# Endogenous private information structures

## Sjaak Hurkens[a],*, Nir Vulkan[b]

[a]*Department of Economics and Business and CREA, Universitat Pompeu Fabra, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain*
[b]*The Saïd Business School, Oxford University, Park End Street, Oxford OX1 1HP, UK*

## Abstract

We formally incorporate the option to gather information into a game and thus endogenize the information structure. We ask whether models with exogenous information structures are robust with respect to this endogenization. Any Nash equilibrium of the game with information acquisition induces a Nash equilibrium in the corresponding game with an exogenous structure. We provide sufficient conditions on the structure of the game for which this remains true when 'Nash' is replaced by 'sequential'. We characterize the (sequential) Nash equilibria of games with exogenous information structures that can arise as a (sequential) Nash equilibrium of games with endogenous information acquisition.
© 2004 Elsevier B.V. All rights reserved.

## 1. Introduction

Modern economic theory emphasizes the importance of private information in industrial organization, principal–agent relationships, bargaining, contract theory, auctions and financial markets. In those models the private information of players is

---

*Corresponding author. Tel.: +34 935 422 556; fax: +34 935 421 746.
*E-mail address:* sjaak.hurkens@upf.edu (S. Hurkens).

an *exogenous* part of the model. Such models explain what will happen *given* a certain private information structure, but do not explain where that information structure *comes from*.[1] For some economic situations assuming an exogenous private information structure may be more sensible than for others. For example, in Akerlof's famous Lemons market it seems to make good sense to assume that owners of a used car have private information about the quality of their car, since they are supposed to have used this car in past years. On the other hand, in a principal–agent model of procurement it is less compelling to assume that the agent has private information about his cost as this cost would depend on the requirements of the project imposed by the principal. Hence, the agent will have to process those requirements and make estimates of input costs, i.e. the agent will have to acquire information about his cost. The private information structure of the agent would thus be endogenous. Similar stories apply for bidders in an auction, firms and entrants in imperfect competition models, and speculators in stock markets.

Now, continuing with the procurement example, one might argue that the information about cost can be much easier acquired by the agent than by the principal, and that it is obviously in the interest of the agent to do so at the (relative) small effort it takes him. This argument would then seem to justify the use of the model with an exogenously given privately informed agent. Similar arguments would of course then apply in similar situations.

However, some recent papers have shown that endogenizing the information structure leads to surprising but not necessarily unintuitive results and, more importantly, to new insights in varying fields in economic theory. In the case of principal–agent theory, Kessler (1998) has shown that the agent may prefer to stay uninformed, thereby crucially changing the incentive compatibility and individual rationality constraints of the principal's problem. Cremer and Khalil (1992) show that the principal may offer contracts that induce the agent to stay uninformed and that those contracts will be very different from those that a principal would offer in case the uninformed agent would not have the possibility to gather information. In the context of auctions Persico (2000) shows that the incentives for bidders to gather information are different in first- and second-price auctions so that the revenue equivalence theorem breaks down. Finally, Hurkens and Vulkan (2003) show that the number of entrants in a new market with uncertain demand when firms' private information about demand is exogenously given need not coincide with the number of entrants when firms have to decide whether or not to become informed. Surprisingly, this holds even though the private information structures that arise in the endogenous model are identical to the ones assumed in the exogenous one.

---

[1] Notable exceptions are Vives (1988), Li et al. (1987), Hwang (1993, 1995), Hauk and Hurkens (2001), and Ponssard (1979) that consider information acquisition in oligopolies, Lee (1982), Milgrom (1981), Matthews (1984), and Persico (1997) who consider the incentives for information acquisition in auctions and Cremer and Khalil (1992,1994), Cremer et al. (1998), and Kessler (1998) that consider principal–agent relationships where the agent is initially uninformed but can acquire information, before or after the principal has offered him a contract. Grossman and Stiglitz (1980) initiated research on the incentives to acquire costly information on the value of a stock in financial markets. Hurkens and Vulkan (2001, 2003) considered information acquisition by potential entrants.

It is not unlikely that this list of examples can be extended with applications from other fields. That is however not the main purpose of this paper. The aim of this paper is to investigate, at an abstract and general level, which type of games are robust and immune with respect to the endogenization of the private information structure and which type of games are susceptible to this endogenization. Games from the first class can thus be analysed using an exogenous private information structure even when they in fact describe economic situations where information gathering seems a natural part of players' strategic options. Games that belong to the second class, however, will have to be considered with care as they are possibly affected by having endogenous rather than exogenous information structures.

Our investigation of robustness starts out of two important questions. First, can any exogenous private information structure arise endogenously? Second, does behaviour in the game depend on whether the private information structure is exogenously given or is endogenously determined? In fact, if the answers to these questions were 'yes' and 'no' respectively, we would be fully justified to use models with exogenous private information structures. From our examples discussed before we know that in fact the answers will not always be 'yes' and 'no', and we need to investigate in which cases the answers may differ.

At first sight the answer to the first question would seem to be 'yes'. Namely, it suffices to choose the costs for different information structures such that the one information structure we want to 'explain' has the highest value, *net of costs*. However, one is not completely free in the choice of costs: any reasonable information cost function should be (weakly) increasing, i.e. more information should cost more. A more fundamental problem is that the notion of 'value of information' in a game is rather vague. (See also Neyman (1991).) How much a player can improve his payoff by obtaining some piece of information will depend on how the other players in the game will behave and respond. Hence, the value of information in a game (if we can even speak about such a concept) must be closely linked to (equilibrium) behaviour in the game. Clearly, existence of multiple equilibria and issues of equilibrium refinements will affect and possibly complicate our task. As a matter of fact, this issue will turn out to be very important.

The answer to the second question would seem to be 'no', especially since we will be interested in equilibrium behaviour. In equilibrium of the game with endogenous information acquisition each player 'knows' how much information other players acquire, similar to the case of games with exogenous private information structures. There is a small but important difference though. Namely, a player will have beliefs about how much information others acquired, and in equilibrium these beliefs must be correct. However, *off* the equilibrium path a player may hold incorrect beliefs which may affect his behaviour *off* the equilibrium path. Of course, behaviour *off* the equilibrium path will in fact determine the behaviour *on* the equilibrium path. It is therefore not obvious at all that the answer to this second question is always 'no'. In fact, the entry example discussed before constitutes a counterexample. Since beliefs seem to play a role here the issue of equilibrium refinements (such as sequential equilibrium) again turns up.

