Reprinted from The Bell Journal of Economics
Vol. 12, No. 2, Autumn 1981

Good news and bad news: representation theorems and applications

Paul R. Milgrom

Good news and bad news: representation theorems and applications

Paul R. Milgrom*

This is an article about modeling methods in information economics. A notion of "favorableness" of news is introduced, characterized, and applied to four simple models. In the equilibria of these models, (1) the arrival of good news about a firm's prospects always causes its share price to rise, (2) more favorable evidence about an agent's effort leads the principal to pay a larger bonus, (3) buyers expect that any product information withheld by a salesman is unfavorable to his product, and (4) bidders figure that low bids by their competitors signal a low value for the object being sold.

1. Introduction

■ Information economics is the study of situations in which different economic agents have access to different information. Many kinds of institutions and patterns of behavior have been treated as attempts to cope with such informational asymmetries. For example, Spence (1973) has treated higher education as an attempt by talented workers to signal their talents to employers. Akerlof (1976) has offered a similar analysis of the "rat race," in which employees work faster than the socially optimal pace to distinguish themselves from less talented coworkers. Milgrom and Roberts (1979) offer a signaling analysis of the phenomenon of limit pricing, in which an established firm sets its price below the monopoly price in an attempt to discourage potential competitors. In each of these signaling models, the analysis is driven by a monotonicity property: more talented workers buy more education (Spence) or work faster (Akerlof) than their less talented counterparts, and lower cost firms set lower prices.

Monotonicity also plays a key role in models of adverse selection. For example, in the insurance market models of Rothschild and Stiglitz (1976), C. Wilson (1977), and Pauly (1974) in which each individual knows his probability of suffering a loss but the insurers do not, the individuals with the greatest likelihood of loss buy the most comprehensive insurance coverage. Similarly, in Akerlof's (1970) famous "lemons" model, higher prices in the used car market result in a higher average quality of the cars available, since owners of good cars will simply keep them if the prevailing prices are too low.

---

* Northwestern University.

I am pleased to acknowledge the many useful suggestions of Sanford Grossman, Bengt Holmstrom, Alvin Klevorick, Roger Myerson, Mark Satterthwaite, Robert Weber, and Ward Whitt. This research was partially supported by the Center for Advanced Studies in Managerial Economics, by NSF grant SES-8001932, and by ONR grant N00014-79-C-0685.

Additional examples of the role of monotonicity can be found in the literatures on search, advertising, and bidding. In bidding, for example, the typical analysis proceeds on the basis of the intuition that a buyer's bid should be an increasing function of his true reservation price. This price, of course, is known only to the buyer. For example, see Vickrey (1961, 1962) and Ortega-Reichert (1968).

In view of the role of monotonicity in so much of information economics, it is surprising that studies of rational expectations equilibria and of the problem of moral hazard make no use of any such property. One might guess, for example, that in a rational expectations model the arrival of good news about a firm's prospects would cause the price of its stock to rise. Such results have, unfortunately, been out of reach because no device has been available for modeling "good news." The purpose of this article is to introduce such a device.

In the formal model treated in Section 2, there is a single, unknown, real-valued parameter θ which is of interest to a decisionmaker. The variable θ might represent "quality" or "intrinsic value" in a rational expectations or adverse selection model. The decisionmaker observes an informative signal x. Depending on the nature of θ, an appropriate signal might be an array of experimental data, a financial or geological report, a road map, a satellite photograph, or a television news show. In the absence of extra assumptions, the form that a signal takes is theoretically irrelevant to its ability to convey information.

Thinking of θ as "effort" or "ability" or "quality," I shall say that observation x is *more favorable than* observation y if for every nondegenerate prior distribution on θ the posterior corresponding to x dominates that corresponding to y in the sense of strict first-order stochastic dominance. In Section 2, I characterize the "more favorable than" relation and develop some related ideas.

The usefulness of the ideas is illustrated by a series of four applications. The first of these is a simple security market model in which the announcement of good news about a security's future returns causes its price to rise.

The second application is to a model in which a principal must design a fee schedule for his agent in an uncertain venture. The principal is unable to observe directly the effort expended by the agent, but he can observe the random profit of the venture which is influenced by the agent's effort. The agent is assumed to be risk averse and to have a reservation level of utility, reflecting his other opportunities. The principal's problem is to design a fee schedule (in which the agent's fee may depend on the profit of the venture) that trades off the necessity of providing the agent with appropriate work incentives against the desire to provide some risk sharing. It has been something of a puzzle in the earlier analyses of this model that the resulting fee schedule may not be increasing in the venture's profits. It turns out that nonmonotonicity in the fee schedule can arise only when higher profits can be evidence of lower effort on the part of the agent. When higher profits are evidence of greater effort, the optimum fee schedule is steeper than any efficient risk-sharing fee schedule.

For the third application, I introduce *games of persuasion*, in which an interested party (such as a salesman or a regulated firm) tries to influence a decisionmaker (such as a consumer or a regulator) by selectively providing data relevant to the decision. In one version of the model, at equilibrium, the interested party reports the information that is most favorable to his case, while withholding less favorable information. If communication between the parties is costless and if the decisionmaker can detect any withholding of information, then, at equilibrium, the decisionmaker adopts a strategy of extreme skepticism: he assumes that information is withheld only if it is very unfavorable. In response, the interested party's best strategy is one of full disclosure.

