Chapter 25

SIGNALLING

DAVID M. KREPS

Stanford University and Tel Aviv University

JOEL SOBEL

University of California at San Diego

Contents

Handbook of Game Theory, Volume 2, Edited by R.J. Aumann and S. Hart

1. Introduction

One of the most important applications of game theory to micro-economics has been in the domain of market signalling. The standard story is a simple one: Two parties wish to engage in exchange; the owner of a used car, say, wants to sell the car to a potential buyer. One party has information that the other lacks; e.g., the seller knows the quality of the particular car. (Think in terms of a situation where the quality of a car is outside of the control of the owner; quality depends on things done or undone at the factory when the car was assembled.) Under certain conditions, the first party wishes to convey that information to the second; e.g., if the car is in good condition, the seller wishes the buyer to learn this. But direct communication of the information is for some reason impossible, and the first party must engage in some activity that indicates to the second what the first knows; e.g., the owner of a good car will offer a limited warranty. The range of applications for this simple story is large: A worker wishes to signal his ability to a potential employer, and uses education as a signal [Spence (1974)]. An insuree who is relatively less risk prone signals this to an insurer by accepting a larger deductable or only partial insurance [Rothschild and Stiglitz (1976), Wilson (1977)]. A firm that is able to produce high-quality goods signals this by offering a warranty for the goods sold [Grossman (1981)]. A plaintiff with a strong case demands a relatively high payment in order to settle out of court [Reinganum and Wilde (1986), Sobel (1989)]. The purchaser of a good who does not value the good too highly indicates this by rejecting a high offer or by delaying his own counteroffer [Rubinstein (1985), Admati nd Perry (1987)]. A firm that is able to produce a good at a relatively low cost signals this ability to potential rivals by charging a low price for the good [Milgrom and Roberts (1982)]. A strong deer grows extra large antlers to show that it can survive with this handicap and to signal its fitness to potential mates [Zahavi (1975)].

The importance of signalling to the study of exchange is manifest, and so the original work on market signalling [Spence (1974), Rothschild and Stiglitz (1976), Wilson (1977)] received a great deal of attention. But the early work, which did not employ formal game theory, was inconclusive; different papers provided different and sometimes contradictory answers, and even within some of the papers, a welter of possible equilibria was advanced. While authors could (and often did) select among the many equilibria they found with informal, intuitive arguments, the various analyses in the literature were ad hoc. Formal game theoretic treatments of market signalling, which began to appear in the 1980s, added discipline to this study. It was found that the different and conflicting results in the early literature arose from (implicit) differences in the formulation of the situation as a game. And selections among equilibria that were previously based on intuitive arguments were rationalized with various refinements of Nash equilibrium, most especially refinements related to Kohlberg and Merten's notion of a stable equilibrium (cf. the Chapter on strategic stability in Volume III of this Handbook) and to out-of-equilibrium beliefs in the sense of a sequential equilibrium. This is not to say that game theory showed that some one of the early analyses was correct and the others were wrong. Rather, game theory has contributed a language with which those analyses can be compared and evaluated.

In this chapter, we survey some of the main developments in the theory of market signalling and its connection to noncooperative game theory. Our treatment is necessarily restricted in scope and in detail, and in some concluding remarks we point the reader toward the vast number of topics that we have omitted.

2. Signalling games – the canonical game and market signalling

Throughout this chapter, we work with the following canonical game. There are two players, called **S** (for sender) and **R** (for receiver). This is a game of incomplete information: **S** starts off knowing something that **R** does not know. We assume that **S** knows the value of some random variable t whose support is a given set T. The conventional language, used here, is to say that t is the *type* of **S**. The prior beliefs of **R** concerning t are given by a probability distribution ρ over T; these prior beliefs are common knowledge. Player **S**, knowing t, sends to **R** a signal s, drawn from some set S. (One could have the set of signals available to **S** depend on t.) Player **R** receives this signal, and then takes an action a drawn from a set A (which could depend on the signal s that is sent). This ends the game: The payoff to **S** is given by a function $u: T \times S \times A \to R$, and the payoff to **R** is given by $v: T \times S \times A \to R$.

This canonical game captures some of the essential features of the classic applications of market signalling. We think of **S** as the seller or buyer of some good: a used car, labor services, an insurance contract; **R** represents the other party to the transaction. **S** knows the quality of the car, or his own abilities, or his propensity towards risk; **R** is uncertain a priori, so t gives the value of quality or ability or risk propensity. **S** sends a signal r that might tell **R** something about t. And **R** responds – in market signalling, this response could be simply acceptance/rejection of the deal that **S** proposes, or it could be a bid (or ask) price at which **R** is willing to consummate the deal.

To take a concrete example, imagine that **S** is trying to sell a used car to **R**. The car is either a *lemon* or a *peach*; so that $T = \{\text{lemon, peach}\}$. Write t_0 for "lemon" and t_1 for "peach". **S** knows which it is; **R** is unsure, assessing probability $\rho(t_1)$ that the car is a peach. **S** cannot provide direct evidence as to the quality of the car, but **S** can offer a warranty – for simplicity, we assume that **S** can offer to cover all repair expenses for a length of time s that is one of: zero, one, two, three or four months. We wish to think of the market in automobiles as being competitive in the sense that there are many buyers; to accommodate this within the framework of our two-person game is nontrivial, however. Two ways typically used to do this are: (1) Imagine that there are (at least) two identical buyers, called $\mathbf{R}_1$ and $\mathbf{R}_2$. The game is then structured so that $S = \{0, 1, 2, 3, 4\}$ (the possible warranties that could be offered); **S** announces one of these warranty lengths. The $\mathbf{R}_1$ and $\mathbf{R}_2$ respond by simultaneously and independently naming a price $a_i \in [0, \infty)$ $(i = 1, 2)$ that they are willing to pay for the car, with the car (with warranty) going to the high bidder at that bid. In the usual fashion of Bertrand competition or competitive bidding when bidders have precisely the same information, in equilibrium $a_1 = a_2$ and each is equal to the (subjective) valuation placed on the car (given the signal s) by the two buyers. If we wished to follow this general construction but keep precisely to our original two-player game formulation, we could artificially adopt a utility function for **R** that causes him to bid his subjective valuation for the car. For example, we could suppose that his utility function is minus the square of the difference between his bid and the value he places on the car. (2) Alternatively, we can imagine that $S = \{0, 1, 2, 3, 4\} \times [0, \infty)$. The interpretation is that **S** makes a take-it-or-leave-it offer of the form: a length of warranty (from $\{0, 1, 2, 3, 4\}$) and a price for the car (from $[0, \infty)$). And **R** either accepts or rejects this offer. Since **S** has all the bargaining power in this case, he will extract all the surplus. We hereafter call this the *take-it-or-leave-it* formulation of signalling.

