int_0^T \left\{ \left[ 1 - \delta(S) \right] \frac{K}{r} + \delta(S) \bar{J} \right\} dG(S)
+ \left\{ \left[ 1 - \delta(T) \right] \frac{K}{r} + \delta(T) [D + F(T) C + J(T, K, C)] \right\}
\times \left[ 1 - G(T) \right],
\]

where \( \delta(S) = e^{-rs} \) is the discount factor associated with a time interval of length \( S \). The integral in (1) accounts for the fact that \( X \) may fail before the checkup date \( T \) is reached. Specifically, \( [1 - \delta(S)](K/r) \) is the present value of the maintenance cost annuity that ends if \( X \) fails at time \( S < T \), and \( \delta(S) \bar{J} \) is the present value of the terminal cost.
The expression in braces following the integral accounts for the fact that the checkup date is reached with probability \(1 - G(T)\). First, the maintenance cost annuity runs all the way until \(T\). Next, the expression \(\delta(T)[D + F(T)C + J(T, K, C)]\) is the expected present value of a checkup. It consists of the diagnostic cost \(D\), the expected treatment cost \(F(T)C\), and the cost function \(J(T, K, C)\), which indicates that the whole process “starts over” immediately following the checkup whether or not treatment was needed. If we rewrite (1), the owner’s optimization problem can be expressed as

\[
\min_{T,K,C} J(T, K, C) = \frac{K}{r} + \int_0^T \delta(S)[\bar{J} - (K/r)]dG(S) + \delta(T)[1 - G(T)]D + F(T)C \\
1 - \delta(T)[1 - G(T)]
\]

**Proposition 1.** Suppose that the owner adheres to the service plan \((K, C) \in [(K_L, C_H), (K_H, C_L)]\). If

\[
\frac{\lambda r}{(\lambda + r)^2} \left( \bar{J} - \frac{K}{r} \right) > D + C,
\]

then the unique cost-minimizing checkup interval, \(T^*\), is defined implicitly by

\[
H(T^*)[\bar{J} - [D + F(T^*)C + J(T^*, K, C)]] + F'(T^*)C + K = r[D + F(T^*)C + J(T^*, K, C)];
\]

if (2) is violated, then checkups should never be performed.

The value of a checkup rises over time because it becomes ever more likely that \(X\) is diseased. The left side of (3) is the marginal cost of delaying a checkup by \(dT\). Specifically, \(H(T)\) is the instantaneous failure rate of \(X\), and \(\bar{J} - [D + F(T)C + J(T, K, C)]\) is the net cost if a failure occurs during the delay. The term \(F'(T)C\) accounts for the increased likelihood that treatment will be necessary after the checkup is delayed, and \(K\) is the cost of maintenance during the delay. The right side of (3) is the marginal benefit of delaying a checkup by \(dT\). It is just the time value of money \(r\) times the expected cost of a checkup, \(D + F(T)C + J(T, K, C)\).

Expression (2) is also easy to interpret. It can be rewritten as

\[
\left( \frac{\lambda}{\lambda + r} \right) \bar{J} + \left[ 1 - \left( \frac{\lambda}{\lambda + r} \right) \right] \frac{K}{r} \\
> \left( \frac{\lambda}{\lambda + r} \right)^2 \bar{J} + \left[ 1 - \left( \frac{\lambda}{\lambda + r} \right)^2 \right] \frac{K}{r} + D + C.
\]
Suppose that it is known that $X$ is diseased. Then the left side of $(2')$ gives the discounted expected cost of not obtaining treatment, and the right side gives the discounted expected cost of obtaining treatment now and never again. As the time since $X$ was last known to be healthy, $S$, approaches infinity, the probability that $X$ is diseased conditional on its survival $F(S)$ goes to one. Expression $(2')$ says, therefore, that treatment is cost effective “at infinity.” Moreover, since the inequality is strict, there must be some finite $T^*$ at which checkups become cost effective.

Since this paper is concerned with procurement of treatment under asymmetric information, it is assumed below that $(2)$ holds, so that $(3)$ defines the first-best checkup interval. It is also assumed that the first-best policy calls for high maintenance and low treatment costs. That is, $J(T^*, K_H, C_L) < J(T^*, K_L, C_H)$, where $T^*$ is the optimal checkup interval under the service plan $(K_H, C_L)$ and $T^*$ is optimal under $(K_L, C_H)$.