In the final application, a sealed-bid auction is studied. It is shown that winning the auction is "bad news," that is, the winner's estimate of the value of his prize tends to be too high. Winning with a relatively low bid is especially bad news, since it implies that no competitor has tendered even a moderate bid.

2. Representation theorems

■ Let Θ be a subset of $\mathbb{R}$, representing possible values of the random parameter $\tilde{\theta}$. The set of possible signals about $\tilde{\theta}$ is denoted by X which, for expositional simplicity,[1] is taken to be a subset of $\mathbb{R}^m$. Let $f(x|\theta)$ denote the conditional density (or probability mass) function on X when $\tilde{\theta}$ takes the particular value θ. With this set-up, let us say that a signal x is *more favorable than* another signal y if for every nondegenerate[2] prior distribution G for θ, the posterior distribution $G(\cdot|x)$ dominates the posterior distribution $G(\cdot|y)$ in the sense of strict first-order stochastic dominance.

Recall that a distribution G_1 is said to dominate G_2 in the sense of first-order stochastic dominance if for every increasing function U,[3]

$$\int U(\theta)dG_1(\theta) > \int U(\theta)dG_2(\theta).$$

Intuitively, G_1 dominates G_2 if every decisionmaker whose utility is increasing in θ prefers gamble G_1 to gamble G_2. It is well known that G_1 dominates G_2 in this sense if and only if for every θ, $G_1(\theta) \leq G_2(\theta)$, with strict inequality for some value of θ.[4]

To investigate the "more favorable than" relation, let G be a prior distribution for $\tilde{\theta}$ that assigns probabilities $g(\theta)$ and $g(\bar{\theta})$ to two possible values θ and $\bar{\theta}$ of $\tilde{\theta}$. By Bayes' theorem,

$$\frac{g(\bar{\theta}|x)}{g(\theta|x)} = \frac{g(\bar{\theta})\,f(x|\bar{\theta})}{g(\theta)\,f(x|\theta)}, \tag{1}$$

[1] I also assume for simplicity that the densities are positive everywhere. The propositions in this section are true exactly as stated for general measurable spaces and general density functions.

[2] A distribution is *degenerate* if it assigns probability one to a single point γ, and *nondegenerate* otherwise.

[3] More precisely, the strict inequality must hold for all increasing functions U such that both $\int UdG_1$ and $\int UdG_2$ are finite.

[4] One could also define "more favorable than" by using second-order stochastic dominance. A distribution G_1 dominates G_2 in this sense if for every increasing concave function U,

$$\int UdG_1 < \int UdG_2.$$

When G has two-point support, these concepts of dominance are identical; so (2) is necessary to conclude that x is more favorable than y in either sense. As Proposition 1 shows, it is also sufficient.

and a similar expression describes the posterior odds given y. In particular, if $\theta < \bar{\theta}$, if $g(\theta) = g(\bar{\theta}) = ½$, and if x is more favorable than y, then it follows that

$$\frac{f(x \mid \bar{\theta})}{f(x \mid \theta)} > \frac{f(y \mid \bar{\theta})}{f(y \mid \theta)}. \tag{2}$$

Proposition 1. x is more favorable than y if and only if for every $\bar{\theta} > \theta$,

$$f(x \mid \bar{\theta}) f(y \mid \theta) - f(x \mid \theta) f(y \mid \bar{\theta}) > 0. \tag{2a}$$

Proof: Equation (2a) generalizes (2) by allowing for the possibility that $f(y \mid \bar{\theta}) = 0$, a possibility that I shall henceforth ignore. The derivation of (2) constitutes the proof that it is necessary.

For sufficiency, fix some nondegenerate G and choose θ^* for which $0 < G(\theta^*) < 1$. For $\theta \leq \theta^*$, it follows from (2) that

$$\frac{\int_{\bar{\theta}>\theta^*} f(x \mid \bar{\theta}) dG(\bar{\theta})}{f(x \mid \theta)} > \frac{\int_{\bar{\theta}>\theta^*} f(y \mid \bar{\theta}) dG(\bar{\theta})}{f(y \mid \theta)}$$

or equivalently,

$$\frac{f(x \mid \theta)}{\int_{\bar{\theta}>\theta^*} f(x \mid \bar{\theta}) dG(\bar{\theta})} < \frac{f(y \mid \theta)}{\int_{\bar{\theta}>\theta^*} f(y \mid \bar{\theta}) dG(\bar{\theta})}. \tag{3}$$

Integrating (3) over θ for $\theta \leq \theta^*$ yields

$$\frac{\int_{\theta\leq\theta^*} f(x \mid \theta) dG(\theta)}{\int_{\bar{\theta}>\theta^*} f(x \mid \bar{\theta}) dG(\bar{\theta})} < \frac{\int_{\theta\leq\theta^*} f(y \mid \theta) dG(\theta)}{\int_{\bar{\theta}>\theta^*} f(y \mid \bar{\theta}) dG(\bar{\theta})}. \tag{4}$$

The last expression is equivalent to

$$G(\theta^* \mid x)/[1 - G(\theta^* \mid x)] < G(\theta^* \mid y)/[1 - G(\theta^* \mid y)],$$

which implies that $G(\theta^* \mid x) < G(\theta^* \mid y)$. *Q.E.D.*

Definition. Let $X \subset \mathbb{R}$. The densities $\{f(\cdot \mid \theta)\}$ have the *strict* monotone likelihood ratio property (strict MLRP) if for every $x > y$ and $\bar{\theta} > \theta$, (2a) holds. If the strict inequality in (2a) is changed to a weak inequality, then the adjective "strict" is dropped from the definition.