Either of these two games captures market signalling institutional details in a manner consistent with our canonical game. But there are other forms of institutions that give a very different game-theoretic flavor to things. For example, we could imagine that **R** (or many identical **R**s) offer a menu of contracts to **S**; that is, a set of pairs from $\{0, 1, 2, 3, 4\} \times [0, \infty)$, and then **S** chooses that contract that he likes the most. Models of this sort, where the uninformed party has the leading role in setting the terms of the contract, are often referred to as examples of market *screening* instead of *signalling*, to distinguish them from institutions where the informed party has the leading role. We will return later to the case of market screening.

3. Nash equilibrium

We describe a Nash equilibrium for the canonical signalling game in terms of behavior strategies for **S** and **R**. First take the case where the sets T, S and A are all finite. Then a behavior strategy for **S** is given by a function $\sigma: T \times S \to [0, 1]$, such that $\sum_s \sigma(t, s) = 1$ for each t. The interpretation is that $\sigma(t, s)$ is the probability that **S** sends message s if **S** is of type t. A behavior strategy for **R** is a function $\alpha: S \times A \to [0, 1]$ where $\sum_a \alpha(s, a) = 1$ for each s. The interpretation is that **R** takes action a with probability $\alpha(s, a)$ if signal s is received.

Proposition 1. Behavior strategies α for **R** and σ for **S** form a Nash equilibrium if and only if

$$\sigma(t,s) > 0 \text{ implies } \sum_a \alpha(s,a)u(t,s,a) = \max_{s' \in S}\left(\sum_a \alpha(s',a)u(t,s',a)\right), \tag{3.1}$$

and, for each s such that $\sum_t \sigma(t,s)\rho(t) > 0$,

$$\alpha(s,a) > 0 \text{ implies } \sum_t \mu(t;s)v(t,s,a) = \max_{a'} \sum_t \mu(t;s)v(t,s,a'), \tag{3.2a}$$

where we define

$$\mu(t;s) = \frac{\sigma(t,s)\rho(t)}{\sum_{t'} \sigma(t',s)\rho(t')} \text{ if } \sum_t \sigma(t,s)\rho(t) > 0. \tag{3.2b}$$

Condition (3.1) says that σ is a best response to α, while (3.2) says (in two steps) that α is a best response to σ; (3.2b) uses Bayes' rule and α to compute **R**'s posterior beliefs $\mu(\cdot;s)$ over T upon hearing signal s (if s is sent with positive probability), and (3.2a) then states that $\alpha(s,\cdot)$ is a conditional best response (given s). Note well that the use of Bayes' rule when applicable [in the fashion of (3.2b)] and (3.2a) is equivalent to α being a best response to σ in terms of the ex ante expected payoffs of **R** in the associated strategic-form game.

All this is for the case where the sets T, S and A are finite. In many applications, any or all of these sets (but especially S and A) are taken to be infinite. In that case the definition of a Nash equilibrium is a straightforward adaptation of what is given above: One might assume that the spaces are sufficiently nice (i.e., Borel) so that a version of regular conditional probability for t given s can be fixed (where the joint probabilities are given by ρ and σ), and then (3.2a) would use conditional expectations computed using that fixed version.

We characterize some equilibria as follows:

Definition. An equilibrium (σ, α) is called a *separating* equilibrium if each type t sends different signals; i.e., the set S can be partitioned into (disjoint) sets $\{S_t; t \in S\}$ such that $\sigma(t, S_t) = 1$. An equilibrium (σ, α) is called a *pooling* equilibrium if there is a single signal s^* that is sent by all types; i.e., $\sigma(t, s^*) = 1$ for all $t \in T$.

Note that we have not precluded the possibility that, in a separating equilibrium, one or more types of **S** would use mixed strategies. In most applications, the term is used for equilibria in which each type sends a single signal, which we might term a pure separating equilibrium. On the other hand, we have followed convention in using the unmodified expression *pooling equilibrium* for equilibria in which all types use the same pure behavior strategy. One can imagine definitions of pooling equilibria in which **S** uses a behaviorally mixed strategy (to some extent), but by virtue of Proposition 2 following, this possibility would not come up in standard formulations. Of course, these two categories do not exhaust all possibilities. We can have equilibria in which some types pool and some separate, in which all types pool with at least one other type but in more than one pool, or even in which some types randomize between signals that separate them from other types and signals that pool them with other types.

4. Single-crossing

In many of the early applications, a "single-crossing" property held. The sets T, S and A each were simply ordered, with $\geqslant$ used to denote the simple order in each case. (By a simple order, we mean a complete, transitive and antisymmetric binary relation. The single-crossing property has been generalized to cases where S is multi-dimensional; references will be provided later.) We will speak as if T, S and A are each subsets of the real line, and $\geqslant$ will be the usual "greater than or equal to" relationship. Also, we let $\mathscr{A}$ denote the set of probability distributions on A and, for each $s \in S$ and $T' \subseteq T$, we let $\mathscr{A}(s, T')$ be the set of mixed strategies that are best responses by **R** to s for some probability distribution with support T'. Finally, for $\alpha \in \mathscr{A}$, we write $u(t, s, \alpha)$ for $\sum_{a \in A} u(t, s, \alpha)\alpha(a)$.

Definition. The data of the game are said to satisfy the *single-crossing property* if the following holds: If $t \in T$, $(s, \alpha) \in S \times \mathscr{A}$ and $(s', \alpha') \in S \times \mathscr{A}$ are such that $\alpha \in \mathscr{A}(s, T), \alpha' \in \mathscr{A}(s', T), s > s'$ and $u(t, s, \alpha) \geqslant u(t, s', \alpha')$, then for all $t' \in T$ such that $t' > t$, $u(t', s, \alpha) > u(t', s', \alpha')$.

In order to understand this property, it is helpful to make some further assumptions that hold in many examples. Suppose that S and A are compact subsets of the real line and that u is defined for all of $T \times \mathrm{co}(S) \times \mathrm{co}(A)$, where $\mathrm{co}(X)$ denotes the convex hull of X. Suppose that u is continuous in its second two arguments, strictly decreasing in the second argument and strictly increasing in the third. In terms of our example, think of *peach* $>$ *lemon*, and then the monotonicity assumptions are that longer warranities are worse for the seller (if $s \in S$ is the length of the warranty) and a higher price is better for the seller (if $a \in A$ is the purchase price).