III. The Agency Environment and the Contracting Game

Now, suppose that it is not possible for the owner to conduct checkups or treatment himself. Instead, he must hire an expert (a firm) to perform these tasks on his behalf. Firms are risk-neutral expected profit maximizers and also discount at rate $r$. It is assumed that the owner cannot observe whether a firm performs a diagnostic check or treatment. Moreover, the outcome of a diagnostic check is “soft” information. In other words, a firm has no objective means for substantiating its claim of whether $X$ is diseased. This might occur because either the owner does not possess the expertise needed to evaluate the relevant evidence or it might be possible for a firm to fabricate misleading evidence. It is also assumed that a diagnostic check reveals to the firm the owner’s choice of service plan, $(K, C)$, but that this information is too costly to verify to a court and, hence, cannot be contracted on directly.

These informational assumptions give rise to four potential incentive problems on the part of a firm. (I) A firm might shirk on performance of a diagnostic check and declare $X$ to be healthy. (II) A firm might shirk on performance of a diagnostic check and declare $X$ to be diseased, in which case it will also shirk if authorized to perform treatment.10 (III) A firm might faithfully perform a diagnostic check, learn that $X$ is diseased, and then shirk if authorized to perform

10 Recall that a diagnostic check is assumed to be necessary before treatment can be performed.
treatment. (IV) A firm might faithfully perform a diagnostic check, learn that X is healthy, and declare it to be diseased so that it might be authorized (and paid) to perform unnecessary treatment.

The first three of these incentive problems are situations of standard moral hazard. Since the owner and the firms are risk-neutral, it is possible to eliminate these three potential problems with conventional incentive compatible warranties. For instance, a firm might sell a warranty to the owner that commits the firm to make an indemnity payment of $V$ if $X$ fails before some specified time $\hat{T}$. It is straightforward to show that $V$ and $\hat{T}$ can be calibrated to eliminate incentive problem II and that this is sufficient for the elimination of incentive problems I and III as well. Moreover, as long as the warranty is priced at its expected value, it is individually rational for a firm to sell it and the owner to buy it.

The focus of this paper is incentive problem IV, which cannot be eliminated with a warranty. To see why, suppose that a firm declares $X$ to be diseased. Also, suppose that it backs this claim by offering the following warranty contract. If the owner does not obtain treatment and $X$ does not fail before some specified time $\hat{T}$, then the firm will make a payment of $W$ to the owner. There are two problems with this scheme. First, imagine that $W$ and $\hat{T}$ are calibrated so that the firm makes an honest diagnosis if it expects the owner to buy the warranty. Then, if it declares $X$ to be diseased, the owner's best response is to authorize treatment, and he will not buy the warranty. However, if the owner does not buy the warranty, then the firm's best response is to declare $X$ to be diseased regardless of its true condition. So there exists no equilibrium in which firms are always honest, and the owner always obeys their recommendations. Even if it is possible to approximate this outcome in a mixed-strategy equilibrium, there is another important problem. The owner can profit at the expense of the original firm by cheating on the terms of the warranty and secretly obtaining treatment from another firm. Unless this can be prevented, firms will be unwilling to warrant their diagnoses of disease. Throughout the remainder of the paper, therefore, it is assumed that incentive problems I–III are eliminated with warranties but problem IV is not. An important consequence of eliminating problems I–III is that a firm has no incentive except to be honest whenever it gives $X$ a clean bill of health (whether or not it has per-

11 Dybvig and Lutz (1993) call this common type of contract a “block” warranty because it provides constant coverage for an initial block of time. They show that such warranties often possess excellent incentive properties in situations in which the rate of product failure depends on both the quality choice of the manufacturer and the care taken by the consumer.

12 Pitchik and Schotter (1987) analyze a mixed-strategy equilibrium in a similar game.
formed treatment). So it is assumed below that treatment is faithfully performed whenever it is needed and that a diagnosis of health is genuine.

Suppose that the industry consists of \( N \) identical competitive firms with which the owner may do business. Let \( S \) denote the amount of time since \( X \) was last known to be healthy, and suppose for convenience that the firms as well as the owner observe \( S \).\(^{13}\) Given the discussion of the previous paragraph, \( S = 0 \) can occur in one of three ways: it is the beginning of \( X \)'s life, \( X \) has just been diagnosed as healthy, or a firm has just claimed to have treated \( X \) for disease.

The timing is as follows. At \( S = 0 \), firms simultaneously choose their diagnostic and treatment price schedules \( \Pi^i(\cdot) \) and \( P^i(\cdot) \), \( i = 1, \ldots, N \). Then, at some time of his choosing, \( S = T \leq \infty \), the owner pays \( \Pi^i(\cdot) \) and obtains a checkup from one of the firms. The chosen firm submits a diagnosis \( \hat{\theta}^i \in \{0, 1\} \). If it declares \( X \) to be healthy, \( \hat{\theta}^i = 0 \), then the current round of play ends and \( S = 0 \) again.\(^{14}\) If it diagnoses \( X \) as diseased, \( \hat{\theta}^i = 1 \), then the firm offers to perform treatment at price \( P^i(\cdot) \).