The monotone likelihood ratio property takes its name from the fact that the likelihood ratio $f(x \mid \theta)/f(x \mid \bar{\theta})$ is monotone in x, increasing if $\theta > \bar{\theta}$ and decreasing otherwise. This property plays a major role in statistical theory, as described in most basic textbooks on the subject. Among the families of densities and probability mass functions with this property are the normal (with mean θ), the exponential (with mean θ), the Poisson (with mean θ), the uniform (on $[0, \theta]$), the chi-square (with noncentrality parameter θ), and many others. With the definition of strict MLRP and Proposition 1, we have:

Proposition 2. The family of densities $\{f(\cdot \mid \theta)\}$ has the strict MLRP iff $x > y$ implies that x is more favorable than y.

Two signals x and y are called *equivalent* if for every θ and $\tilde{\theta}$,

$$f(x|\tilde{\theta})f(y|\theta) - f(x|\theta)f(y|\tilde{\theta}) = 0. \tag{5}$$

In view of (1), it is apparent that starting from any prior, equivalent signals lead to identical posterior beliefs about $\tilde{\theta}$. Two signals are called *comparable* if they are equivalent or if one is more favorable than the other.

Families of densities with the strict MLRP have the convenient property that any two signals are comparable. The next proposition establishes that any information system with this comparability property can be modeled using a real-valued variable with the MLRP.

Proposition 3. Let X be general and suppose that any two signals in X are comparable. Then there exists a function $H: X \to I\!R$ such that $H(\tilde{x})$ is a sufficient statistic for $\tilde{x}$ and such that the densities of $H(\tilde{x})$ have the strict MLRP.

Proof: Let $h: I\!R \to I\!R$ be any bounded increasing function and define

$$H(x) = \int h(\theta)dG(\theta|x), \tag{6}$$

where G is any nondegenerate prior for θ. Since signals are comparable, $H(x) > H(y)$ if and only if x is more favorable than y. Therefore, by Proposition 2, the densities of $H(\tilde{x})$ have the strict MLRP. Also, since $H(x) = H(y)$ iff x and y are equivalent, $H(\tilde{x})$ is a sufficient statistic. *Q.E.D.*

Occasionally, interesting situations arise in which there are natural notions of good news, neutral news, and bad news. A signal x is called *neutral* if for every prior distribution G, $G = G(\cdot|x)$. It follows immediately from (1) that a signal x is neutral if and only if for every θ and $\tilde{\theta}$, $f(x|\theta) = f(x|\tilde{\theta})$. A signal x is *good news* if it is more favorable than neutral news (i.e., if $f(x|\theta)$ is increasing in θ) and it is bad news if it is less favorable. For example, if $1 - \tilde{\theta}$ is the unknown failure rate of a certain system, then a period without failure is good news.

This section ends with a proposition that will be useful in the subsequent applications.

Proposition 4. Let $\tilde{x}$ be a random variable whose densities have the strict MLRP. For any two intervals $[a, b]$ and $[c, d]$ with $a \geq c$ and $b \geq d$, where at least one inequality is strict, the signal $\{\tilde{x} \in [a, b]\}$ is more favorable than $\{\tilde{x} \in [c, d]\}$.

Proof: Let U be any bounded increasing function and fix any nondegenerate prior distribution for $\tilde{\theta}$. Given any numbers α and β with $\alpha < \beta$, let $f_{\alpha,\beta}$ denote the conditional density of $\tilde{x}$, given $\tilde{x} \in [\alpha, \beta]$. The first step is to show that $\{\tilde{x} = \beta\}$ is more favorable than $\{\tilde{x} \in [\alpha, \beta]\}$, which, in turn, is more favorable than $\{\tilde{x} = \alpha\}$. Using Proposition 2,

$$E[U(\tilde{\theta})|\tilde{x} \in [\alpha, \beta]] = \int_{\alpha}^{\beta} E[U(\tilde{\theta})|\tilde{x} = x]f_{\alpha,\beta}(x)dx$$

$$< \int_{\alpha}^{\beta} E[U(\tilde{\theta})|\tilde{x} = \beta]f_{\alpha,\beta}(x)dx = E[U(\tilde{\theta})|\tilde{x} = \beta].$$

The inequality $E[U(\tilde{\theta})|\tilde{x} \in [\alpha, \beta]] > E[U(\tilde{\theta})|\tilde{x} = \alpha]$ has an analogous proof.