Then we can draw in the space $\mathrm{co}(S) \times \mathrm{co}(A)$ indifference curves for each type t. The monotonicity and continuity assumptions will guarantee that these indifference curves are continuous and increasing. And the single-crossing property will imply that if indifference curves for types t and t' cross, for $t' > t$, then they cross once, with the indifference curve of the higher type t' crossing from "above on the left" to "below on the right". The reader may wish to draw this picture. Of course, the single-crossing property says something more, because it is not necessarily restricted to pure strategy responses. But in many standard examples, either u is linear in the third argument [i.e., $u(t, s, a) = u'(t, s) + a$] or v is strictly concave in a and A is convex [so that $\mathscr{A}(s, T')$ consists of degenerate distributions for all s], in which case single-crossing indifference curves will imply the single-crossing property given above.

The single-crossing property can be used to begin to characterize the range of equilibrium outcomes.

Proposition 2. Suppose that the single-crossing property holds and, in a given equilibrium (σ, α), $\sigma(t, s) > 0$ and $\sigma(t', s') > 0$ for $t' > t$. Then $s' \geqslant s$.

That is, the supports of the signals sent by the various types are "increasing" in the type. This does not preclude pooling, but a pool must be an interval of types pooling on a single signal, with any type in the interior of the interval sending only the pooling signal (and with the largest type in the pool possibly sending signals larger than the pooling signal and the smallest type possibly sending signals smaller than the pooling signal).

While the single-crossing property holds in many examples in the literature, it does not hold universally. For example, this property may not hold in the entry deterrence model of Milgrom and Roberts (1982) [see Cho (1987)], nor it is natural in models of litigation [see Reinganum and Wilde (1986) and Sobel (1989)].

More importantly to what follows, this entire approach is not well suited to some variations on the standard models. Consider the example of selling a used car with the warranty as signal, but in the variation where **S** offers to **R** a complete set of terms (both duration of the warranty and the purchase price), and **R**'s response is either to accept or reject these terms. In this game form, S is not simply ordered, and so the single-crossing property as given above makes no sense at all.

Because we will work with this sort of variation of the signalling model in the next section, we adapt the single-crossing property to it. First we give the set-up.

Definition. A signalling game has the *basic take-it-or-leave-it setup* if: (a) T is finite and simply ordered; (b) $S = M \times (-\infty, \infty)$ for M an interval of the real line [we write $s = (m, d)$]; (c) $A = \{yes, no\}$; (d) $u(t, (m, d), a) = U(t, m) + d$ if $a = yes$ and $u(t, (m, d), a) = 0$ if $a = no$; and (e) $v(t, (m, d), a) = V(t, m) - d$ if $a = yes$ and $= 0$ if $a = no$; where U, W and V are continuous in m; V is strictly increasing in t and nondecreasing in m; and U is strictly decreasing in m.

The interpretation is that m is the message part of the signal s (such as the length of the warranty on a used car) and d gives the proposed "dollar price" for the exchange. Note that we assume proposed dollar prices are unconstrained. For expositional simplicity, we have assumed that the reservation values to S and to R if there is no deal are constants set equal to zero.

Definition. In a signalling game with the basic take-it-or-leave-it setup, the *single-crossing property* holds if $t' > t$, $m' > m$ and $U(t, m') + d' \geqslant U(t, m) + d$, then $U(t', m') + d' > U(t', m) + d$.

Proposition 2′. Fix a signalling game with the basic take-it-or-leave-it setup for which the *single-crossing property* holds. In any equilibrium in which type t proposes (m, d), which is accepted with probability p, and type t' proposes (m', d'), which is accepted with probability $p', t' > t$ and $p' \geqslant p$ imply $m' \geqslant m$.

5. Refinements and the Pareto-efficient separating equilibrium

Proposition 2 (and 2′) begins to characterize the range of Nash equilibria possible in the standard applications, but still there are typically many Nash equilibria in a given signalling game. This multiplicity of equilibria arises in part because the response of R to messages that have zero prior probability under σ are not constrained by the Nash criterion; i.e., (3.2a) restricts $\alpha(s, \cdot)$ only for s such that $\sum_t \sigma(t, s)\rho(t) > 0$. However the response of **R** to so-called *out-of-equilibrium* messages can strongly color the equilibrium, since the value of sending out-of-equilibrium messages by **S** depends on the equilibrium response to those messages. That is, **S** may not send a particular message s° because **R** threatens to blow up the world in response to this message; since (therefore) **S** will not send s°, the threatened response is not disallowed by (3.2a). This is clearly a problem coming under the general heading of "perfection", as discussed in the Chapters on "Strategic Equilibrium" and "Conceptual Foundations of Strategic Equilibrium" in volume III of this Handbook. And one naturally attacks this problem by refining the Nash criterion. The first step in the usual attack is to look for sequential equilibria. Hereafter, μ will denote a full set of beliefs for **R**; i.e., for each $s \in S$, $\mu(\cdot, s)$ is a probability distribution on T.

Proposition 3. Behavior strategies (σ, α) and beliefs μ for a signalling game constitute a sequential equilibrium if (3.1) holds, (3.2a) holds for every $s \in S$, and (3.2b) holds as a condition instead of a definition.

(The only thing that needs proving is that strategies and beliefs are consistent in the sense of sequential equilibrium, but in this very simple setting this is true automatically. The formal notion of a sequential equilibrium does not apply when T, S, or A is infinite; in such cases is it typical to use the conditions of Proposition 3 as a definition.)

Restricting attention to sequential equilibria does reduce the number of equilibria, in that threats to "blow up the world" are no longer credible. But multiple equilibria remain in interesting situations. For example, in applications where the types $t \in T$ and the responses $a \in A$ are simply ordered, u is increasing in the response a, and the responses by **R** at each signal s increase with stochastic increases in **R**'s assessment as to the type of **S**, we can construct many sequential equilibrium outcomes where **R** "threatens with beliefs" – for out-of-equilibrium messages s, **R** holds beliefs that put probability one on the message coming from the worst (smallest) type $t \in T$, and **R** takes the worst (smallest) action consistent with those beliefs and with the message. This, in general, will tend to keep **S** from sending those out-of-equilibrium messages.