After receiving a diagnosis of disease, the owner may either authorize the current firm to perform treatment at the specified price or seek the services of another firm. For simplicity, it is assumed that there is no search or transportation cost associated with obtaining a second opinion and that no time elapses between consecutive checkups. The current round of play ends when the owner receives a diagnosis of health for the first time, when he authorizes treatment, or when \( X \) dies. Payoffs in each round are assigned in the natural way with the important observation that a firm authorized to perform unnecessary treatment bears no treatment cost. So, its payoff is \( \Pi^i(\cdot) + P^i(\cdot) - D \).

### IV. Short-Term Contracts

This section considers situations in which the behavior of the owner and firms may depend only on actions and outcomes in the current round of play, not on the history of previous actions or outcomes. Begin by considering ex ante contracts \([\Pi^i(S), P^i(S)]\) in which all firms \( i = 1, \ldots, N \) post diagnostic and treatment prices as functions of \( S \). This type of contract is called “ex ante” because firms commit to treatment prices before learning whether \( X \) is diseased. While ex

\(^{13}\) If firms cannot observe \( S \), then the analysis below is complicated by adverse selection. Specifically, the firms must anticipate the checkup interval that will be chosen by the owner in equilibrium and price checkups and treatment accordingly.

\(^{14}\) It is important to note that \( S \) is the state variable measuring time and \( \theta \) is the physical state of \( X \); thus \( S = 0 \) implies \( \theta = 0 \), but \( \theta = 0 \) does not imply \( S = 0 \).
ante contracts are very natural types of agreements, they are seldom observed in practice. Car mechanics and other technicians almost always perform diagnostic checks before making binding treatment bids. Medical and dental services are most often provided under a multiperiod scheme involving either an ongoing relationship between the provider and patient or some more explicit type of arrangement. An important shortcoming of ex ante contracts that helps to explain their infrequent use is identified below.

First, observe that a firm never possesses an incentive to diagnose X as healthy under an ex ante contract.\textsuperscript{15} So whenever the owner obtains a checkup, he knows that he will receive a diagnosis of disease, $\hat{\theta}^i = 1$, and that it will be false with probability $1 - F(T)$. This would not present a problem if the owner could commit himself to following the first-best service plan $(K_H, C_L)$.

Because the owner knows that he will always receive a diagnosis of disease, he places no value on checkups alone and will obtain one only when he intends to authorize treatment. Moreover, because of the warranties discussed in the previous section, the owner knows that if he authorizes treatment, X will be returned to him in state 0, whether or not it was in this state already. So he views the total price $\Pi'(S) + P'(S)$ of each firm as a price for spot insurance against the presence of disease. Since the firms are identical, an equilibrium in ex ante contracts requires the total price of each firm to be actuarially fair.\textsuperscript{16} If the owner could commit to the first-best service plan, this would yield

$$\Pi'(S) + P'(S) = D + F(S)C_L, \quad i = 1, \ldots, N.$$  

It is easy to verify that if this holds and the owner follows the first-best service plan, then he will choose the first-best checkup interval, $T^*_L$. Unfortunately, ex ante contracts will not implement the first-best policy because they give the owner an incentive to do only minimal maintenance.

PROPOSITION 2. Under an equilibrium with ex ante contracts, the owner implements the inefficient service plan $(K_L, C_H)$ and the corresponding checkup interval $T^*_H$.

Under an ex ante contract, the owner pays the same price, $P'(T)$, for treatment regardless of which service plan he implements. So his

\textsuperscript{15} The contract space could be enriched to allow a firm to be rewarded for submitting a diagnosis of health, but since the reward would have to be at least as big as the rents the firm could get by recommending unnecessary treatment, paying an incentive compatible reward is economically equivalent to operating under ex ante contracts.

\textsuperscript{16} Strictly speaking, Bertrand competition requires only two of the firms to price at expected cost in equilibrium. The other $N - 2$ firms may post higher total prices, although they will get no business. Without loss of generality, only symmetric pricing equilibria are considered here.
only incentive is to choose the plan with the lower maintenance cost, \((K_L, C_H)\). Of course, in equilibrium the firms will anticipate this moral hazard on the part of the owner and will price accordingly:

\[
\Pi^i(S) + P^i(S) = D + F(S)C_H.
\]

The fact that ex ante contracts provide poor incentives for owners to perform maintenance is an important reason behind the infrequent use of this type of pricing mechanism.