Next, if $a \geq d$, then $E[U(\tilde{\theta})|x \in [a, b]] > E[U(\tilde{\theta})|\tilde{x} = a] \geq E[U(\tilde{\theta})|\tilde{x} = d] > E[U(\tilde{\theta})|\tilde{x} \in [c, d]]$, and the proof is complete. If $a < d$, then

$$E[U(\tilde{\theta})|\tilde{x} \in [a, b]] = P\{\tilde{x} \in [a, d]|\tilde{x} \in [a, b]\}E[U(\tilde{\theta})|\tilde{x} \in [a, d]] + P\{\tilde{x} \in [d, b]|\tilde{x} \in [a, b]\}E[U(\tilde{\theta})|\tilde{x} \in [d, b]].$$

But since $E[U(\tilde{\theta})|\tilde{x} \in [d, b]] > E[U(\tilde{\theta})|\tilde{x} = d] > E[U(\tilde{\theta})|\tilde{x} \in [a, d]]$, we conclude that $E[U(\tilde{\theta})|\tilde{x} \in [a, b]] > E[U(\tilde{\theta})|\tilde{x} \in [a, d]]$ and similarly that $E[U(\tilde{\theta})|\tilde{x} \in [c, d]] < E[U(\tilde{\theta})|\tilde{x} \in [a, d]]$. These last two inequalities establish the result. *Q.E.D.*

3. Application: securities markets

■ The first example is a simple model of a securities market in which the public announcement of good news about the future returns on a security causes its price to rise.

Let there be two securities: a riskless security for which the return will be 1 and a risky security with the random return $\tilde{\theta}$. All investors are assumed to be identical with a concave, differentiable utility-of-wealth function U. Each investor is endowed with one unit each of the risky and riskless securities. Clearly, no trading takes place, so that setting the price of the riskless security at one, the price, p, of the risky security can be computed from the typical investor's first-order condition:

$$p = \frac{E[\tilde{\theta}U'(1 + \tilde{\theta})]}{E[U'(1 + \tilde{\theta})]}.$$

Let $g(\cdot)$ denote the density (or mass function) for $\tilde{\theta}$, and define another density $\bar{g}$ by

$$\bar{g}(\theta) = g(\theta)U'(1 + \theta)/E[U'(1 + \tilde{\theta})].$$

Letting $\bar{E}$ denote the expectations operator corresponding to the density $\bar{g}(\theta)$, we can express the price as:

$$p = \bar{E}[\tilde{\theta}].$$

Now suppose that some news x is publicly revealed. Then reasoning as before and applying Bayes' theorem lead to the following expressions for a new market-clearing price, $p(x)$:

$$p(x) = \frac{E[\tilde{\theta}U'(1 + \tilde{\theta})|x]}{E[U'(1 + \tilde{\theta})|x]} = \bar{E}[\tilde{\theta}|x].$$

Let x and y be signals with x more favorable than y. The definition of "more favorable" requires that for *any* prior distribution for $\tilde{\theta}$, including the prior $\bar{g}$, the posterior, given x, stochastically dominates the posterior, given y. It follows that $\bar{E}[\tilde{\theta}|x] > \bar{E}[\tilde{\theta}|y]$; more favorable news leads to higher prices. In particular, good news x causes the price to rise ($p(x) > p$) and bad news causes it to fall.

Expectation expressions of the form $p(x) = E[\tilde{\theta}|x]$ are abundant in financial market theory. In them, x usually represents the information available at some point in time (cf., Cox and Ross (1976) and Harrison and Kreps (1979)).

4. Application: moral hazard

■ In Holmström's (1979) treatment of the principal-agent problem, an agent expends effort θ to influence the profit of a venture. Let π denote the profit and let $\tilde{\alpha}$ denote the random state of nature. Realized profits depend on both θ and $\tilde{\alpha}$: $\tilde{\pi} = \pi(\tilde{\alpha}, \theta)$. It is assumed that effort always improved profits ($\partial\pi/\partial\theta > 0$), but there are diminishing returns to effort ($\partial^2\pi/\partial\theta^2 < 0$).[5] The agent dislikes expending effort; his payoff $U(x) - \theta$ is an increasing function of his wealth x and a decreasing function of effort. In addition, the agent is risk averse: $U'' < 0$. The principal has utility for wealth only. His payoff is denoted by $G(x)$, where $G' > 0$ and $G'' \leqq 0$. A *fee schedule or sharing rule* $s(\cdot)$ is a function that specifies the agent's compensation for each possible profit level of the venture. Notice that s depends *only* on π, because the variables θ and $\tilde{\alpha}$ cannot be observed by the principal.

It is the fact that θ and $\tilde{\alpha}$ are unobservable that leads to the moral hazard problem. If, for example, the principal were risk neutral (i.e., $G'' = 0$) and θ were observable, then any Pareto optimal sharing rule would involve the agent's receiving a fixed fee for undertaking a specified level of effort, and the principal would bear all risks (Spence and Zeckhauser, 1971). Since θ is assumed not to be observable, however, a contract based on a specified level of effort is not enforceable, and the agent must be given some incentive to expend effort.