We conclude that we cannot answer the above two questions without fully analysing and comparing equilibrium behaviour in games with exogenous and endogenous information structures. That is what we will do in the paper. We will consider equilibrium behaviour in the game with endogenous information acquisition and compare that to equilibrium behaviour in the game with an exogenous information structure, where the information structure is in fact the one that endogenously arises in the first game. In the other direction, we will consider equilibrium behaviour in a game with an exogenously given private information structure and we ask ourselves whether this behaviour and this information structure can prevail in some equilibrium of the endogenous information acquisition game. We also investigate the importance of equilibrium refinements.

Let us now outline the results obtained in this paper. We will focus on the class of games where information about fundamentals (i.e. states of Nature that affect the payoffs but not the strategic options of the players) is acquired before players enter into a strategic situation. Hence, there is an information gathering stage followed by a game playing phase.[2] Information acquisition is modelled as a choice between different partitions of the set of states of Nature. Information costs are exogenously specified but may differ amongst players to reflect the asymmetry in the underlying game. Since the actions of a player are chosen after information acquisition decisions are taken, they cannot influence those decisions. To each (pure or mixed) profile of choices of partitions corresponds a game with an exogenous information structure.

Our first result says that any equilibrium of the endogenous game induces an equilibrium in the corresponding exogenous game. The result is simple and intuitive but also important. Namely, it follows that games with exogenous information structures that have a unique Nash equilibrium are quite robust to endogenization of the information in the sense that behaviour in the game does not depend on whether the information structure is exogenously given or endogenously determined. (Of course, as seen in the principal–agent examples, it might be true that the given information structure could not have arisen endogenously for any reasonable specification of information costs.) We then consider equilibrium refinements, and in particular we consider sequential equilibria. We show that any sequential equilibrium of the endogenous information acquisition game induces a sequential equilibrium in the corresponding game with exogenous structure if (i) the information acquisition decisions are perfectly observed, or (ii) the game playing phase only involves simultaneous moves. A counterexample illustrates that if neither of these conditions is satisfied a sequential equilibrium of the endogenous information acquisition game need not induce a sequential equilibrium in the exogenous game. A consequence of this is that games with exogenous information structures that have a unique (plausible) sequential equilibrium but multiple Nash equilibria may be affected by the endogenization of information.

---

[2]This excludes the case analysed by Cremer and Khalil (1992) where the principal offers a contract before the agent may gather information. It also excludes the case where a player can get information about *unobserved actions* chosen by other players. See Perea y Monsuwé (1997) for a setting in which players buy information about actions chosen by other players.

We then turn our attention to equilibrium behaviour in games with exogenous information structures and ask ourselves whether this behaviour and information structure can also arise in the endogenous information acquisition game for some reasonable assumptions on information costs. If we do not exclude that information is available for free then any Nash equilibrium of a game with an exogenous private information structure may arise as a Nash equilibrium of the game with endogenous information acquisition. When information acquisition is unobserved and information is not free of cost then this holds in case information is used in the equilibrium of the exogenous game under consideration. Hence, pooling in signalling games or bunch-bidding in auctions are excluded in this case. On the other hand, when information acquisition is observed and we restrict attention to sequential equilibria then the result holds if information does not 'hurt', that is, if no player is better off (in any equilibrium) with less information.

The rest of the paper is set up as follows. Section 2 models private information and defines two modes of information acquisition, *unobserved* and *observed*. We start the analysis in Section 3 by means of an example to illustrate the difficulties that may arise in comparing endogenous and exogenous information structures. We then consider equilibria of the endogenous games and analyse the behaviour it induces in the corresponding game with an exogenous information structure. The final part of this section goes in the other direction, that is, we consider an equilibrium of some exogenous game and investigate when the same information structure and behaviour may arise when information acquisition is endogenous. Section 4 summarizes the results and hints at some possible applications for future research.

## 2. The model

In this paper, we want to compare behaviour in games where players are exogenously endowed with information about fundamentals of the game with that in games where players can only become informed about fundamentals if they decide to (costly) search for this type of information. Obviously, in both types of games players may have (at some point in the game) information about fundamentals (to which we will refer as "information structure") as well as information about what other players have done up to this point in the game (to which we will refer as "action information sets"). This will be indicated (as usual) by means of information sets. In order to compare the two types of games it will be helpful to break up the situation in three parts: In the first part Nature determines the values of the fundamentals. In the second part the information of players about fundamentals (either exogenously or endogenously) is determined. Finally, in the third part the players interact strategically. We will start with the last part since it will be easier to describe the information sets in the games with endogenous and exogenous information about fundamentals when we already know how the action information sets look like.

There are $n$ players involved in a strategic interaction represented by a game form $G$, which can be thought of as a tree where nodes represent players who have to

choose actions and where arcs represent the possible actions. This game form describes the actions that players may take, and the order in which they take them. It includes "action information sets", indicating that players may not know which actions have been chosen before by other players. (We will assume perfect recall so that every player recalls the choices it made before as well as all the information it had about what other players did before.) It is not, however, a fully defined extensive form game because there are no payoffs associated to the set of end points of the tree, $Z$. Let $K^i$ denote the set of nodes at which player $i$ must take an action and let $H^i$ denote the partition[3] of $K^i$ into sets of nodes that cannot be distinguished by the player. Each $h \in H^i$ is an action information set of player $i$. A pure (game form) strategy for player $i$ is a mapping $s_i : H^i \to A$ that assigns to each action information set $h \in H^i$ one of the actions available. We will denote by $A(h)$ the set of actions available at $h$ and $A$ denotes the set of all actions. Alternatively, we can say that a pure strategy for a player is a mapping that assigns to each of his decision nodes an available action, with the restriction that this mapping is measurable with respect to the action information set structure. We denote by $S_i$ the finite set of pure (game form) strategies of player $i$ in $G$ (which does not depend on the state of Nature) and write $S = \prod_i S_i$ as the product set of strategy profiles.

The payoff functions to evaluate outcomes depend on the realized value of the state of Nature. We denote by $\Omega = \{\omega_1, \ldots, \omega_N\}$ the finite set of states of Nature and will write $\omega$ to denote a generic state. We denote by $u_i^\omega(z)$ the payoff obtained by player $i$ when end point $z$ is reached and the state of Nature is $\omega$. If the strategy profile $s$ induces end point $z$ to be reached we will also write $u_i^\omega(s)$ instead of $u_i^\omega(z)$. The players initially hold a common prior about the likelihood of each state of Nature which is represented by a probability distribution $\rho$ on $\Omega$. However, players can either endogenously improve their information or will be exogenously endowed with better information.