Accordingly, much of the attention in refinements applied to signalling games has been along the lines: At a given equilibrium (outcome), certain out-of-equilibrium signals are "unlikely" to have come from types of low index. These out-of-equilibrium signals should therefore engender out-of-equilibrium beliefs that put relatively high probability on high index types. This then causes **R** to respond to those signals with relatively high index responses. And this, in many cases, will cause the equilibrium to fail.

There are a number of formalizations of this line of argument in the literature, and we will sketch only the simplest one here, which works well in the basic take-it-or-leave-it setup. (For more complete analysis, see Banks and Sobel (1987) and Cho and Kreps (1987). See also Farrell (1993) and Grossman and Perry (1987) for refinements that are similar in motivation.)

The criterion we use is the so-called *intuitive criterion*, which was introduced informally in Grossman (1981), used in other examples subsequently, and was then codified in Cho and Kreps (1987).

Definition. For a given signalling game, fix a sequential equilibrium (σ, α, μ). Let $u^*(t)$ be the equilibrium expected payoff to type t in this equilibrium. Define $B(s) = \{t \in T : u(t, s, \alpha) < u^*(t) \text{ for every } \alpha \in \mathscr{A}(s, T)\}$. The fixed sequential equilibrium *fails the intuitive criterion* if there exist $s \in S$ and $t^* \in T \backslash B(s)$ such that $u(t^*, s, \alpha) > u^*(t^*)$ for all $\alpha \in \mathscr{A}(s, T \backslash B(s))$.

The intuition runs as follows. If $t \in B(s)$, then, relative to following the equilibrium, type t has no incentive to send signal s; no matter what **R** makes of s, **R**'s response leaves t worse off than if t follows the equilibrium. Hence, if **R** receives the signal s, **R** should infer that this signal comes from a type of **S** drawn from $T \backslash B(s)$. But if **R** makes this inference and acts accordingly [taking some response from $\mathscr{A}(s, T \backslash B(s))$], type t^* is sure to do better than in the equilibrium. Hence, type t^* will defect, trusting **R** to reason as above.

This criterion (and others from the class of which it is representative) puts very great stress on the equilibrium outcome as a sort of status quo, against which defections are measured. The argument is that players can "have" their equilibrium values, and defections are to be thought of as a reasoned attempt to do better. This aspect of these criteria has been subject to much criticism: If a particular equilibrium is indeed suspect, then its values cannot be taken for granted. See, for example, Mailath et al. (1993). Nonetheless, this criterion can be justified on theoretical grounds.

Proposition 4. [Banks and Sobel (1987), Cho and Kreps (1987)]. If T, S and A are all finite, then for generically chosen payoffs any equilibrium that fails the intuitive criterion gives an outcome that is not strategically stable in the sense of Kohlberg and Mertens (1986).

This criterion is very strong in the case of basic take-it-or-leave-it games that satisfy the single-crossing property and that are otherwise well behaved. Please note the conflict with the hypothesis of Proposition 4. In the take-it-or-leave-it games, S is uncountably infinite. Hence, the theoretical justification for the intuition criterion provided by Proposition 4 only applies "in spirit" to the following results. Of course, the intuitive criterion itself does not rely on the finiteness of S.

The result that we are headed for is that, under further conditions to be specified, there is a single equilibrium outcome in the basic take-it-or-leave-it game that satisfies the intuitive criterion. This equilibrium outcome is pure separating, and can be loosely characterized as the first-best equilibrium for **S** subject to the separation constraints. This result is derived in three steps. First, it is shown that pooling is impossible except at the largest possible value of m. Then, more or less by assumption, pooling at that extreme value is ruled out. This implies that the equilibrium outcome is separating, and the unique separating equilibrium outcome that satisfies the intuitive criterion is characterized.

Proposition 5. Fix a basic take-it-or-leave-it game satisfying the single-crossing property. Then any equilibrium in which more than one type sends a given signal (m, d) with positive probability to which the response is *yes* with positive probability can satisfy the intuitive criterion only if m equals its highest possible value.

A sketch of the proof runs as follows: Suppose that, in an equilibrium, more than one type pooled at the signal (m, d) with m strictly less than its highest possible value, the response to which is *yes* with positive probability. Let t^* be the highest type in the pool. Then it is claimed that for any $\varepsilon > 0$, there is an unsent (m', d') with (a) $m < m' < m + \varepsilon$ and $d < d' < d + \varepsilon$, (b) $U(t^*, m') + d' > U(t^*, m) + d$, and (c), for all $t < t^*$, $U(t, m') + d' < U(t, m) + d$. (The demonstration of this is left to the reader. The single-crossing property is crucial to this result.)

Suppose that type t^* proposes (m', d'). By (c), $t \in B(s)$ for all $t < t^*$. Hence, **R** in the face of this deviation must hold beliefs that place probability one on types t^* or greater. Since V is strictly increasing in t, for ε sufficiently small, if **R** was willing to accept (m, d) with positive probability at the pool, with beliefs restricted to types t^* or greater, **S** must accept (m', d') with probability one. But then (b) ensures that t^* would deviate; i.e., the intuitive criterion fails.

Proposition 5 does not preclude the possibility that pooling occurs at the greatest possible level of m. To rule this possibility out, the next step is to make assumptions sufficient to ensure that this cannot happen. The simplest way is to assume that M is unbounded above, and we proceed for now on that basis. Hence by Proposition 5, attention can be restricted to equilibria where types either separate or do not trade with positive probability.

Proposition 6. Fix a basic take-it-or-leave-it game that satisfies the following three supplementary assumptions. (a) For all t, $\max_m U(t,m) + V(t_0,m) > 0$, where t_0 is the type with lowest index. (b) For all t, $U(t,m) + V(t,m)$ is strictly quasi-concave in m. (c) For each type t, if t' is the next-lowest type, then the problem $\max_m U(t,m) + V(t,m)$ subject to $U(t',m) + V(t,m) \leqslant u$ has a solution for every value of u. (If t is the lowest type, ignore the constraint.) Then there is a unique equilibrium outcome for sequential equilibria that satisfy the intuitive criterion and in which types either separate or do not trade.