In this setting, it might seem reasonable that the sharing rule should be increasing, since a rule with some decreasing segments is undesirable from a risk-sharing point of view and appears to reduce the agent's effort incentive. As the following example shows, this appearance is misleading. Let $P\{\tilde{\alpha} = 0\} = P\{\tilde{\alpha} = 1\} = .5$, let $\Theta = [0, .9]$, and let $\tilde{\pi} = \tilde{\alpha} + \theta$. Then θ can be perfectly inferred from any realization π. If the principal is risk neutral and the agent is risk averse, then the agent's compensation in the optimal contract will depend only on θ. Thus, the agent's share when $\pi = 1.0$ can quite sensibly be smaller than his share when $\pi = .1$.

A plausible model in which the sharing rule is increasing results if one assumes that π has the MLRP as information about θ. To formalize this, let $f(\pi|\theta)$ denote the conditional distribution of output, given effort. Assume that f is differentiable and let f_θ denote $\partial f/\partial\theta$. Holmström showed that the optimal sharing rule must satisfy the following relationship for some θ^*, b, and c ($c > 0$):

$$\frac{G'(\pi - s(\pi))}{U'(s(\pi))} = b + c\,\frac{f_\theta(\pi|\theta^*)}{f(\pi|\theta^*)}\,. \tag{7}$$

From the concavity of U and G, it is apparent that s is increasing in π if $f_\theta(\pi|\theta^*)/f(\pi|\theta^*)$ is increasing in π. This latter condition is a local characterization of the monotone likelihood ratio property.

Proposition 5. The family $\{f(\pi|\theta)\}$ has the MLRP if and only if for every θ^*, $f_\theta(\pi|\theta^*)/f(\pi|\theta^*)$ is increasing.

Proof. Notice that $f_\theta/f = \partial \ln f/\partial\theta$. It follows that for any θ' and θ'',

$$f(\pi|\theta')/f(\pi|\theta'') = \exp\left\{-\int_{\theta'}^{\theta''} [f_\theta(\pi|\theta)/f(\pi|\theta)]d\theta\right\}.$$

The conclusion follows easily. *Q.E.D.*

In this principal-agent model, the MLRP assumption captures the intuitive idea that greater profits are evidence of greater effort by the agent, so that the fee schedule should slope upwards to provide the correct incentives. Indeed,

[5] The assumption of diminishing returns to effort was absent from Holmström's analysis, but it is needed, as Grossman and Hart (1980) have shown.

one can deduce from (7) and the MLRP that s crosses any first-best sharing rule at most once and only from below. That reflects the intuition that the desire to provide incentives results in steeper fee schedules than would be desirable for pure risk-sharing.

5. Application: the persuasion game

■ The two previous sections display routine applications of the more-favorable-than relation to well-known models. This powerful modeling tool can also be used to render tractable a whole range of new problems.

The model considered in this section is a simple version of what I call a *persuasion game*, in which one or more interested parties provide information to a decisionmaker in an attempt to influence his decision. Persuasion games can be used to model regulatory decisions, courtroom battles, and sales encounters. The kinds of questions that these games help to answer are: How effectively does an adversary system provide useful information to decisionmakers? When should a buyer rely on a salesman, and when should he incur costs to gather his own information?

Let us consider a simple sales encounter in which a commodity of unknown quality $\tilde{\theta}$ is to be exchanged for money. If the buyer purchases q units of the commodity at price p, his payoff is $\tilde{\theta}F(q) - pq$. The salesman's payoff is his commission, which is some increasing function of q. It is assumed that F is bounded, increasing, concave, and differentiable, and that $F'(0) = +\infty$.

Let the salesman have N pieces of data about his product, represented by $\tilde{x} = (\tilde{x}_1, \ldots, \tilde{x}_N)$. The salesman may report or conceal any of these variables, but he cannot misreport them. Such a feature might arise if the information is verifiable by a product demonstration or if there are truth-in-advertising laws.

The sales encounter can be conveniently modeled as a game with incomplete information. In this game a *report* by the salesman is a closed nonempty subset of $\mathbb{R}^N$; the report S is to be interpreted as an assertion by the salesman that $\tilde{x} \in S$. A *reporting strategy* r is a function from $\mathbb{R}^N$ to the closed nonempty subsets of $\mathbb{R}^N$ with the property that $x \in r(x)$ for all $x \in \mathbb{R}^N$. The condition $x \in r(x)$ models the constraint that the salesman must report truthfully. The salesman's report can be very precise, as when $r(x) = \{x\}$, or it can be very vague, as when $r(x) = \mathbb{R}^N$, but it can never be false.

A *purchase decision* is a nonnegative real number q, representing the quantity purchased. A *purchasing strategy*, b, is a function from reports to purchase decisions. Thus, $b(S)$ specifies how much to buy when the salesman reports S. A pair (b, r) is a Nash equilibrium if holding r fixed, b is optimal for the buyer and, holding b fixed, r is optimal for the salesman.

Some Nash equilibria of the sales encounter game are unnatural. For example, at one equilibrium, the buyer resolves to ignore the salesman's report and the salesman makes only uninformative reports, i.e., $r(x) \equiv \mathbb{R}^N$ and $b(S) \equiv q^*$, where q^* maximizes $E[\tilde{\theta}F(q) - pq]$. It seems unreasonable, however, that the buyer would actually choose to ignore even very precise information, and that the salesman would expect such behavior. For the sales encounter game, a more sensible solution concept than the Nash equilibrium is the *sequential equilibrium* introduced by Kreps and Wilson (1980). At a sequential equilibrium, the buyer must always act in his own self-interest; he cannot resolve to ignore a report that is relevant to his decision. Every sequential equilibrium is Nash, but not every Nash equilibrium is sequential.