A pure information structure of player $i$ is given by some partition $P^i$ of $\Omega$. We write $P^i(\omega)$ for the element of $P^i$ that contains $\omega$. The interpretation of the partition is that with information structure $P^i$ player $i$ can distinguish between two states $\omega$ and $\omega'$ if and only if $P^i(\omega) \neq P^i(\omega')$. We say that partition $P^i$ is finer (more informative) than partition $\tilde{P}^i$ if $P^i(\omega) \subset \tilde{P}^i(\omega)$ for all $\omega \in \Omega$ (with strict inclusion for some $\omega$). In the game with endogenous information acquisition each player $i$ will have to choose among a set of available pure information structures, which we will denote by $\wp_i$. Not all pure information structures need to be feasible but we assume that $P^{no} = \{\Omega\} \in \wp_i$, that is, players can always choose not to become informed at all. Since we do not want to rule out that players randomize between several pure information structures we should not impose that in games with exogenous information structures, each information structure is pure, i.e. a partition of $\Omega$. We will therefore allow for random information structures, represented by moves of Nature: $\tilde{\rho}^i$ is the probability distribution over $\wp_i$. We impose that the probability mechanisms for assigning pure private information structures to the players are uncorrelated amongst themselves and with the state of Nature. (Because in the

---

[3] $X = \{X_1, \ldots, X_k\}$ is a partition of $Y$ if $\cup_j X_j = Y$ and $X_j \cap X_l = \emptyset$ if and only if $j \neq l$.

endogenous information acquisition game we will assume that players acquire information independently and we assume that players start off with no information about Nature.)

Each player $i$ will either be endogenously or exogenously endowed with some information structure $\tilde{\rho}^i$. Let $\tilde{\rho} = (\tilde{\rho}^1, \ldots, \tilde{\rho}^n)$. We will denote the game with the (random) exogenous information structure $\tilde{\rho}$ by $G^{\tilde{\rho}}$. (In this game each player $i$ will learn the outcome of the randomization $\tilde{\rho}^i$.) Note that in the case of a pure information structure the probability distribution $\tilde{\rho}$ has a unit mass at a pure information structure $P = (P^1, \ldots, P^n)$. In this case we denote the game with the exogenous information structure by $G^P$. Both $G^{\tilde{\rho}}$ and $G^P$ are standard games with private information that can be represented by an extensive form. We will first describe the extensive form of $G^P$. Here first Nature chooses one of the $N$ possible states in $\Omega$ after which a "copy" of $G$ is attached (together with the corresponding payoffs at the end points). It will be useful to label the copy of $G$ at $\omega$ by $G(\omega)$. Similarly, for each action information set $h$ in $G$ we will denote by $h(\omega)$ the corresponding action information set in $G(\omega)$. Such an action information set would be a real information set for player $i$ in case this player has perfect information about the state of Nature, i.e. his information structure is the finest partition of $\Omega$. However, in general, if his information structure is $P^i$ then the true information sets that correspond to the action information set $h$ are the unions over all indistinguishable states of Nature, i.e. for each $P^i_k \in P^i$ there is one information set in $G^P$ that corresponds to the action information set $h$ of $G$, and it is $\cup_{\omega \in P^i_k} h(\omega)$. We will denote this information set by $P^i_k(h)$.

In case of a random information structure $\tilde{\rho}^i$ things are slightly more complicated as this cannot be represented as before with Nature choosing a state, attaching a copy of $G$ after each state and uniting action information sets appropriately into an information set. In this case we need Nature also to determine the pure information structure for each player (according to the probabilities given by $\tilde{\rho}^i$). Now we have to attach a copy of $G$ after each outcome of the moves of Nature, that is, for each combination of state of Nature $\omega$ and pure information structure $P \in \wp = \wp_1 \times \cdots \times \wp_n$. We will denote the corresponding copy of $G$ by $G(\omega, P)$ and the corresponding copy of $h$ by $h(\omega, P)$. The information set for player $i$ that corresponds to an action information set $h$ and each $P^i_k \in P^i$ is now $\cup_{P^{-i}} \cup_{\omega \in P^i_k} h(\omega, (P^i, P^{-i}))$, where the union over all $P^{-i}$ represents the fact that player $i$ does not know the exact information structure of other players.[4] Observe that it is possible for a player to have an information structure that is a randomization between two pure information structures $P^i$ and $\tilde{P}^i$ where some set $X \subset \Omega$ is an element of both partitions. Our

---

[4] The case of a pure information structure $P$ can also be modelled as described here, with the restriction that all other feasible information structures have probability zero (that is, $\rho(P') = 0$ for all $P' \neq P$). When we now delete all moves of Nature that have zero probability (as is usually done) we get again our first description of $G^P$. However, it may be useful to leave the zero probability moves in, because of its similarity with the case of endogenous information acquisition where players can choose between all feasible information structures. In particular, the information sets of the game with exogenous information structures then coincide with the information sets in the case of endogenous unobserved information acquisition.

modelling choice implies that we assume that this player can distinguish between the case where his information structure is $P^i$ and the one where it is $\tilde{P}^i$ when the true state is in $X$. This is very natural in the case of endogenous information acquisition, when it is the player himself who randomizes. In the case of exogenous information structures it seems more natural to assume the opposite and merge the two information sets into one. In terms of strategies it makes a difference which assumption one adopts since in the first case a player could choose different actions in the two information sets. However, in terms of *equilibrium* strategies the difference is innocuous: if it is optimal for a player to randomize between choosing the information structure $P^i$ and to play action $a$ in case the true state of Nature is in $X$ and choosing the information structure $\tilde{P}^i$ and to play action $a'$ in case the true state of Nature is in $X$, then it must be also optimal to randomize between these information structures and to randomize between the actions $a$ and $a'$ in case the true state of Nature is in $X$.

We will compare the equilibria of $G^{\tilde{p}}$ and $G^P$ with those of the super-game where information structures are endogenously determined. This super-game has two stages. In the first stage all players choose simultaneously one of the feasible information structures while in the second stage $G$ is played. Without loss of generality we assume that the extensive form of the endogenous information acquisition games start with the move of Nature and then players choose their information structure in order of increasing player index (although no player is informed about previous moves). This means that in this phase of the game each player has one information set. Information acquisition is costly and costs are player specific. Player $i$ needs to pay $c(i, P)$ to acquire partition $P$, where $c(i, P) \geqslant c(i, P')$ if $P$ is more informative than $P'$. When information is not freely available the inequalities will be strict. For convenience, we assume $c(i, P^{\text{no}}) = 0$.