The simplest way of eliminating the moral hazard and concomitant inefficiency under ex ante contracts is to let the treatment price depend directly on the service plan, \(P^i(S, K, C)\). Recall, however, the assumption made in Section III that contracting directly on the service plan is infeasible because of high verification costs. Another method for inducing the owner to select the efficient service plan is to use ex post contracts. Specifically, when \(S = 0\), all firms \(i = 1, \ldots, N\) post three price schedules \([\Pi^i(S), P^i_L(S), P^i_H(S)]\). Then at some time \(T \leq \infty\) of his choosing, the owner pays a firm \(\Pi^i(T)\) to perform a checkup. If the firm diagnoses \(X\) as diseased (as it always will in equilibrium), it offers to perform treatment at either the low price \(P^i_L(T)\) or the high price \(P^i_H(T)\), or it refuses to treat \(X\) at either price. If the firm makes a treatment bid, the owner may either authorize treatment at the specified price or take his business elsewhere.

This type of arrangement, in which a firm commits to its treatment price only after performing a diagnostic check, is often used in industries such as car and equipment repair, pest control, and personal financial planning. Ex post contracts do provide incentives for the owner to do efficient maintenance, but the fact that a firm obtains private information about whether \(X\) is diseased before committing to its treatment price can sometimes lead to another problem. Specifically, the owner will attempt to reduce the payment of any information rents by adopting an inefficiently long checkup interval. Before this is demonstrated, it is necessary to derive the equilibrium price schedules under ex post contracts.

**Proposition 3.** The equilibrium prices are given by

\[
\Pi^i(S) = 0, \quad i = 1, \ldots, N,
\]

and

\[
P^i_M(S) = \begin{cases} 
C_M, & S \leq T^0_M \\
D + F(S)C_M, & S \geq T^0_M,
\end{cases}
\]

\(M \in \{L, H\}, i = 1, \ldots, N\), where

\[
T^0_M = \max \left\{0, \frac{C_M - D}{\lambda D} \right\}, \quad M \in \{L, H\}.
\]
Also, in equilibrium a firm that performs a checkup bids $P^*_L(T)$ on treatment if the owner followed $(K_H, C_L)$ and $P^*_H(T)$ if he followed $(K_L, C_H)$.

Two aspects of this result deserve comment. First, the conclusion that firms offer free diagnostic services in equilibrium with ex post contracts is attractive. In fact, this occurs widely in industries that use ex post pricing arrangements: auto and equipment repair, pest control, personal financial planning, and so forth.

The second notable conclusion of proposition 3 is more subtle. First, it is easy to establish

$$T \approx T^*_M \Leftrightarrow C_M \approx D + F(T)C_M, \quad M \in \{L, H\}.$$  

This, along with (4), indicates that if the owner adopts a checkup interval $T \approx T^*_M$, he will get actuarially fair insurance against disease. In other words, a firm authorized to perform treatment makes an expected profit of zero in this case. If $T < T^*_M$, however, a firm authorized to perform treatment makes a positive expected information rent of

$$R(T) = C_M - [D + F(T)C_M], \quad M \in \{L, H\}.$$  

Under an ex post contract, a firm cannot commit itself to treat X below cost because it learns whether treatment is really needed before deciding what and whether to bid. This creates a positive expected rent since X needs the recommended treatment only with probability $F(T)$. The rent persists in spite of the competitive environment because it cannot be dissipated without charging negative diagnostic prices $\Pi'(T) = -R(T)$. These prices are infeasible since they would create a money pump for the owner, who could get rich by procuring checkups and never authorizing treatment. These observations lead to the following result.

**Proposition 4.** If $T^*_L \geq T^*_L$, then the owner implements the first-best policy $(T^*_L, K_H, C_L)$. If $T^*_L < T^*_L$ and

$$J(T^*_L, K_H, C_L) < J(T^*_H, K_L, C_H),$$

the owner implements the first-best service plan but waits too long between checkups.

Let $\bar{J}(T, K, C)$ denote the expected present value of costs to the owner under ex post contracts. When $T \geq T^*_M$, a firm authorized to treat X receives an expected profit of zero. In other words, the cost functions $\bar{J}(T, K^*_M', C_M)$ and $J(T, K^*_M', C_M)$ coincide to the right of $T^*_M$. So when $T^*_L \geq T^*_L$, the first-best policy $(T^*_L, K_H, C_L)$ minimizes $\bar{J}(T, K, C)$, and ex post contracts are efficient (see fig. 1).
When, on the other hand, $T_L^* < T_L^0$, the first-best checkup interval gives positive expected information rents to a firm. The owner can (and will) reduce these rents by waiting until $T_L^{**} > T_L^*$ to get a checkup. By delaying, the owner raises the probability that X needs the recommended treatment from $F(T_L^*)$ to $F(T_L^{**})$, and this dilutes the value of the firm's private information. Casual empiricism suggests that procrastination by owners in obtaining checkups is very common. Proposition 4 indicates that procrastination by informationally disadvantaged owners may be privately rational.