Consider how the buyer interprets the reports he receives. When the salesman makes a report S, the buyer can safely conclude that $\tilde{x} \in S$, but he may choose to draw a sharper conclusion. For example, if the salesman reports that his product "meets or exceeds" a certain standard, the buyer might infer that the product does not substantially exceed the standard. This idea can be formalized as follows. Given a report S, let $c(S)$ be a nonempty subset of S representing the conclusion or conjecture reached by the buyer. The interpretation is that if the seller reports S, the buyer will conclude that $\tilde{x} \in c(S)$.

For the sales encounter game, a sequential equilibrium is a triple (b, r, c) satisfying three conditions:

(i) For every possible report S, $b(S)$ solves $\max_q E[\tilde{\theta}F(q) - p \cdot q \,|\, \tilde{x} \in c(S)]$.
(ii) For every $x \in \mathbb{R}^N$, $r(x)$ solves $\max_S b(S)$, subject to $x \in S$.
(iii) For every S in the range of r, $c(S) = r^{-1}(S)$.

Condition (i) states that the buyer will maximize his expected payoff, given his conjectures, in response to any report the salesman makes. Condition (ii) is the usual best response condition for the salesman. Condition (iii) is a rational expectations condition. It asserts that the buyer's conjectures are consistent with the salesman's strategy, or, more informally, that the buyer takes the salesman's motives into account in considering the report.

For any Nash equilibrium (b, r), the triple (b, r, c) with $c = r^{-1}$ satisfies conditions (ii), (iii), and (i)′:

(i)′ For every report S in the range of r, $b(S)$ solves $\max_q E[\tilde{\theta}F(q) - pq \,|\, \tilde{x} \in c(S)]$.

The distinguishing feature of the sequential equilibrium is that, for *any* report the salesman may make, the buyer is obliged to listen to the salesman's report, form a conjecture consistent with that report, and base his purchase decision upon that conjecture.

One more definition is required for the statement of the next proposition. A reporting strategy r is called a strategy of *full disclosure* if r together with any optimal response (b, c) satisfies $b(r(x)) \equiv b(\{x\})$. Intuitively this condition means that r does not conceal any information relevant to the buyer's decision. In the present context, it is direct to show that the only information relevant to the buyer's decision is $E[\tilde{\theta}|x]$. Consequently, r is a strategy of full disclosure if $E[\tilde{\theta}|\tilde{x} = x] \equiv E[\tilde{\theta}|r(\tilde{x}) = r(x)]$.

Proposition 6.[6] At every sequential equilibrium of the sales encounter game, the salesman uses a strategy of full disclosure.

Proof: Let (b, r, c) be an equilibrium and let x be an arbitrary signal in $\mathbb{R}^N$. From condition (ii), it follows that $b(\{x\}) \leq b(r(x))$. From (i) and the fact that $c(\{x\}) = \{x\}$, the foregoing inequality can hold only if $E[\tilde{\theta}|\tilde{x} = x] \leq E[\tilde{\theta}|\tilde{x} \in c(r(x))]$. Using (iii), this becomes $E[\tilde{\theta}|\tilde{x} = x] \leq E[\tilde{\theta}|r(\tilde{x}) = r(x)]$. Since x was arbitrary, the inequality can be written $E[\tilde{\theta}|\tilde{x}] \leq E[\tilde{\theta}|r(\tilde{x})]$. (The expressions on the left- and right-hand sides of this inequality represent random variables whose values depend on the particular realization of $\tilde{x}$. The inequality between these two random variables was just shown to hold for all possible realizations x of $\tilde{x}$.) If the inequality were ever strict, we could conclude that $E[E[\tilde{\theta}|\tilde{x}]] < E[E[\tilde{\theta}|r(\tilde{x})]]$. But by a well-known identity of probability theory,

[6] The argument given here is essentially the same as the one given by Grossman (1980), though the equilibrium concept is slightly different.

$E[E[\tilde{\theta}|\tilde{x}]] = E[\tilde{\theta}] = E[E[\tilde{\theta}|r(\tilde{x})]]$. Hence, $E[\tilde{\theta}|\tilde{x}] \equiv E[\tilde{\theta}|r(\tilde{x})]$, so that r is a strategy of full disclosure. *Q.E.D.*

It can be shown that, at a sequential equilibrium, the buyer suspects that any information withheld is unfavorable to the product, i.e., the conjecture $c(S)$ minimizes $E[\tilde{\theta}|\tilde{x} \in c(S)]$ subject to $c(S) \subset S$.[7] When the buyer is so suspicious, the salesman's best strategy is one of full disclosure.

In the formulation given above, the idea that reports can be verified takes an extreme form. In effect, it is assumed that the buyer can verify both product information and statements like: "I have reported everything I know." In other words, the buyer can detect when the seller is concealing information. One promising approach to modeling the salesman's ability to conceal information is to let N be a random variable whose realization cannot be verified. That approach, however, is not explored here.