Information acquisition can either be *observed* or *unobserved*. In the super-game where information acquisition is unobserved, denoted by $\Gamma^{\text{unobs}}$, the choice of partition $P^i$ by player $i$ is not observed by players $j \neq i$. This means that information sets (in the second stage) are as in the case of $G^{\tilde{p}}$: The information set for player $i$ that corresponds to an action information set $h$ and each $P^i_k \in P^i$ is $\cup_{P^{-i}} \cup_{\omega \in P^i_k} h(\omega, (P^i, P^{-i}))$. A pure super-game strategy for player $i$ is thus a pair $(P^i, \sigma_i)$, where $P^i$ is the partition chosen in the first stage and $\sigma_i : \Omega \to S_i$ maps states of Nature into strategies of $G$. Of course, the mapping $\sigma_i$ must be measurable with respect to $P^i$, i.e. $\sigma(\omega) = \sigma(\omega')$ if $P^i(\omega) = P^i(\omega')$. On the other hand, in the super-game where information acquisition is observed, denoted by $\Gamma^{\text{obs}}$, the choice of partition $P^i$ by player $i$ is observed by all players. In case players use mixed strategies, we assume that the outcome of the realization is observed, and not the mixture itself. Hence, a pure super-game strategy for player $i$ is now a pair $(P^i, \tau_i)$, where $P^i$ is again the partition chosen by $i$ and where $\tau_i(\omega, P) \in S_i$ denotes the strategy chosen in $G$ when the state of Nature is $\omega$ and the (total) information structure is $P$. Again, the mapping $\tau_i$ must be measurable with respect to $P^i$.

We will be interested in comparing the results of the endogenous information acquisition games $\Gamma^{\text{unobs}}$ and $\Gamma^{\text{obs}}$ with those obtained in games with an exogenous information structure, $G^{\tilde{p}}$. Since the strategy spaces in those three types of games are

different, we need to make precise how we will compare the equilibrium results. Because of the sequential structure of the information acquisition games we will sometimes want to restrict attention to sequential equilibria so that our results do not rely on incredible out of equilibrium threats or beliefs. Recall that a sequential equilibrium is a pair $(\mu, \sigma)$ where strategy $\sigma$ is optimal given the beliefs $\mu$, and the beliefs are consistent, i.e., $(\mu, \sigma) = \lim_n(\mu_n, \sigma_n)$, where $\sigma_n$ is a completely mixed strategy profile and $\mu_n$ are beliefs determined by Bayes' rule based on $\sigma_n$. Finally, we will compare the strategies induced in the second stage by the (sequential) equilibria of the information acquisition games with the (sequential) equilibria of the games with an exogenous information structure, where this information structure is the one that endogenously emerged in the equilibrium at hand. Note that in the case of observed information acquisition the strategy of each player is contingent on the actual pure information structure that arises.

## 3. Results

### 3.1. Comparing endogenous and exogenous information structures

Before we come to our formal results, let us consider an example to illustrate how comparisons are made. The example also illustrates the importance of allowing for random information structures.

Suppose that Nature determines which of the two bimatrix games of Fig. 1 is going to be played, I or II. Game I is picked with probability 1/2.

Suppose that only player 1 can unobservably learn whether game I or II is chosen, at some cost $c \in (0, 1)$. (We could assume that player 2 can also learn this but that his cost is larger than 5, which would make it a strictly dominated strategy for him to learn the outcome of the move of Nature.) Note that learning and not using the information is a strategy strictly dominated by not learning and playing B. Also, learning and playing MT (i.e. M in game I, T in game II), MB, TB, BM, or BT are dominated. (For example, (learn, play TB) is strictly dominated by the mixed strategy that assigns equal weight to (not learn, play B) and (learn, play TM).) The only undominated strategies for player 1 are (not learn, play B) and (learn, play TM). Of course, if player 1 plays B, player 2 prefers to choose R, against which player 1 prefers TM. But if player 1 plays (learn, TM), player 2 prefers to choose L,

|   | L | R |   | L | R |
|---|---|---|---|---|---|
| T | 3,1 | 6,0 | T | 0,0 | 0,1 |
| M | 0,0 | 0,1 | M | 3,1 | 6,0 |
| B | 4,4 | 5,5 | B | 4,4 | 5,5 |
|   | I |   |   | II |   |

Fig. 1.

against which it is optimal for player 1 to choose (not learn, B). Hence, the game with unobserved information acquisition has no pure Nash equilibrium and there is a unique mixed Nash equilibrium: player 1 randomizes equally between (learn, TM) and (not learn, B), while player 2 chooses L with probability $(1 - c)/2$.

Note that the information structure that endogenously arises is mixed since player 1 can be informed or uninformed.[5] This can be interpreted as a game in which three types of player 1 exist, type I, type II, and a third uninformed type with exogenous probabilities $1/4, 1/4$, and $1/2$, respectively. Note that these probabilities do not depend on $c$ (as long as $c \in (0, 1)$), but on the payoffs of player 2. In the game with three types a unique equilibrium component exists: player 1 chooses T (if type I), M (if type II) and B (if type III), while player 2 plays L with probability at most 1/2. We see that the Nash equilibrium of the game with unobserved information acquisition induces a Nash equilibrium in the game with the exogenous (mixed) information structure and that each of the Nash equilibria of the game with the exogenous information structure can be "explained" by making the adequate assumption on the cost of information acquisition. Also note that the model where player 1 is informed with probability $p \neq \frac{1}{2}$ cannot be generated from endogenous and unobserved information acquisition, as long as information cost is positive but small.[6]

## 3.2. Endogenous information acquisition equilibria

In this subsection, we consider equilibria of games with endogenous information acquisition. In particular, we investigate the behaviour induced by such equilibria in the game playing phase. We will consider both unobserved and observed information acquisition games and will also discuss the role of equilibrium refinements. We start with ordinary Nash equilibria.

**Theorem 1.** (i) *Let $(\tilde{\rho}, \sigma)$ be a Nash equilibrium of the unobserved information acquisition game $\Gamma^{\text{unobs}}$. Then $\sigma$ is a Nash equilibrium of $G^{\tilde{p}}$.*

(ii) *Let $(\tilde{\rho}, \tau)$ be a Nash equilibrium of the observed information acquisition game $\Gamma^{\text{obs}}$. Then $\tau(P)$ is a Nash equilibrium of $G^P$ for all $P$ with $\tilde{\rho}(P) > 0$.*

**Proof.** (i) Suppose not. Then some player strictly prefers to deviate from $\sigma$ in $G^{\tilde{p}}$. But then this same player would also like to deviate in the game playing phase of $\Gamma^{\text{unobs}}$, contradicting the assumption that $(\tilde{\rho}, \sigma)$ is a Nash equilibrium of $\Gamma^{\text{unobs}}$.