We sketch the argument (and show how to construct this unique equilibrium outcome). Fix a sequential equilibrium that survives the intuitive criterion and that is separating (or where some types trade with probability zero). Let m_0 solve $\max_m U(t_0,m) + V(t_0,m)$, where t_0 is the lowest index type. [By supplementary assumption (b), this solution is unique.] By proposing $(m_0, V(t_0,m_0) - \varepsilon)$ for $\varepsilon > 0$, type t_0 can be sure of acceptance, because **R**'s beliefs can be no worse than that this proposal comes from t_0 with certainty. Hence, t_0 can be sure to get utility $u^*(t_0) := U(t_0,m_0) + V(t_0,m_0)$ in any sequential equilibrium. [Moreover, a similar argument using supplementary assumption (a) ensures that the no-trade outcome will not be part of the equilibrium.] Since the equilibrium is separating, this is also an upper bound on what t_0 can get; so we know that the equilibrium outcome for t_0 in the equilibrium gives $u^*(t_0)$ to t_0, and (hence) t_0, in this equilibrium, must propose precisely $(m_0, V(t_0,m_0))$, which is accepted with probability one. Let t_1 be the type of second-lowest index. By use of the intuitive criterion, it can be shown that this type can be certain of utility $\max_m\{U(t_1,m) + V(t_1,m)\colon U(t_0,m) + V(t_1,m) \leqslant u^*(t_0)\}$, since by proposing the (unique) maximizing m and a payment a bit less than $V(t_1,m)$ for this m, acceptance is guaranteed; since t_0 receives strictly less than $u^*(t_0)$ whether or not **R** accepts the proposal, $t_0 \in B((m, V(t_1,m)^-))$. This is also an upper bound for t_1's payoff in any separating equilibrium in which t_0 gets $u^*(t_0)$, and so this gives the outcome for t_1 in the equilibrium. And so on. We can build up the equilibrium outcome by, for each type, maximizing its payoff, assuming it is separated and that it is constrained to send a signal that the next-lower type would not (given its equilibrium value, determined a stage earlier).

Remarks. (1) This particular equilibrium, the optimal equilibrium for **S** subject to a separation constraint, goes back in the literature to well before its "justification" by stability-related refinements. The notion that, in equilibrium, there will be overinvestment in signals for purposes of separation can be found in the very first work on market signalling, by Spence (1974) and by Rothschild and Stiglitz (1976). This precise equilibrium is singled out by Riley (1979) on other grounds (see Section 6 following).

(2) Supplementary assumption (a) in Proposition 6 can be paraphrased as "trade is beneficial", even at terms appropriate for the lowest index type. It is rather

strong and should (and can) be relaxed, however there is no general treatment of this written down yet.

(3) We avoided equilibria in which there is pooling at the highest value of m by fiat, by assuming that M is unbounded. But then we needed to assume that U and V are sufficiently well-behaved so that supplementary assumption (c) holds. Alternatively, we could assume that M is compact [so that (c) is no longer required] and either make assumptions sufficient to guarantee that there is no pooling at the "top" or deal with the possibility of such a pool. For a careful treatment of this, see Cho and Sobel (1990).

(4) This argument is finely tuned to the take-it-or-leave-it game form, since in this game form, any type of **S**, having the ability to propose an entire "deal", has very fine ability to distinguish himself from lower types. In the more standard game form, **S** proposes only m (this becomes the signal), and **R** (or more than one **R**) responds with a proposed price d (which becomes a). When there are more than two types of **S**, the intuitive criterion is generally not strong enough to imply full separation. In such cases, stronger restrictions than the intuitive criteria such as universal divinity [Banks and Sobel (1987)] are required. The basic idea is that for each signal s, the set $B(s)$ is enlarged to include any type t for which there is some other type t' with: If $u(t,s,\alpha) \geqslant u^*(t)$ for $\alpha \in \mathscr{A}(s,T)$, then $u(t',s,\alpha) > u^*(t')$. [This is not quite universal divinity, but instead is the slightly weaker D1 restriction of Cho and Kreps (1987).] Or, in words, $t \in B(s)$ if for any response to s that would cause t to defect, t' would defect. Using this enlarged $B(s)$ amounts to an assumption that **R**, faced with s, infers that there is no chance that s came from t; if type t of **S** were going to send s instead of sticking to the equilibrium, than type t' would certainly do so. Although less intuitive than the intuitive criterion, this is still an implication (in generic finite signalling games) of strategic stability. Together with the single-crossing property (and some technical conditions), it implies that Proposition 2 can be extended to out-of-equilibrium signals: If type t' sends signal s with positive probability, then for all $t < t'$ and $s' > s$, we must be able to support the equilibrium with beliefs that have $\mu(t,s') = 0$. With further monotonicity conditions, it then gives results similar to Propositions 5 and 6 above. See Cho and Sobel (1990) and Ramey (1988) for details.

(5) Perhaps the intuitively least appealing aspect of these results is that a very small probability of bad types can exert nonvanishing negative externality on good types. That is, if ninety-nine percent of all the cars in the world are peaches and only one percent are lemons, still the peach owners must offer large warranties to distinguish their cars from the lemons. You can see this mathematically by the fact that if t_0 is a lemon and t_1 a peach, then the peach owners must choose an m_1 satisfying $U(t_0,m_1) + V(t_1,m_1) \leqslant u^*(t_0)$. this constraint is independent of the number of peaches and lemons; it is there as long as there is a single lemon in the world. But if there are no lemons, then this constraint vanishes. Pooling equilibria, in which the lemon owners exert a negative externality on the peach owners that vanishes at the fraction of lemons goes to zero, seem intuitively more reasonable.

Mailath et al. (1993) require that the receiver interpret an out-of-equilibrium message as an attempt by some types of sender to shift to another, preferred equilibrium whenever such an interpretation is possible. This leads to what they call *undefeated equilibria*; in the current context, this selects a pooling outcome.

(6) The argument given depends on m being one-dimensional; more generally, the single-crossing condition developed in the previous section is formulated for the case of a one-dimensional signal. Engers (1987) generalizes the single-crossing condition to multi-dimensional signals, following earlier specific examples. Although Engers works with a screening model (see the next section), Cho and Sobel (1990) and Ramey (1988) use versions of single-crossing conditions for multi-dimensional signals to characterize equillibrium outcomes in the canonical signalling game.

Humans are not the only animals that communicate through indirect signals. The songs or plumage of birds, the antlers of deer, and the colors of fish may be viewed as signals designed to inform a potential mate about reproductive fitness. Zahavi (1975, 1977) argues that meaningful signals arise in biological settings when they can be reliably interpreted by their intended audience, and he informally proposes that a single-crossing property holds. Grafen (1990) provides a game theoretic treatment of Zahavi's ideas. Grafen's model is virtually identical with the basic Spence model with a continuum of types. He argues that the same pure separating equilibrium outcome given prominence in the economics literature is the unique evolutionarily stable outcome.