Another interesting way to generalize the sales encounter game is to allow for costly communications or for constraints on the player's abilities to transmit, receive, or process information. A particularly simple model with constraints on communication is studied below.

Specifically, consider a modification of the sales encounter game in which the buyer can assimilate only k observations, where $k < N$. To formalize that restriction, let the salesman's reports be limited to sets of the form $S = S_1 \times \cdots \times S_N$, where at most k of the S_j's can be different from $\mathbb{R}$.

For this model, it is useful to assume that $\tilde{x}_1, \ldots, \tilde{x}_N$ are (conditionally on $\tilde{\theta}$) independent and drawn from a common family of distributions $\{F(\cdot|\theta)\}$ with the strict monotone likelihood ratio property. Let $\mathbb{R}$ be defined to include $-\infty$ and let the salesman be constrained to report only *closed* sets S. With these restrictions, one can speak of the *least favorable* observation x_i in each S_i.

Proposition 7. The modified sales encounter game has a sequential equilibrium in which the salesman always reports the k most favorable observations.

Proof: Define $m_i = \min S_i$ and let $\bar{m}$ be the kth smallest element of $\{m_1, \ldots, m_N\}$. Let $M_i = \max\{m_i, \bar{m}\}$. For any S, let $c(S) = [m_1, M_1] \times \cdots \times [m_N, M_N]$.

This specification of $c(\cdot)$ can be stated less formally—but more clearly—as follows. For the k sets S_i that are different from $\mathbb{R}$, the buyer conjectures that $\tilde{x}_i = m_i = \min S_i$, i.e., he makes the least favorable conjecture consistent with S_i. For the other $N - k$ sets $S_j = \mathbb{R}$, he conjectures that $\tilde{x}_j \le \bar{m}$. If the salesman reports more than $N - k$ sets $S_j = \mathbb{R}$, then the buyer conjectures that $\tilde{x}_j = -\infty$ for each such j.

Define b by condition (i) and define r to be the strategy given in the proposition. We must show that (b, r, c) is an equilibrium, i.e., that (i)–(iii) hold. It follows directly from the specifications of c and r that $c = r^{-1}$ on the range of r. Thus, condition (iii) is satisfied. Condition (i) holds by definition.

Since $c(S)$ depends only on the minima m_i of the sets S_i, the report $[m_1, \infty) \times \cdots \times [m_N, \infty)$ leads to the same purchase decision as does the report S. Hence, the salesman's problem reduces to one of selecting an N-tuple $(m_1, \ldots, m_N)$, with at least $N - k$ components equal to $-\infty$, to maximize $b(S)$. This is equivalent to maximizing $E[\tilde{\theta}|\tilde{x} \in c(S)] = E[\tilde{\theta}|\tilde{x}_1 \in [m_1, M_1], \ldots, \tilde{x}_N \in [m_N, M_N]]$. By Proposition 4, the expectation is an increasing function of

[7] If for any closed set, S, this minimum does not exist, then no sequential equilibrium exists. One sufficient condition for existence of an equilibrium is that $f(x|\theta)$ be continuous and have compact support. The conditions used for Proposition 7 are also sufficient.

$(m_1, \ldots, m_N, M_1, \ldots, M_N)$. Then, since $\tilde{x}_1, \ldots, \tilde{x}_N$ are independent and identically distributed and since the M_i's are nondecreasing functions of $(m_i, \ldots, m_N)$, the expectation is a symmetric increasing function of the m_i's alone. Consequently, the specified reporting strategy is the one that maximizes the expectation $E[\tilde{\theta}|c(S)]$, and condition (ii) is satisfied. *Q.E.D.*

In both variations of the sales encounter game studied here, the buyer's attitude at equilibrium is one of extreme skepticism, and that feature makes the analysis tractable and intuitively comprehensible.

6. Application: auction theory

■ An interesting variation of adverse selection is the phenomenon known as the *winner's curse* which arises in competitive bidding. Intuitively, the idea is that a bidder is more likely to win an auction when he overestimates the value of the object being sold than when he underestimates it. Consequently, bidders who make unbiased value estimates will find that, on average, they have overestimated the value of the objects they win at auction. As noted by Milgrom (1979, 1981), bidders can earn positive profits, despite the winner's curse, by adjusting their bids downwards and by gathering extra information to improve the accuracy of their estimates.

Let us now be more specific. Consider a sealed-bid tender auction for the mineral rights on some unexplored tract of land. The U.S. Department of the Interior periodically conducts such auctions for potential oil-bearing tracts.

Let $\tilde{\theta}$ denote the value of the oil on the tract and let $\tilde{x}_i$ represent bidder i's estimate of that value. Let there be n competitors in the auction and suppose competitor i tenders a bid of $B_i(\tilde{x}_i)$. Assume that, conditioned on $\tilde{\theta}$, $\tilde{x}_1, \ldots, \tilde{x}_n$ are independent and drawn from families of distributions with the MLRP. Then one can sensibly assume that each B_i is nondecreasing.[8]

If bidder 1 were to win the auction with a bid of b, he would find, on average, that the value of the oil was $E[\tilde{\theta}|\tilde{x}_1, B_2(\tilde{x}_2) < b, \ldots, B_n(\tilde{x}_n) < b]$. Plainly, the news $\{B_i(\tilde{x}_i) < \infty\}$ is neutral (because it conveys no information), so that by Proposition 4, $\{B_i(\tilde{x}_i) < b\}$ is bad news for any finite b. Thus, the estimate $E[\tilde{\theta}|\tilde{x}_1]$ exceeds the average value expectation $E[\tilde{\theta}|\tilde{x}_1, B_2(\tilde{x}_2) < b, \ldots, B_n(\tilde{x}_n) < b]$; this is precisely the winner's curse.