(ii) Suppose not. Then some player strictly prefers to deviate from $\tau(P)$ in $G^P$ for some $P$ with $\tilde{\rho}(P) > 0$. But then this same player would also like to deviate in the game playing phase of $\Gamma^{\text{obs}}$ contradicting the assumption that $(\tilde{\rho}, \tau)$ is a Nash equilibrium of $\Gamma^{\text{obs}}$.   □

---

[5]This result is somewhat reminiscent of the principal–agent models by Kessler (1998) and Cremer and Khalil (1992) where it is shown that the agent will (with positive probability) be uninformed about his cost.

[6]For values of $c > 1$ information acquisition is dominated and we would have support for the model with an uninformed player 1, whereas if information is costless ($c = 0$) it is an equilibrium of the unobserved information acquisition game to acquire information and then use the information with probability a half, while player 2 uses both strategies with equal probability.

The result is straightforward but has important consequences. Namely, it follows that a game with an exogenous information structure that has a unique Nash equilibrium is at least partly robust with respect to the endogenization of the information structure in the sense that behaviour in the game playing phase does not depend on whether the information structure is exogenous or endogenous. What of course remains to be seen is whether that information structure would arise endogenously. We will come back to this issue in the next subsection.

The reason that the result is straightforward is that we consider any Nash equilibrium. That is, we only ask whether each player's action is optimal given other players' actions. Actions off the equilibrium path of play are always optimal, while actions on the equilibrium path of play are, by definition, optimal. Of course, some actions off the equilibrium path may seem unreasonable or incredible. Hence, it is natural to consider subgame perfect or, more generally, sequential equilibria.[7] One might guess that Theorem 1 would break down when replacing 'Nash' by 'sequential' as players will make inferences about other players' informedness in the game with endogenous information acquisition whereas they do not in a game with exogenous information structure. It could be that in the former game some action off the proposed path of play can be 'rationalized' by some beliefs about informedness while no such rationalization exists in a game where the information structure is given and common knowledge. As Theorem 2 below shows, it turns out that this argument can only apply in the case of games with sequential moves and when information acquisition is not observed. Namely, in the case of observed information acquisition the information structure will become common knowledge even though it arises endogenously. Furthermore, in the case of unobserved information acquisition beliefs about informedness only play a role when actions are taken after beliefs have been revised. In the case of simultaneous move games belief revision can thus not play any role!

**Theorem 2.** (i) *Let $(\tilde{\rho}, \sigma)$ be part of a sequential equilibrium of the unobserved information acquisition game $\Gamma^{\mathrm{unobs}}$. Then $\sigma$ is part of a sequential equilibrium of $G^{\tilde{\rho}}$ if $G$ is a simultaneous move game.*

(ii) *Let $(\tilde{\rho}, \tau)$ be part of a sequential equilibrium of the observed information acquisition game $\Gamma^{\mathrm{obs}}$. Then $\tau(P)$ is part of a sequential equilibrium of $G^P$ for all $P \in \wp$.*

**Proof.** (i) By Theorem 1(i) we know that $\sigma$ is a Nash equilibrium of $G^{\tilde{\rho}}$. Since $G^{\tilde{\rho}}$ has only simultaneous moves it is sequential as well (where beliefs are determined by equilibrium strategies and moves of nature).

(ii) Each information set in the game playing phase $\Gamma^{\mathrm{obs}}$ coincides with one information set in $G^P$ for exactly one information structure $P$. We can therefore use exactly the same beliefs that support the sequential equilibrium of $\Gamma^{\mathrm{obs}}$ to support $\tau(P)$ as a sequential equilibrium of $G^P$: $\tau(P)$ is optimal given these beliefs and the beliefs are the limit of a sequence of beliefs that are generated from completely mixed strategies in $G^P$. $\quad\square$

---

[7]The game with information acquisition may have no or few proper subgames as Nature moves first.

In order to see that Theorem 2(i) is, in general, not true for sequential games consider the following situation. Nature chooses between game I and game II with equal probability. Only player one has the possibility to become informed about Nature's move (Fig. 2).

We will show that the unobserved information acquisition game has a unique Nash equilibrium outcome (and therefore also a unique sequential equilibrium outcome) where player 1 does not learn and chooses out. By Theorem 1(i) we know that this must also be a Nash equilibrium outcome in the game where it is common knowledge that player 1 is not informed. However, the only sequential equilibrium in that game is where players enter the subgame and obtain $(5, 5)$. Hence, the sequential equilibrium of the unobserved information acquisition game does not induce a sequential equilibrium in the corresponding game. As we will see shortly, the reason is that in order to support the sequential equilibrium of $\Gamma^{\text{unobs}}$ player 2 should believe that player 1 is informed when player 2 (unexpectedly) gets to move.

Let us now analyse the game. Player 1 can decide to learn the outcome of Nature's move at cost $c \in (0, 1/4)$. Then he must decide to end the game (yielding gross payoff vector $(2, 0)$) or give the move to player 2. If given the move, player 2 can decide to end the game (yielding gross payoff vector $(1, 4)$) or continue the game. If the game continues both players engage in a $3 \times 2$ bimatrix game. The payoffs in this game depend on whether Nature has chosen I or II. The structure of the payoffs in this 'subgame' is such that if player 1 is informed he has a strictly dominant strategy: play $t$ if Nature chooses I, choose $m$ if Nature chooses II. Player 2 will then want to play $l$
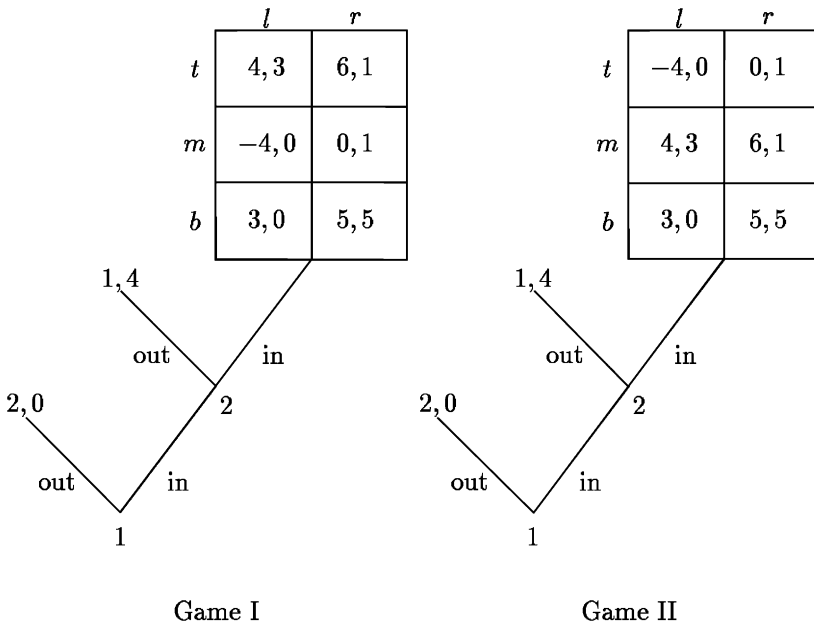


