

Cognitive Games and Cognitive Traps*

Jean Tirole[†]

July 21, 2015

Abstract

The paper defines “cognitive games” as games in which players covertly choose their information structures, and play a normal- or extensive-form game under the resulting information structures. It introduces the concept of “expectation conformity”, according to which the vector of information structures is potentially the object of self-fulfilling prophecies; loosely, each information-acquiring player has more incentive to select a given information structure if he expected to do so.

For example, games of pure conflict (zero-sum games) never give rise to self-fulfilling cognition while games of pure alignment (coordination games) always do. The paper defines a class of games for which a direct characterization of the expectation conformity property in terms of rotation points can be obtained. This class comprises many games of interest to economists, starting with the cognition-augmented lemons model.

Keywords: cognition, expectation conformity, adverse selection.

JEL numbers: C72; C78; D82; D83; D86.

1 Introduction

Cognition is costly; thinking and memorizing, obtaining financial, engineering and legal expertise and brainstorming with others consumes resources (in amounts that depend on the urgency to act, the cognitive load or the context) and is strategic. This paper considers the implications of costly cognition in a multi-player context. It defines cognitive games as games in which players privately select information structures and then play an arbitrary, normal or extensive form game under the resulting information structures. For convenience, we will label the latter game the “stage-2 game” and the information acquisition stage “stage 1”. If the stage-1 choice of cognition and the choice of stage-2 strategy can, due to their unobservability, be viewed as simultaneous from a game-theoretic perspective, our characterizations will purport to classes of second-stage games, and so we keep the two stages separate for the exposition.

*The author is very grateful to participants at conferences and seminars at MIT, EARIE, Games Toulouse, Game Theory Society World Congress, Econometric Society meetings, the University of Chicago, and the University of Tokyo, and to Gabriel Carroll, Drew Fudenberg, Christian Gollier, Navin Kartik, Alessandro Pavan, Di Pei, Jean-Charles Rochet, Joel Sobel, Pierre-Luc Vautrey and Olivier Wang for helpful discussions and comments, and to a European Research Council advanced grant (European Community’s Seventh Framework Programme (FP7/2007-2013) Grant Agreement no. 249429) for financial support.

[†]Toulouse School of Economics (TSE) and Institute for Advanced Study in Toulouse (IAST), jean.tirole@tse-fr.eu

Individually optimal and anticipated cognitions can, under some circumstances, be expected to be “strategic complements”; for instance, a higher level of anticipated cognitive activity may incentivize the player to indeed engage in more cognition. Section 2 accordingly introduces expectation-conformity (EC). EC holds if there exist information structures $\mathcal{F} = \{\mathcal{F}_i\}_{i \in I}$ and $\widehat{\mathcal{F}} = \{\widehat{\mathcal{F}}_i\}_{i \in I}$ such that each player $i \in I$ has more incentive to acquire information $\widehat{\mathcal{F}}_i$ rather than \mathcal{F}_i when all players expect information structure $\widehat{\mathcal{F}}$ rather than \mathcal{F} . A special case of interest arises when information structures are ordered and a finer information structure is costlier than a coarser one. A cognitive trap obtains when two ordered equilibria co-exist and information-acquiring players are better-off in the low-cognition equilibrium.

EC straightforwardly leads to a multiplicity of equilibrium information structures for appropriate information acquisition cost functions. By contrast, equilibrium uniqueness prevails if EC is never satisfied.

EC is violated when the stage-2 game is a two-person zero-sum game. It is always satisfied in coordination games.

Section 3 studies “generalized lemons environments”, in which one of the players acquires information and then picks between an interaction-free option and one whose return depends on the other player’s beliefs; for example, he may decide whether to play a game with the other player or to opt out. The key assumption is that news that makes the first player want to interact with the second player also makes the second player react in an unfriendly manner. This class of games includes the lemons model of Akerlof (augmented with one-sided information acquisition of soft or hard information) and a number of other games of interest in economics. Section 3 derives a sufficient condition for such games to satisfy expectation conformity. It assumes that the class of information structures is composed of rotations and shows that EC holds for sufficient “gains from interaction”, although perhaps not for low gains from interacting. The condition for EC to hold is much easier to check than verifying directly that expectation conformity prevails. It then applies the general result to the cognition-augmented lemons game under directed and non-directed search. It also derives results for a general “anti-lemons environment” in which what makes the first player want to interact makes the second player want to behave in a friendly manner. Section 4 concludes with alleys for future research.