(7) The construction of the unique equilibrium outcome that survives the intuitive criterion made use of the assumption that T is a finite set. (The argument that no equilibrium with pooling would survive the intuitive criterion, on the other hand, did not require this assumption.) Many of the early examples, however, assumed that T is an interval, so it is useful to provide extensions in this direction. Mailath (1987) provides an analysis of separating equilibria when there is an interval of types, giving useful characterizations and an existence result. Ramey's (1988) analysis is posed in this sort of setting.

(8) It is also desirable to extend this sort of analysis to cases where types of the sender are drawn from a set which is multi-dimensional (i.e., only partially ordered). Not much has been done along these lines, although Kohlleppel (1983) and Quinzii and Rochet (1985) provide interesting analyses of specific cases.

6. Screening

In the canonical signalling game of Sections 2 and 3 and in the variations on this game described above, the party with private information takes the lead in deciding which signals will be sent. That is to say, the set of signals that can be sent is given exogenously, as part of the description of the game, and the uninformed party reacts to the signal that is chosen by the informed party.

In many applications to market signalling, it may seem more appropriate to model things with the uninformed party taking the lead. The simplest game form that one can imagine for this, specialized to the very concrete model of the take-it-or-leave-it setup, has a number (>1) of identical buyers (the uninformed side to a transaction) simultaneously and independently proposing a set of contracts, each contract taking the form (m, d). (Assume T is finite and that buyers are restricted to proposing a finite set of pairs.) Sellers (or the single seller, the informed side) look over the set of contracts on offer and choose at most one – the deal is then consummated at the terms chosen. Since the choice of a contract is a "signal", this formulation falls under the general rubric of signalling, although in the literature this sort of formulation, where the uninformed take the lead, is referred to as *screening* [Stiglitz and Weiss (1990)].

For this game form, the only possible pure-strategy equilibrium outcome is the separating outcome sketched in the previous section. The argument is very similar to the argument given in previous section, except that one starts with the observation that, by the usual Bertrand argument, each contract that is taken with positive probability must just "break even", and then think of a uninformed party making the sorts of offers that we had the informed parties making in the argument to break any other pure-strategy equilibrium. But, as observed first by Rothschild and Stiglitz (1976) and as refined by Riley (1979), this separating outcome is (often) not an equilibrium outcome either; it is often possible for an uninformed party to propose a pooling a contract which attracts a number of types and which is profitable.

By general existence results [Dasgupta and Maskin (1986)], we know an equilibrium exists for this game (under certain regularity conditions that hold in many applications). Hence, the only equilibria are in mixed strategies. This lacks appeal in many of the economic contexts which this is meant to model. Hence early authors were moved to modify the equilibrium concept employed.

There are two basic modifications that have been studied. The first basic modification is Wilson's (1977) E2 equilibrium. Here, roughly put, an outcome is an equilibrium if there is no contract that can be added to the set of contracts that are offered that will be profitable *after all contracts which then sustain losses are removed from the set on offer.* Wilson shows that an E2 equilibrium always exists (under standard conditions), that pooling can be an E2 equilibrium, and that there are robust examples with multiple E2 equilibria.

The second is Riley's (1979) *reactive equilibrium.* Again roughly put, an outcome is a reactive equilibrium if no additional contract added to the set on offer would make a profit if other firms are given the opportunity to react and *add* still more offers. Engers and Fernandez (1987) study this equilibrium notion in substantial generality, showing that under standard conditions, the only reactive equilibrium is the separating equilibrium of Section 5.

Both Wilson and Riley advance notions of equilibrium that are meant to be reduced form solution concepts for an unspecified dynamic process of competition.

The linkages to noncooperative (Nash) equilibrium of completely well-specified game forms came later. Hellwig (1986) studies the game form where buyers make offers simultaneously, sellers choose contracts, but then buyers have the right to refuse any seller. This game admits many Nash (and even sequential) equilibria, but Hellwig announces the result that the only stable equilibrium outcome is the pooling contract which is most favorable to the highest type, at least for the case where there are only two types. Engers and Fernandez (1987), on the other hand, study a game form where buyers simultaneously name contracts; their contract choices are revealed, and then buyers are given the opportunity to add to the set of contracts offered; if any additions are made, buyers are given another opportunity to add contracts; and so on, until on some given round buyers all "pass" on the opportunity to add contracts, at which point sellers (or the single seller) chooses his most preferred contract. Any reactive equilibrium of the original game is a Nash equilibrium for this game, but there are many other Nash equilibria; and Engers and Fernandez do not explore whether any of the standard refinements will help pin things down.

A further variation on market screening concerns the case of a single uninformed party who offers the informed party a menu of contracts on a take-it-or-leave-it basis. That is, all the bargaining power is given to the uninformed party. The seminal reference to this sort of analysis is Stiglitz (1977), who shows that this reduces formally to a problem of a monopolist supplier who is able to offer nonlinear prices. He therefore conducts analysis analogous to that in the literature on optimal income taxation [Mirrlees (1971)] to characterize the monopolist's optimal set of contracts. Complications due to strategic interactions disappear and there is no problem guaranteeing the existence of equilibrium. In contrast to the outcomes obtained in signalling games or in competitive screening models, the monopolist's optimal set of contracts need not be fully separating (when there are at least three types of informed party).

7. Costless signalling and neologisms

We next specialize to the case where messages are costless; i.e., where $u(t,s,a) = u(t,a)$ and $v(t,s,a) = v(t,a)$. In this case, even though "talk is cheap", signalling may be of use, insofar as there is some commonality of interests between **S** and **R**.