Fig. 2.

|  | out | (in, $l$) | (in, $r$) |
|---|---|---|---|
| (not learn, out, $\star$) | 2 , 0 | 2 , 0 | 2 , 0 |
| (not learn, in, $b$) | 1 , 4 | 3 , 0 | 5 , 5 |
| (learn, in, $tm$) | 1-$c$ , 4 | 4-$c$ , 3 | 6-$c$ , 1 |

Fig. 3.

in this subgame, yielding him a payoff of 3 (which is less than 4 which he can obtain by choosing 'out' if given the move). If player 1 is not informed his strictly dominant strategy (in the subgame) is to play $b$. In that case player 2 will want to play $r$, yielding him a payoff of 5. Also note that it is a strictly dominated strategy for player 1 to learn the move of Nature and then choose 'out', since information acquisition is costly. Any Nash equilibrium of the game with endogenous information acquisition must therefore be a Nash equilibrium of the game restricted to the undominated strategies of Fig. 3.

There is a component of equilibria where player 1 chooses (not learn, out, $\star$) and player 2 chooses out with sufficiently high probability. There are no equilibria in which player 1 uses any other strategy with positive probability, because against such strategies it is not optimal to play (in, $l$) for player 2. Moreover, against the mixture between 'out' and (in, $r$) needed to make player 1 indifferent between his last two strategies it is optimal for player 1 to choose his first strategy. We conclude that in the game with costly and unobserved information acquisition the unique equilibrium outcome is that player 1 does not learn and chooses out with probability 1.

Consider now the game where it is common knowledge that player 1 is uninformed. Of course, also in this game it is a Nash equilibrium for both players to choose 'out'. However, the only sequential equilibrium is of course ((in, $b$), (in, $r$)). Namely, it is a dominant strategy for player 1 to choose $b$ in the final stage so in any sequential equilibrium player 1 must do so. That implies that in any sequential equilibrium player 2 must play $r$ in the final stage and 'in' when he is given the move. But that in turn implies that player 1 must choose 'in' as well.

This example is in the same spirit as our[8] application to entry in a new market. We assume there that information acquisition is unobserved and since firms compete in quantities after having taken entry decisions, that application has a sequential structure that makes revising beliefs possible and important. There exists a sequential equilibrium in which firms become perfectly informed about demand but less entry takes place than in a model where it is common knowledge that firms are perfectly informed. Unexpected entry is believed to be caused by uninformed entrants which makes the 'incumbent' act more aggressively, thereby deterring entry altogether.

---

[8]Hurkens and Vulkan (2003).

|   | L | R |     |   | L | R |
|---|---|---|-----|---|---|---|
| T | 4,2 | 1,3 |   | T | 2,2 | 0,0 |
| B | 2,2 | 0,0 |   | B | 4,2 | 1,3 |
|   | I |   |     |   | II |   |

Fig. 4.

### 3.3. Exogenous information structures

In this subsection we consider games with pure[9] exogenous information structures $G^P$ and their (sequential) equilibria $\sigma$ and ask ourselves whether the same information structure and behaviour could arise with endogenous information acquisition when information is costly. Of course, the answer will depend on which information structures are available to the players and how costly the different information structures are. When many different information structures are assumed to be available then it becomes in principle more difficult to justify the choice of a particular one. Of course, it is easy to rule out the choice of partitions finer than $P^i$ by assuming that these are very costly. We can do the same for partitions that are not coarser than $P^i$. We cannot do so for partitions coarser than $P^i$ because they cannot cost more than $P^i$. We will thus assume throughout that $\{P^{no}, P^i\} \subset \wp_i$ and that only information structures coarser than $P^i$ are possibly available. If we cannot justify the choice under these assumptions, then it will be impossible to do so if more information structures are available.

It will be instructive to consider some examples before coming to the formal results. Nature determines which of the two bimatrix games of Fig. 4 is going to be played, I or II. Game I is picked with probability 2/5. Let us see whether endogenous information acquisition can explain the information structure where only player 1 is informed, i.e. $P^1 = \{\{I\}, \{II\}\}$ and $P^2 = \{\{I, II\}\} = P^{no}$. We assume that $\wp_1 = \{P^{no}, P^1\}$ and that $\wp_2 = \{P^{no}\}$, so that only player 1 has the opportunity to learn the state of Nature.

If neither player knows which game is played, both players have a dominant strategy and the outcome will be (B, L), yielding an expected payoff of (16/5,2). However, if player 1 observes which matrix is chosen, it is a dominant strategy to play T in game I and B in game II. If player 2 knows that player 1 observes the outcome of the move of Nature, he will then play R and the resulting payoff vector will be (1,3). Player 1 is better off in the game where it is common knowledge he does not observe Nature's move than in the game where it is common knowledge that he does. This is an example where information hurts the player possessing it. In the

---

[9]The example of Fig. 1 has shown that equilibrium behaviour in games with random information structures can be explained by referring to endogenous information acquisition only under very specific assumptions about the information costs and for very specific probabilities on the information structures. We therefore restrict ourselves in this section to pure information structures.

unique sequential equilibrium of the game with observed information acquisition player 1 will thus choose (commit) not to obtain information, and the outcome will be (B, L) for any nonnegative cost of acquiring information. If information cost is small, the observed information game possesses a *Nash* equilibrium in which player 1 observes Nature's move and player 2 threatens to play R in any case (that is, whether or not player 1 observes Nature's move or not). We conclude that in this example the information structure where player 1 is informed can arise in a Nash equilibrium of the observed information acquisition game but not in a sequential equilibrium because information hurts.

In the case of unobserved information acquisition, however, player 2 cannot condition his action on whether player 1 learned or not. He must choose the same (mixed) strategy in both games (as he cannot distinguish them). Whatever this strategy of player 2 may be, for low enough information costs $c$ player 1 will choose to acquire information and choose T in game I and B in game II. Player 2 will foresee this in any equilibrium and play R and the final payoffs will be $(1 - c, 3)$.

We can draw some general conclusions from this example. In the case of unobserved information acquisition, players cannot condition their strategies in the game playing phase on the information acquisition decision of any player. Hence, when making the information acquisition decision a player takes the strategies of others in the game playing phase as given and can actually calculate the value of information. This value is always non-negative and if it is strictly positive he will acquire the information for low enough (positive) information costs. Of course, the value of information will be strictly positive only if the player will actually use it. That is, there must be some states of Nature that the player can distinguish only in case he learns and in which he will play distinct strategies resulting in distinct outcomes.