Of course, waiting until $T_L^{**}$ is socially inefficient because of the increased likelihood of failure over the interval $(T_L^*, T_L^{**})$. The deadweight loss due to procrastination is the difference $J(T_L^{**}, K_H, C_L) - J(T_L^*, K_H, C_L)$ (see fig. 2).

Although ex post contracts sometimes lead to inefficiently long checkup intervals (especially when the treatment cost is large relative to the diagnostic cost), inequality (6) ensures that the efficiency loss from ex post contracts never exceeds that from ex ante contracts. This seems reasonable in light of the profusion of ex post pricing arrangements. The next section explores long-term contracts that are also widely used and tend to be efficient precisely when ex post contracts perform worst.

V. Long-Term Contracts

The last section focused on equilibria of the infinite horizon contracting game corresponding to repeatedly playing equilibria of the
one-shot stage game. This section considers a type of equilibrium that depends on the opportunity for repeated interaction between the owner and a firm. These multiperiod arrangements are naturally interpreted as health maintenance or extended service plans.

The reason for the failure of ex ante contracts in the last section was that there existed no way to penalize the owner for implementing the inefficient service plan. In a long-term relationship, however, a firm may be able to induce the owner to comply with the first-best plan by threatening to sever the relationship if he does not. Suppose that the firms in the industry offer two types of contracts, short-term ex post contracts and a long-run version of ex ante contracts. Under the long-term contract, the owner agrees to obtain checkups and treatment from the same firm in each round unless and until either he or the firm exercises an option to terminate the relationship. The long-term contract also specifies the length $T$ of the checkup interval the owner must employ. The total price of a "visit" in any given round is stipulated to be

$$\Pi^i(T) + P^i(T) = D + F(T)C_L, \quad i = 1, \ldots, N,$$

and the firm is allowed to terminate the contract only if it warrants $X$ as healthy. This means that if the owner implements the inefficient

---

17 Fudenberg, Holmstrom, and Milgrom (1990) have shown that the achievable outcomes are the same in short- and long-term complete contracts. The contracts in this paper are incomplete since the service plan cannot be contracted on directly.

18 This idea was first formalized by Klein and Leffler (1981).
service plan and X is diseased, the firm must perform treatment before it can terminate the contract.\textsuperscript{19} Finally, it is assumed that if a firm terminates a long-term contract with the owner, then no other firm offers him a new one.\textsuperscript{20} So the owner must procure checkups and treatment using short-term ex post contracts, following severance of a long-term relationship. Of course, this disciplines the owner only to the extent that he is harmed by short-term contracts.

**Proposition 5.** A long-term contract of the type described above can implement the first-best policy \((T_L^*, K_H, C_L)\) if and only if

\[
\delta(T_L^*)[1 - G(T_L^*)][\hat{J}(T_L^{**}, K_H, C_L) - J(T_L^*, K_H, C_L)] = 0. 
\]

The left side of (7) is the cost to the owner of having a long-term contract terminated because he has performed low maintenance. Specifically, X survives until the checkup date \(T_L^*\) with probability \(1 - G(T_L^*)\), at which time the owner’s malfeasance is detected and he is forced thereafter to procure checkups and treatment with short-term contracts. The expression \(\hat{J}(T_L^{**}, K_H, C_L) - J(T_L^*, K_H, C_L)\) is the net cost to the owner of being forced to use short-term contracts rather than continuing with the long-term agreement. The right side of (7) is the benefit from following the inefficient service plan under a long-term contract. The integral accounts for the difference in the present value of the maintenance cost annuities if X fails before \(T_L^*\), and the expression following the integral gives this difference if X survives to \(T_L^*\) times the probability, \(1 - G(T_L^*)\), of survival.

Proposition 5 complements the findings of the last section. If ex post contracts are efficient, then efficient long-term equilibrium contracts of this type do not exist because the threat of termination carries no bite:

\[
\hat{J}(T_L^{**}, K_H, C_L) - J(T_L^*, K_H, C_L) = 0. 
\]

On the other hand, if short-term contracts are inefficient, then the threat of reverting to them may be sufficient to support an efficient long-term agreement. Also, if the difference in maintenance costs \(K_H - K_L\) is small, then the threat of reverting to short-term contracts

\textsuperscript{19} If the firm could terminate the contract without treating X, then it would have an incentive to do so if \(C_L > D + F(T)C_L\), since the price of treatment would not cover its cost, even under the efficient service plan.

\textsuperscript{20} If the owner implements the efficient service plan, the expected profit to a firm under a long-term contract is zero. So if the owner is terminated, no other firm has a strict incentive to offer him a new long-term contract.
may support an efficient long-term agreement, even when short-term contracts are not very inefficient.