Green and Stokey (1980) and Crawford and Sobel (1982) provide the first analyses of this case. Crawford and Sobel assume: (a) $T = [0,1]$, and the prior distribution of t on T is absolutely continuous. (b) The space A is a connected interval from the real line, for every t, $u(t,\cdot)$ and $v(t,\cdot)$ are concave, and, denoting partials of u and v by subscripts in the usual fashion, $u_2(t,\cdot) = 0$ and $v_2(t,\cdot) = 0$ have solutions in A. (c) $u_{12} > 0$ and $v_{12} > 0$, so that the solutions of $u_2(t,\cdot) = 0$ and $v_2(t,\cdot) = 0$ as functions of t both increase with t. (d) Finally, if we write $a_u(t)$ for the solution of $u_2(t,\cdot) = 0$ and $a_v(t)$ for the solution of $v_2(t,\cdot) = 0$, then $a_u(t) \neq a_v(t)$ for all t. With these assumptions, they show that all equilibria are essentially of the following form: There is a finite partition of T into intervals, and S's signal indicates which of the cells of the partition contains the true value of t, with **R** responding according to the posterior so generated. The strong monotonicity assumptions guarantee that the response taken "increases" with increases in the cell of the partition. Moreover, with a further monotonicity assumption, there are (essentially) finitely many of these equilibria; one for a partition which is trivial (no information is communicated), and one for partitions of each integer size up to some largest-sized partition. Further, the expected utility of both **S** and **R** increases as the number of cells in the partition increases. Finally, they give a sense in which one can say that preferences of **S** and **R** are more closely aligned, and they establish that the more closely aligned are these preferences, the greater the cardinality of the maximal partition.

All cheap-talk games have a no-communication or babbling equilibrium in which S's signal is not informative, and **R** responds to all signals (on the equilibrium path) with the action that is optimal given a posterior that is equal to the prior. [To be precise, this is so if the game with no possibility of communication has an equilibrium; cf. Seidman (1992) and Van Damme (1987).] There may be other equilibria. In particular, if the preferences of **S** and **R** coincide, there exists a separating equilibrium. Yet the refinements discussed in Section 5, which depend on different signals being more costly for some types of **S** than for others, do not help at all to reduce the set of sequential equilibria. Nothing prevents **R** from interpreting an out-of-equilibrium signal s in exactly the same way as a signal s' that is sent with positive probability. When signalling is costless, this sort of response establishes that the signal s is not "bad" for any type that sends s'. Hence no sequential equilibrium can fail the intuitive criterion (or any more restrictive refinements derived from strategic stability). Put another way, in cheap-talk games signals have no natural meaning at all, so no interpretation of them can be ruled out.

It is possible to refine the equilibrium set for costless signalling games by requiring that **R** believe what **S** says if it is in S's interest to speak the truth. Farrell (1993) allows **S** to invent a new signal, called a neologism, which is credible if and only if there exists a nonempty set J of T such that precisely the types in J prefer the neologism to candidate equilibrium payoffs when **R** responds optimally to the literal meaning of the new signal (that is, "$t \in J$"). Formally, take a sequential equilibrium (σ, α, μ), and let $u^*(t)$ be the equilibrium payoff of type t. $J \subseteq T$ can send a *credible neologism* if and only if $J = \{t : u(t, \alpha(J)) > u^*(t)\}$, where $\alpha(J)$ is **R**'s (assumed unique, for simplicity) optimal response to the prior distribution conditioned on $t \in J$. (When ρ is not nonatomic, things are a bit more complex than this, since some types may be indifferent between their equilibrium payoff and the payoff from the neologism.) If **R** interprets a credible neologism literally, then some types would send the neologism and destroy the candidate equilibrium. Accordingly, Farrell emphasizes sequential equilibria for which no subset of T can send a credible neologism. These *neologism proof* equilibria do not exist in all cheap-talk games, including the game analyzed in Crawford and Sobel (1982) for some specification of preferences.

Rabin (1990) defined credibility without reference to a candidate equilibrium. He assumed that a statement of the form "t is an element of J" is credible if (a) all types in J obtain their best possible payoff when **R** interprets the statement literally, and (b) **R**'s optimal response to "t is an element of J" does not change when he takes into account that certain types outside of J might also make the statement. Roughly, a sequential equilibrium is a credible message equilibrium if no type of S can increase his payoff by sending a credible message. He proves that credible message equilibria exist in all finite, costless signalling games, and that "no communication" need not be a credible message equilibrium in some cases.

Dynamic arguments may force cheap talk to take on meaning in certain situations. Wärneryd (1993) shows that in a subset of cheap-talk games in which the interests of the players coincide, only full communication is evolutionarily stable. Aumann and Hart (1990) and Forges (1990) show that allowing many rounds of communication can enlarge the set of equilibrium outcomes.

8. Concluding remarks

In this chapter, we have stayed close to topics concerned with applications to market signalling. Even this brief introduction to the basic concepts and results from this application has exhausted the allotment of space that was given. In a more complete treatment, it would be natural to discuss at least the following further ideas: (1) In the models discussed, the set of possible signals (or, more generally, contracts) is given exogenously. A natural question to ask is whether there might be some way to identify a broader or even universal class of possible contracts and to look for contracts that are optimal among all those that are possible. The literature on mechanism design and the revelation principle should be consulted; see Chapter 24. It is typical in this literature to give the leading role to the uninformed party, so that this literature is more similar to screening than to signalling, as these terms are used above. When mechanism design is undertaken by an informed party, the process of mechanism design by itself may be a signal. Work here is less well advanced; see Myerson (1985) and Crawford (1985) for pioneering efforts, and see Maskin and Tirole (1990, 1992) for recent work done more in the spirit of noncooperative game theory. (2) We have discussed cases where individuals wish to have information communicated, at least partially. One can easily think of situations in which one party would want to stop or garble information that would otherwise flow to a second party, or even where one party might wish to stop or garble the information that would otherwise flow to himself. For an introduction to this topic, see Tirole (1988). (3) Except for a bare mention of multistage cheaptalk, we have not touched at all on applications where information may be communicated in stages or where two or more parties have private information which they may wish to signal to one another. There are far too many interesting analyses of specific models to single any out for mention – this is a broad area in which unifying theory has not yet taken shape. (4) The act of signalling may itself be a signal. This can cut in (at least) two ways. When the signal is noisy, say, it consists of a physical examination which is subject to measurement error, then the willingness to take the exam may be a superior signal. But then it becomes "no signal" at all. Secondly, if the signal is costly and develops in stages, then once one party begins to send the signal, the other may propose a Pareto-superior arrangement where the signal is cut off in midstream. But this then can destroy the signal's separating characteristics. For some analyses of these points, see Hillas (1987).

All this is only for the case where the information being signalled is information about some exogenously specified type, and it does not come close to exhausting that topic. The general notion of signalling also applies to signalling past actions or, as a means of equilibrium selection, in signalling future intentions. We cannot begin even to recount the many developments along these lines, and so we end with a warning to the reader that the range of application of non-cooperative game theory in the general direction of signallng is both immense and rich.