In the case of observed information acquisition, information does not only have an informational value (as in the case of unobserved information acquisition) but it also has some strategic value, as it can affect other players' behaviour. In a sequential equilibrium it will often affect other players' behaviour, while in Nash equilibria other players may (pretend to) ignore the fact that some player has acquired information. As we have seen, however, such ignorance may very well be incredible. We have also seen that the strategic value of obtaining a piece of information may be strictly negative and may more than offset the positive informational value: information may hurt. However, the information observably acquired in a sequential equilibrium never hurts.

We saw that costly information that is unobservably acquired must be used. Hence, equilibria in games with exogenous information structure where private information is not used is not robust to endogenization of the information structure. Examples of such cases are pooling equilibria in signalling games and bunch-bidding in auctions. However, if information is acquired observably it is possible that costly information is acquired but not used.

Consider the example from Fig. 5. Nature chooses with equal probability between game I and II. Again, assume only player 1 can learn Nature's move at some small positive cost $c$. If given the move by player 2, player 1 has to guess
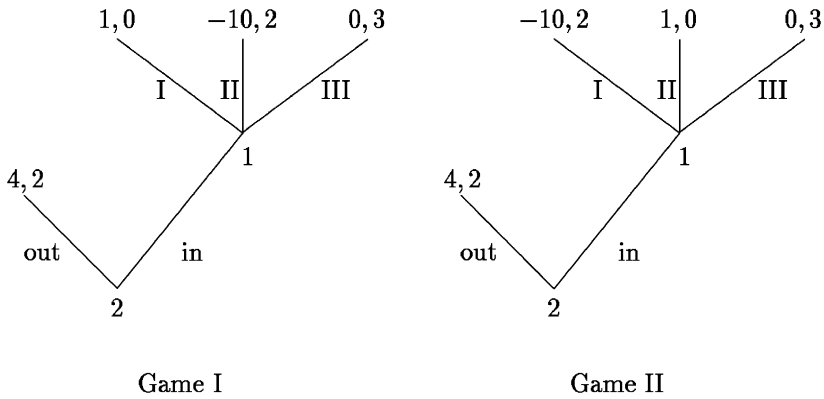
Fig. 5.

Nature's move. If he guesses correctly he is rewarded but if he guesses wrong he is severely punished. In case player 1 does not know it is optimal to admit that and choose option III.

It is clear that the unique sequential equilibrium of the observed information acquisition game is for player 1 to learn Nature's move and guess right if given the move and for player 2 to opt out. Information is not really used as the outcome will be $(4, 2)$ independent of whether nature chooses I or II. Player 1 acquires the costly information in order to credibly threaten to guess Nature's move and thereby forcing player 2 to opt out.

In order to state the observations made in a formal way we need two definitions.

**Definition 1.** We say that *all costly information is used* in $G^P$ by $\sigma$ if for no player $i$ there exists a coarser information structure $\tilde{P}^i \in \wp_i$ with $c(i, P^i) > c(i, \tilde{P}^i)$ with respect to which either

 (i) $\sigma_i$ itself is measurable, or
(ii) some other strategy $\sigma_i'$ is measurable which yields the same payoff as $\sigma_i$ against $\sigma_{-i}$.

**Definition 2.** We say that *information hurts* player $i$ in the equilibrium $\sigma$ of $G^P$ if all sequential equilibria of $G^{\tilde{P}}$ yield player $i$ a weakly higher payoff than $\sigma$ in $G^P$, where $\tilde{P}$ is an information structure obtained from $P$ by replacing $i$'s information structure $P^i$ by some $\tilde{P}^i \in \wp_i$.

**Theorem 3.** *Let $G^P$ be a generic extensive form game with exogenous information structure $P \neq P^{\text{no}}$. Assume information costs of $P$ are small. For all players $i$ let $\wp_i$ be such that $P^i \in \wp_i$ and such that for all $\tilde{P}^i \in \wp_i$, $P^i$ is finer than $\tilde{P}^i$ or $P^i = \tilde{P}^i$.*

 (i) *Let $\sigma$ be a Nash equilibrium of $G^P$. Then $(P, \sigma)$ is a Nash equilibrium of the unobserved information acquisition game $\Gamma^{\text{unobs}}$ if all costly information is used in $\sigma$.*

(ii) *Let $\sigma$ be part of a sequential equilibrium of $G^P$. Then there exist beliefs that support $(P, \sigma)$ as a sequential equilibrium of the unobserved information acquisition game $\Gamma^{\text{unobs}}$ if all costly information is used in $\sigma$.*

(iii) *Let $\sigma$ be a Nash equilibrium of $G^P$ and suppose that all costly information is used. Then there exists a Nash equilibrium $(P, \tau)$ of the observed information acquisition game $\Gamma^{\text{obs}}$ with $\tau(P) = \sigma$.*

(iv) *Let $\sigma$ be part of a sequential equilibrium of $G^P$. Then there exist beliefs that support a sequential equilibrium $(P, \tau)$ of the observed information acquisition game $\Gamma^{\text{obs}}$ where $\tau(P) = \sigma$ if information does not hurt any player.*

**Proof.** (i) Since $\sigma$ is a Nash equilibrium of $G^P$, information acquisition is not observed, and $P^i$ is the finest partition available for each player $i$, the only unilateral deviation from $(P, \sigma)$ that could possibly improve upon a player's payoff would involve this player choosing a coarser and cheaper information structure $\tilde{P}^i$. However, since all costly information is used in $\sigma$ this implies that this player must then use a strategy $\sigma_i'$ which yields a strictly lower payoff against $\sigma_{-i}$ than $\sigma_i$. In principle this could be compensated by the lower cost of the coarser information structure $\tilde{P}^i$. However, the cost saving is bounded by the cost of the finest partition $P^i$, which is assumed to be small. Hence, no unilateral deviation is profitable.

(ii) This follows from (i) and the fact that the information sets in $G^P$ correspond uniquely to the information sets in the game playing phase of $\Gamma^{\text{unobs}}$. Namely, for each action information set $h$ and each element $P_k^i$ of $i$'s information structure there is one information set for player $i$. The difference is that in $G^P$ player $i$ knows the information structure of other players whereas in $\Gamma^{\text{unobs}}$ player $i$ does not know. But we can define beliefs in the latter information sets as in $G^P$ by letting all nodes corresponding to information structures $\tilde{P}^j \neq P^j$ have zero probability. That is, each player expects other players to have acquired $P$. Finally, these beliefs are consistent as they can be obtained as limits of beliefs deduced from Bayes' rule from completely mixed strategy profiles by having information structures $\tilde{P}^j \neq P^j$ be played with probability of order $\varepsilon^K$ with $K$ sufficiently large.