Curiously, Proposition 4 implies that the average value expectation is an increasing function of b: higher bids alleviate the winner's curse. Intuitively, when a very high bid wins an auction, little can be inferred from the competitors' failure to place higher bids. But when a low bid wins, one can infer that the others had relatively low estimates of $\tilde{\theta}$.

7. Conclusion

■ This article introduces and develops the idea that individual pieces of information can be ordered by favorableness. Four applications are considered. In a securities market model, more favorable news about a security's future returns leads to a higher price for that security. In a principal-agent model, when high profits constitute favorable evidence about the agent's effort, the op-

[8] Equilibrium analyses of such a model have been given by R. Wilson (1977) and Milgrom and Weber (1980). At the equilibria of these models, the bidding strategies are increasing functions.

timal incentive contract entails a steeper fee schedule than does any efficient risk sharing contract. In a model of a sales encounter, the salesman reports the most favorable data about his product and the buyer takes a skeptical view of any information the salesman conceals. In an auction model, the winner reduces his estimate of the value of the object being sold when he learns that he has won, because winning implies that other bidders have relatively low value estimates. The "more favorable than" relation and the related ideas developed in this article make it easy to analyze these models and to interpret the results. Perhaps this fact is, in itself, favorable news about the quality of those ideas.

References

AKERLOF, G. "The Market for Lemons: Qualitative Uncertainty and the Market Mechanism." *Quarterly Journal of Economics*, Vol. 84 (1970), pp. 488–500.

———. "The Economics of Caste, the Rat Race, and Other Woeful Tales." *Quarterly Journal of Economics*, Vol. 90 (1976), pp. 599–617.

COX, J. AND ROSS, S. "The Valuation of Options for Alternative Stochastic Processes." *Journal of Financial Economics*, Vol. 3 (1976), pp. 145–166.

GROSSMAN, S. "The Role of Warranties and Private Disclosure about Product Quality." Working Paper No. 80-28, White Center for Financial Research, University of Pennsylvania, 1980.

——— AND HART, O. "An Analysis of the Principal-Agent Problem." CARESS Working Paper #80-17, University of Pennsylvania, 1980.

HARRISON, J.M. AND KREPS, D. "Martingales and Arbitrage in Multiperiod Securities Markets." *Journal of Economic Theory*, Vol. 20 (1979), pp. 381–408.

HOLMSTRÖM, B. "Moral Hazard and Observability." *Bell Journal of Economics*, Vol. 19 (1979), pp. 74–91.

KREPS, D. AND WILSON, R. "Sequential Equilibria." Stanford University Graduate School of Business Research Paper No. 584, 1980.

MILGROM, P. "A Convergence Theorem for Competitive Bidding with Differential Information." *Econometrica* (1979), pp. 679–688.

———. "Rational Expectation, Information Acquisition, and Competitive Bidding." *Econometrica* (1981), pp. 921–943.

——— AND ROBERTS, D.J. "Limit Pricing and Entry under Incomplete Information: An Equilibrium Analysis." Graduate School of Business Research Paper No. 574, Stanford University, 1980.

——— AND WEBER, R. "A Theory of Auctions and Competitive Bidding." Center for Mathematical Studies in Economics and Management Science, Discussion Paper No. 447, Northwestern University, 1980.

ORTEGA-REICHERT, A. "Models for Competitive Bidding under Uncertainty." Ph.D. dissertation, Stanford University, 1968.

PAULY, M. "Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection." *Quarterly Journal of Economics*, Vol. 88 (1974), pp. 44–54.

ROTHSCHILD, M. AND STIGLITZ, J. "Equilibrium in Competitive Insurance Markets: An Essay in the Economics of Imperfect Information." *Quarterly Journal of Economics*, Vol. 90 (1976), pp. 629–650.

SPENCE, A.M. "Job Market Signalling." *Quarterly Journal of Economics*, Vol. 87 (1973), pp. 355–374.

——— AND ZECKHAUSER, R. "Insurance, Information, and Individual Action." *American Economic Review*, Vol. 61 (1971), pp. 380–387.

VICKREY, W. "Counterspeculations, Auctions, and Competitive Sealed Tenders." *Journal of Finance*, Vol. 16 (1961), pp. 8–37.

———. "Auctions and Bidding Games" in *Recent Advances in Game Theory*, Princeton University Conference, 1962, pp. 15–27.

WILSON, C. "A Model of Insurance Markets with Incomplete Information." *Journal of Economic Theory*, Vol. 16 (1977), pp. 167–207.

WILSON, R. "A Bidding Model of Perfect Competition." *Review of Economic Studies*, Vol. 44 (1977), pp. 511–518.