References

Admati, A. and M. Perry (1987) 'Strategic delay in bargaining', *Review of Economic Studies*, **54**: 345–364.

Aumann, R. and S. Hart (1990) 'Cheap talk and incomplete information', private communication.

Banks, J.S., and J. Sobel (1987) 'Equilibrium selection in signaling games', *Econometrica*, **55**: 647–662.

Cho, I-k. (1987) 'Equilibrium analysis of entry deterrence: A re-examination', University of Chicago, mimeo.

Cho, I-k. and D.M. Kreps (1987) 'Signaling games and stable equilibria', *Quarterly Journal of Economics*, **102**: 179–221.

Cho, I-k. and J. Sobel (1990) 'Strategic stability and uniqueness in signaling games', *Journal of Economic Theory*, **50**: 381–413.

Crawford, V. (1985) 'Efficient and durable decision rules: A reformulation', *Econometrica*, **53**: 817–837.

Crawford, V. and J. Sobel (1982) 'Strategic information transmission', *Econometrica*, **50**: 1431–1451.

Dasgupta, P. and E. Maskin (1986) 'The existence of equilibrium in discontinuous economic games, I: Theory', *Review of Economic Studies*, **53**: 1–26.

Engers, M. (1987) 'Signaling with many signals', *Econometrica*, **55**: 663–674.

Engers, M. and L. Fernandez (1987) 'Market equilibrium with hidden knowledge and self-selection', *Econometrica*, **55**: 425–440.

Farrell, J. (1993) 'Meaning and credibility in cheap-talk games', *Games and Economic Behavior*, **5**: 514–531.

Forges, F. (1990) 'Equilibria with communication in a job market example', *Quarterly Journal of Economics*, **105**: 375–98.

Grafen, A. (1990) 'Biological signals as handicaps,' *Journal of Theoretical Biology*, **144(4)**: 517–46.

Green, J. and N. Stokey (1980) 'A two-person game of information transmission', Harvard University mimeo.

Grossman, S. (1981), 'The role of warranties and private disclosure about product quality', *Journal of Law and Economics*, **24**: 461–83.

Grossman, S. and M. Perry (1987) 'Perfect sequential equilibrium', *Journal of Economic Theory*, **39**: 97–119.

Hellwig, M. (1986) 'Some recent developments in the theory of competition in markets with adverse selection', University of Bonn, mimeo.
Hillas, J. (1987) 'Contributions to the Theory of Market Screening', Ph.D. Dissertation, Stanford University.
Kohlberg, E. and J-F. Mertens (1986) 'On the strategic stability of equilibria', *Econometrica*, **54**: 1003–1038.
Kohlleppel, L. (1983) 'Multidimensional market signalling', University of Bonn, mimeo.
Mailath, G. (1987) 'Incentive compatibility in signaling games with a continuum of types', *Econometrica*, **55**: 1349–1365.
Mailath, G., M. Okuno-Fujiwara and A. Postlewaite (1993) 'On belief based refinements in signaling games', *Journal of Economic Theory*, **60**: 241–276.
Maskin, E. and J. Tirole (1990) 'The principal-agent relationship with an informed principal: The case of private values', *Econometrica*, **58**: 379–409.
Maskin, E. and J. Tirole (1990) 'The principal-agent relationship with an informed principal: The case of common values', *Econometrica*, **60**: 1–42.
Milgrom, P. and J. Roberts (1982) 'Limit pricing and entry under incomplete information: An equilibrium analysis', *Econometrica*, **50**: 443–459.
Mirrlees, J.A. (1971) 'An exploration of the theory of optimum income taxation', *Review of Economic Studies*, **28**: 195–208.
Myerson, R. (1985) 'Mechanism design by an informed principal', *Econometrica*, **47**: 61–73.
Quinzii, M. and J.-C. Rochet (1985), 'Multidimensional signalling', *Journal of Mathematical Economics*, **14**: 261–284.
Rabin, M. (1990) 'Communication between rational agents', *Journal of Economic Theory*, **51**: 144–70.
Ramey, G. (1988) 'Intuitive signaling equilibria with multiple signals and a continuum of types', University of California at San Diego, mimeo.
Reinganum, J. and L. Wilde (1986) 'Settlement, litigation, and the allocation of litigation costs', *Rand Journal of Economics*, **17**: 557–566.
Riley, J. (1979) 'Informational equilibrium', *Econometrica*, **47**: 331–359.
Rothschild, M. and J.E. Stiglitz (1976) 'Equilibrium in competitive insurance markets: An essay on the economics of imperfect information', *Quarterly Journal of Economics*, **80**: 629–649.
Rubinstein, A. (1985), 'A bargaining model with incomplete information about time preferences', *Econometrica*, **53**: 1151–1172.
Seidman, D. (1992) 'Cheap talk games may have unique, informative equilibrium outcomes,' *Games and Economic Behavior*, **4**: 422–425.
Sobel, J. (1989) 'An analysis of discovery rules', *Law and Contemporary Problems*, **52**: 133–59.
Spence, A.M. (1974) *Market Signaling*. Cambridge, MA: Harvard University Press.
Stiglitz, J. (1977) 'Monopoly, non-linear pricing and imperfect information: The insurance market', *Review of Economic Studies*, **94**: 407–430.
Stiglitz, J. and A. Weiss (1990) 'Sorting out the differences between screening and signaling models', in: M. Bachrach et al., eds., *Oxford Mathematical Economics Seminar*, Twenty-fifth Anniversary Volume. Oxford: Oxford University Press.
Tirole, J. (1988) *The Theory of Industrial Organization*. Cambridge, MA: MIT Press.
Wärneyrd, K. (1993) 'Cheap talk, coordination, and evolutionary stability', *Games and Economic Behavior*, **5**: 532–546.
Wilson, C. (1977) 'A model of insurance markets with incomplete information', *Journal of Economic Theory*, **16**: 167–207.
Van Damme, E. (1987) 'Equilibria in noncooperative games', in: H.J.M. Peters and O.J. Vrieze, eds., *Surveys in Game Theory and Related Topics*. Amsterdam: Centre for Mathematics and Computer Science, pp. 1–35.
Zahavi, A. (1975) 'Mate selection – a selection for a handicap', *Journal of Theoretical Biology*, **53**: 205–14.
Zahavi, A. (1977) 'Reliability in communication systems and the evolution of altruism', in: B. Stonehouse and C. Perrins, eds., *Evolutionary Ecology*, London: Macmillan.