(iii) Let $\tau_i(\tilde{P}) = \sigma_i$ for each player $i$ and all information structures $\tilde{P}$ with $\tilde{P}^i = P^i$. We claim that the strategies $(P^i, \tau_i)$ constitute a Nash equilibrium. Obviously, given the strategies of the other players in the information acquisition stage, it does not matter what $\tau_i$ specifies in case of information structures different from $P$. In the case of information structure $P$, $\tau_i$ prescribes optimal actions since $\sigma$ is a Nash equilibrium in $G^P$. Finally, deviating in the information gathering stage to a coarser information structure $\tilde{P}^i$ is not beneficial since the other players do not change their actions as a result of that and since all costly information is used and the costs are assumed to be very small. (See also proof of (i).)

(iv) If information does not hurt any player, we can choose for each 'subgame' $G^{P'}$ (where $P'$ is obtained from $P$ by replacing the information structure of one player with a coarser one) a sequential equilibrium $\sigma'$ that yields the corresponding player a strictly lower payoff. Then it is clear that by threatening to play $\sigma'$ if $P'$ is gathered will keep all players from deviating in the information gathering phase, when

information costs are sufficiently small. It is also clear that this equilibrium is sequential.    □


## 4. Conclusion

In this paper, we investigated the robustness of equilibrium results of game models with respect to the endogenization of the information structure of the players. It turned out that only when information acquisition is unobserved and players choose actions simultaneously in the original game, the results are fully robust as long as all information is used. When information acquisition is observed, the endogenization process eliminates the equilibria of the game with an exogenous information structure where information hurts. On the other hand, when players move sequentially (and information acquisition is unobserved), the endogenization may generate additional sequential equilibria. In the latter case endogenous information acquisition may explain, what seemed to be irrational behaviour in the game with an exogenous information structure, as rational equilibrium outcomes. The fact that we need *unobserved* information acquisition to obtain robustness is somewhat surprising, since in that case the information structure will not become common knowledge, while in the case of observed information acquisition and games with an exogenous information structure the latter is always common knowledge.

Our results once again demonstrate the important difference between unobserved and observed information acquisition.[10] When information acquisition is unobserved, information is acquired for informational purposes only. That is, it is acquired because it allows the person who possesses it to make better decisions. When information acquisition is observed, however, it may be acquired (or not) for strategic reasons: By committing to have (or not have) some piece of information the actions of other players can be influenced in a way that is favorable to the first player. Since the difference between unobserved and observed information acquisition is so important, the choice between the two should be determined by which resembles reality best, and not so much by analytical convenience.

We have considered information acquisition as refining one's partition of the state space. In the literature private information is sometimes modelled by players receiving an imprecise signal about the true state. Also in this case one can endogenize the information structure by having players decide on the precision of information they want. We conjecture that our results also hold in this case: When information acquisition is unobserved and play is simultaneous, there will be a one-to-one correspondence between the results of the exogenous and endogenous information models. With sequential moves (and unobserved information) additional equilibria will appear, while in the case of observed information acquisition some equilibria may disappear. We can even use the same examples to prove the latter statements. Just interpret no learning as information with precision

---

[10]This point was already made, in the context of Cournot competition, by Hauk and Hurkens (2001) and Hurkens and Vulkan (2003) who focussed on the incentives to gather information.

zero, and full learning as information of precision 1 (or infinite). For the case where precisions can be chosen from a continuum, one would have to look further to come up with some examples, but we conjecture that there is no fundamental difference and that such examples can be constructed.

The results obtained in this paper should be useful to come up with more applications of the type discussed in the introduction where information acquisition may lead to new insights and surprising results. Assuming that information acquisition is costly and is not observed by the other players may lead to such conclusions in games with a sequential structure. A potential application would be to compare Dutch and English auctions. The Dutch auction is basically a simultaneous move game, whereas the English auction is sequential as one observes all the bids made. Hence in the English auction there could be some revision of beliefs about the informedness of competing bidders which cannot occur in the Dutch auction. Future research in this direction could prove to be interesting.

## Acknowledgements

## References

Cremer, J., Khalil, F., 1992. Gathering information before signing a contract. American Economic Review 82, 566–578.

Cremer, J., Khalil, F., 1994. Gathering information before the contract is offered: the case with two states of nature. European Economic Review 38, 675–682.

Cremer, J., Khalil, F., Rochet, J.-C., 1998. Strategic information gathering before a contract is offered. Journal of Economic Theory 81, 163–200.

Grossman, S., Stiglitz, J., 1980. On the impossibility of informationally efficient markets. American Economic Review 70, 393–408.

Hauk, E., Hurkens, S., 2001. Secret information acquisition in cournot markets. Economic Theory 18 (3), 661–681.

Hurkens, S., Vulkan, N., 2001. Information acquisition and entry. Journal of Economic Behavior and Organization 44 (4), 467–479.

Hurkens, S., Vulkan, N., 2003. Free entry does not imply zero profits. Economics Letters 81 (3), 285–290.

Hwang, H., 1993. Optimal information acquisition for heterogeneous duopoly firms. Journal of Economic Theory 59, 385–402.

Hwang, H., 1995. Information acquisition and relative efficiency of competitive, oligopoly, and monopoly markets. International Economic Review 36, 325–340.

Kessler, A., 1998. The value of ignorance. RAND Journal of Economics 29, 339–354.

Lee, T.K., 1982. Resource information policy and federal resource leasing. Bell Journal of Economics 13, 561–567.

Li, L., McKelvey, R.D., Page, T., 1987. Optimal research for cournot oligopolists. Journal of Economic Theory 42, 140–166.

Matthews, S., 1984. Information acquisition in discriminatory auctions. In: Boyer, M., Kihlstrom, R.E. (Eds.), Bayesian Models in Economic Theory. Elsevier, Amsterdam.

Milgrom, R., 1981. Rational expectations, information acquisition and competitive bidding. Econometrica 49, 921–944.

Neyman, A., 1991. The positive value of information. Games and Economic Behavior 3, 350–355.

Perea y Monsuwé, A., 1997. Buying information in extensive form games. Mimeo.

Persico, N., 2000. Information acquisition in auctions. Econometrica 68, 135–148.

Ponssard, J.-P., 1979. The strategic role of information on the demand function in an oligopolistic market. Management Science 25, 243–250.

Vives, X., 1988. Aggregation of information in large Cournot markets. Econometrica 56, 851–876.