Proposition 6. Suppose that (7) is violated. Then it still may be possible to support long-term contracts that are more efficient than the sequence of equilibrium short-term contracts by implementing a checkup interval shorter than $T^*_L$. That is, even if (7) is violated, there may exist $T < T^*_L$ such that

$$
\delta(T)[1 - G(T)]\left[\int (T^*_L, K_H, C_L) - J(T, K_H, C_L)\right] > 0
$$

Suppose that equilibrium short-term contracts are not efficient. Then the first-order effects from shortening the checkup interval in a long-term contract are to raise the cost of termination to the owner and to lower the benefit from following the inefficient service plan. Intuitively, a shorter checkup interval means that the firm monitors the owner more closely, since the amount of time he can get away with following the inefficient service plan is diminished. Of course, as $T$ is reduced from $T^*_L$, $J(T, K_H, C_L)$ rises. So there is a limit to how short the checkup interval can be made before the threat of reverting to short-term contracts loses its punch.

The long-term contracts characterized here are just actuarially fair insurance policies. A contract under which the provider of treatment insures the health of $X$ over an indefinite horizon is nothing more than a health maintenance agreement or extended service plan. The owner demands this insurance not because he is risk-averse, but because he avoids the information rents associated with short-term contracting.

Arrangements in the spirit of these long-term contracts are very common in markets for medical and dental services and are sometimes observed in the auto and large-appliance repair and pest control markets as well. These contracts may be explicit as in the case of health maintenance organizations or dental plans, or they may be based more on good faith and reputation, as is usually the case with family practice physicians and neighborhood garages.\(^{21}\) An important point, buried in propositions 5 and 6, is that both $\lambda$ and $r$ must be sufficiently small to support this type of agreement. In situations in

\(^{21}\) Hubbard (1994) presents some very interesting preliminary empirical findings regarding California's vehicle inspection and maintenance program. For instance, he reports that the failure rate on "under-hood inspections" performed by independent stations (neighborhood garages) is nearly half the failure rate on inspections performed by branded stations, 9.9 percent vs. 18.6 percent. These rates may be compared with the 15 percent failure rate on vehicles checked at state-run roadside inspection sites.
which mobility and failure rates are high, it will not be possible to sustain multiperiod equilibrium contracts.

Finally, the equilibria suggested in propositions 5 and 6 are only one type of potential equilibria in history-contingent agreements. The fact that this sort of contract is often observed justifies it as a candidate for study, but probably not as the definitive solution. For instance, there may be multiperiod equilibrium arrangements in which a firm gives honest diagnoses and the owner uses a weighted average of the firm’s prior recommendations to decide whether to sever the relationship. While it seems likely that such an arrangement could be supported as an equilibrium, note that it must be Pareto inferior to the contract of proposition 5, because the owner will sever the relationship following a run of sufficiently bad luck.

VI. Conclusion

This paper provides theoretical underpinnings for several prominent features observed in markets for diagnoses and treatment. For instance, the heavy reliance on ex post contracts, the prevalence of free diagnostic services, and consumer procrastination in obtaining checkups are all consistent with the model developed above. Another contribution of this paper is that it identifies an informational demand for insurance. This is most clear in the long-term health maintenance agreements discussed in Section V, where it is shown that provision of health insurance by a firm eliminates the incentive for it to recommend unnecessary treatment.

Regarding the efficiency of markets for diagnoses and treatment, the findings of this paper are cautiously optimistic. First of all, the common practice of ex post contracting may be fully efficient. Moreover, in cases in which ex post contracts are fairly inefficient, proposition 5 indicates that efficient long-term health maintenance agreements are likely to be supportable in equilibrium. Even when efficient long-term agreements cannot be supported, however, proposition 6 shows that shortening the interval between checkups may allow the support of long-term agreements that outperform the sequence of short-term contracts.

There are, of course, many methods and institutions for dealing with the informational problems associated with markets for diagnostic and treatment services. Several prominent alternatives not discussed in this paper are the use of second opinions, the specialization of firms into either diagnostic or treatment services, and the reliance on professional associations that police their members. Each of these alternatives is important in specific settings, and they were excluded from the analysis above purely for expositional reasons. Nevertheless,
the extent to which the market-based contracts studied here can alone alleviate the predicament facing consumers of expert services is encouraging.

Appendix

Proof of Proposition 1

The owner's objective function in the first-best world can be written as

$$J(T, K, C) = \frac{K}{r} + \frac{\beta(T)[\bar{J} - (K/r)] + \gamma(T)[D + F(T)C]}{1 - \gamma(T)}, \quad (A1)$$

where

$$\beta(T) = \int_0^T \delta(S) dG(S) = \frac{\lambda^2}{(\lambda + r)^2} \{1 - [(\lambda + r)T + 1]e^{-(\lambda + r)T}\}$$

and

$$\gamma(T) = \delta(T)[1 - G(T)] = (1 + \lambda T)e^{-(\lambda + r)T}.$$ 

First it is shown that (3) defines the set of critical points of $J(T, K, C)$. Differentiate (A1) with respect to $T$ to get

$$J'(T, K, C) = \left(\left\{\beta'(T)\left(J - \frac{K}{r}\right) + \gamma'(T)[D + F(T)C] + \gamma(T)F'(T)C\right\}\right) \times [1 - \gamma(T)] + \gamma'(T)\left\{\beta(T)\left(J - \frac{K}{r}\right) + \gamma(T)[D + F(T)C]\right\} \quad \text{(A2)}$$

Then a critical point is any $T^*$ satisfying $J'(T^*, K, C) = 0$, which is equivalent to

$$\beta'(T^*)\left[J - \frac{K}{r}\right] + \gamma'(T^*)[D + F(T^*)C]$$

$$+ \gamma(T^*)F'(T^*)C = -\gamma'(T^*)\left[J(T^*, K, C) - \frac{K}{r}\right]$$

$$\Leftrightarrow \beta'(T^*)\left[J - \frac{K}{r}\right] - \left[r\gamma(T^*) + \beta'(T^*)\right][D + F(T^*)C]$$

$$+ \gamma(T^*)F'(T^*)C = \left[r\gamma(T^*) + \beta'(T^*)\right]\left[J(T^*, K, C) - \frac{K}{r}\right]$$

$$\Leftrightarrow \beta'(T^*)[\bar{J} - D - F(T^*)C - J(T^*, K, C)] + \gamma(T^*)F'(T^*)C$$

$$= r\gamma(T^*)[D + F(T^*)C + J(T^*, K, C)] - \gamma(T^*)K.$$ 

Dividing through by $\gamma(T^*)$ and noting $\beta'(T^*)/\gamma(T^*) = H(T^*)$ show that (3)
defines the critical points of $J(T, K, C)$. Next, it is shown that (2) is equivalent to the existence of a unique critical point. To see this, substitute for $\beta(T)$ and $\gamma(T)$ on the right side of (A2) to get

$$J'(T, K, C) = \frac{[(z - 1)e^z + 1][\beta(\infty)(\bar{r} - K) - \lambda C] - (\lambda z + r)e^zD}{(e^z - \lambda T - 1)^2},$$

where $\beta(\infty) = [\lambda/(\lambda + r)]^2$ and $z = (\lambda + r)T$. Define the functions

$$W(T) = \frac{(\lambda + r)Te^{(\lambda + r)T}}{[e^{(\lambda + r)T} - \lambda T - 1]^2}$$

and

$$V(T) = \left[1 - \frac{1}{(\lambda + r)T} + \frac{1}{(\lambda + r)Te^{(\lambda + r)T}}\right]$$

$$\times [\beta(\infty)(\bar{r} - K) - \lambda C] - \left[\lambda + \frac{r}{(\lambda + r)T}\right]D.$$

Then $T^*$ is a critical point iff it satisfies

$$J'(T^*, K, C) = V(T^*)W(T^*) = 0.$$

Now, $W(T)$ is positive and finite for all positive finite values of $T$. So critical points are defined implicitly by the condition $V(T^*) = 0$. Observe that

$$\lim_{T \to 0} V(T) = -\infty,$$

$$\lim_{T \to 0} V(T) = \beta(\infty)(\bar{r} - K) - \lambda(C + D).$$

Therefore, at least one local minimum of $J(T, K, C)$ exists iff

$$\beta(\infty)(\bar{r} - K) - \lambda(C + D) > 0 \iff \frac{\lambda r}{(\lambda + r)^2} \left(\bar{r} - \frac{K}{r}\right) > D + C,$$

where the right-hand side of the equivalence is (2). To see that (2) also implies uniqueness, note that it implies $\beta(\infty)(\bar{r} - K) - \lambda C > 0$. Also, the expression $e^z - z - 1$ is positive for $z > 0$, since it achieves its global minimum when $z = 0$. Together, these expressions yield the following string of implications:

$$(e^z - z - 1)[\beta(\infty)(\bar{r} - K) - \lambda C] > 0$$

$$\Rightarrow \left(\frac{1}{z^2} - \frac{1 + z}{z^2e^z}\right)[\beta(\infty)(\bar{r} - K) - \lambda C] > 0$$

$$\Rightarrow \left(\frac{1}{z^2} - \frac{1 + z}{z^2e^z}\right)[\beta(\infty)(\bar{r} - K) - \lambda C] + \frac{r}{z^2}D > 0$$

$$\Rightarrow (\lambda + r)\left(\frac{1}{[(\lambda + r)T]^2} - \frac{1 + (\lambda + r)T}{[(\lambda + r)T]^2e^{(\lambda + r)T}}\right)$$

$$\times [\beta(\infty)(\bar{r} - K) - \lambda C] + \frac{r}{[(\lambda + r)T]^2}D > 0.$$
The last line establishes that $V(T)$ is monotone increasing because the left-hand side is the expression for $V'(T)$. Q.E.D.

Proof of Proposition 3

The equilibrium concept is the perfect sequential equilibrium notion of Grossman and Perry (1986) with the added condition that no player employ a weakly dominated strategy. First, note that all firms always diagnose X as diseased in equilibrium under ex post contracts. Next, note that if (4) holds, then a firm performing a checkup makes the "correct" treatment bid in equilibrium. Specifically, it bids $P_i^L(T)$ when the owner follows $(K_H, C_L)$, since the owner would take his business elsewhere if the firm bid $P_i^H(T)$. Also, the firm bids $P_i^H(T)$ when the owner follows $(K_L, C_H)$, since it either loses money or signals to the owner that treatment is unnecessary by bidding $P_i^L(T)$.

Next, it must be shown that the price functions are correct. Observe that in equilibrium

$$P_i^M(S) \geq C_M, \quad M \in \{L, H\}, i = 1, \ldots, N. \quad (A3)$$

The reason is that if X is diseased, a firm would refuse to treat it rather than bid below cost, and the owner would not frequent such a shop. In fact, an owner with treatment cost $C_M$ will procure a checkup at time $S = T$ from the firm possessing the lowest total price $\Pi^i(T) + P_i^M(T)$. If the selected firm bids $P_i^M(T)$, then the owner authorizes treatment. If two or more firms are "tied" with the lowest total price, then the owner has a weakly dominant strategy to procure a checkup from the one with the lowest diagnostic price $\Pi(T)$. If two or more firms are tied with the lowest total price and the lowest diagnostic price, then the owner randomizes between them. If the owner receives a diagnosis (off the equilibrium path) of $\theta^i = 0$, then he buys a warranty against this claim and believes it. Otherwise, he believes that X is diseased with probability $F(T)$ for any behavior of the firms (on or off the equilibrium path). In the absence of any constraints, competition among the firms would force them to post actuarially fair total prices

$$\Pi^i(S) + P_i^M(S) = D + F(S)C_M. \quad (A4)$$

However, for $S < T_M^0$, this and (A3) imply $\Pi^i(S) < 0$, which is not feasible because it creates a money pump for the owner. For $S \geq T_M^0$, it is possible to satisfy (A3) and (A4) without negative diagnostic prices. Finally, since the

22 The restriction to strategies that are not weakly dominated ties down the equilibrium price functions for $S > T_M^0$. Without this condition, diagnostic prices would not necessarily be zero over this region, since any prices $[\Pi^i(S), P_i^L(S), P_i^H(S)]$ satisfying $P_i^M(S) \geq C_M$ and

$$\Pi^i(S) + P_i^M(S) = D + F(S)C_M$$

would be admissible in equilibrium. This would not change equilibrium behavior, however.
owner breaks ties in total prices by selecting the firm with the lowest diagnostic price, the firms will back-load their fees and charge zero for diagnostic services in equilibrium. Q.E.D.

**Proof of Proposition 4**

Let \( J(T, K_{M'}, C_M) \) denote the present expected value of costs to the owner if he employs a checkup interval of length \( T \) and the service plan \((K_{M'}, C_M)\) in an equilibrium with ex post contracts. Then

\[
J(T, K_{M'}, C_M) = \frac{K_{M'}}{r} + \frac{B(T)[\bar{J} - (K_{M'}/\gamma)] + \gamma(T)\bar{P}_{M'}(T)}{1 - \gamma(T)} \tag{A5}
\]

First, suppose \( T_L^* \geq T_L^0 \). Then substituting (4) into (A5) and invoking proposition 1 show that the owner implements the first-best policy, \((T_L^*, K_H, C_L)\).

Now, suppose \( T_L^* < T_L^0 \). Let \( T_M^* \) denote the cost-minimizing checkup interval if the owner implements \((K_{M'}, C_M)\). Now, substituting (4) into (A5) establishes

\[
J(T_M^*, K_H, C_L) \leq J(T_L^0, K_H, C_L),
\]

and condition (6) establishes that the owner implements the first-best service plan. Finally, it is straightforward to prove that \( T_L^* < T_L^0 \) implies \( T_L^* < T_M^* \). This is done by showing that \( J'(T, K, C) < J'(T, K, C) \), so that \( J(T, K, C) \) achieves its minimum to the right of \( J(T, K, C) \). Q.E.D.

**References**


Patterson, Gregory A. "Sears’ Brannen Accepts Blame for Auto Flap." Wall St. J. (June 23, 1992), sec. B.