Broader motivation: Covert investments

The paper’s emphasis will be on information acquisition, a choice motivated both by the applications and by the fact that cognitive investments are the ultimate covert investments. But there is interest in other forms of covert investments: in capacity, learning by doing, arms buildup, etc. To prepare the ground for the subsequent material, it is useful to consider one such environment. Suppose that there are two players, playing a “second-stage” normal-form game with actions $a_i, a_j \in \mathbb{R}$, and picking an “first-stage” investment $\rho_i \in \mathbb{R}$ at increasing investment cost $C_i(\rho_i)$. Payoffs are

$$[\phi_i(a_i, a_j) - \psi_i(a_i, \rho_i)] - C_i(\rho_i),$$

where all functions are C^2 and satisfy for all i, j :

$$\frac{\partial^2 \psi_i}{\partial a_i \partial \rho_i} < 0 \quad \text{and} \quad \frac{\partial^2 \phi_i}{\partial a_i \partial a_j} \begin{cases} > 0 & (SC) \\ \text{or} \\ < 0 & (SS). \end{cases}$$

That is, we assume that the investment ρ_i lowers the marginal cost of action a_i , and that the strategic interaction involves either strategic complementarity (*SC*) or strategic substitutability (*SS*). One may have in mind that a_i is a quantity, ρ_i an investment that lowers the marginal cost of production and production a_j is a strategic complement or substitute to production a_i .

Assume for simplicity that (only) player i invests and that, were her investment common knowledge, the normal-form game in (a_i, a_j) would have a unique and stable equilibrium. In particular, player j 's equilibrium action $\mathbf{a}_j(\rho_i^\dagger)$ under common knowledge of cognition ρ_i^\dagger is increasing in ρ_i^\dagger under (*SC*) and decreasing in ρ_i^\dagger under (*SS*).

Suppose that player i 's actual choice ρ_i is not observed by j (so de facto the game is a simultaneous-move game in actions (ρ_i, a_i) and a_j , respectively). One can also define player i 's optimal action when she deviates from her equilibrium investment. Our assumptions imply that player i 's optimal action $\mathbf{a}_i(\rho_i, \rho_i^\dagger)$ given expected cognition ρ_i^\dagger and actual cognition ρ_i is non-decreasing in ρ_i^\dagger under either (*SC*) or (*SS*).

This environment is similar to that considered in the industrial organization literature on the taxonomy of business strategies¹, except for one twist: The investment choice ρ_i is not observed by firm j and so has no commitment effect; rather, what matters for the outcome of the normal form game is the anticipation ρ_i^\dagger by j of firm i 's choice as well as the actual choice ρ_i (of course, in pure strategy equilibrium $\rho_i^\dagger = \rho_i$).

Letting

$$T_i(\rho_i, \rho_i^\dagger) \equiv \max_{\{a_i\}} \left\{ \phi_i(a_i, \mathbf{a}_j(\rho_i^\dagger)) - \psi_i(a_i, \rho_i) - C_i(\rho_i) \right\}$$

denote player i 's payoff when actual cognition is ρ_i and player j anticipates cognition ρ_i^\dagger . The assumptions imply that, whether (*SC*) or (*SS*) prevails, for all (ρ_i, ρ_i^\dagger) and $(\widehat{\rho}_i, \widehat{\rho}_i^\dagger)$ with $\widehat{\rho}_i \geq \rho_i$ and $\widehat{\rho}_i^\dagger \geq \rho_i^\dagger$, the following increasing differences condition is satisfied:

$$T_i(\widehat{\rho}_i, \widehat{\rho}_i^\dagger) - T_i(\rho_i, \widehat{\rho}_i^\dagger) \geq T_i(\widehat{\rho}_i, \rho_i^\dagger) - T_i(\rho_i, \rho_i^\dagger)$$

We call this property “expectation conformity”.

Consequently, let ρ_i (resp. $\widehat{\rho}_i$) denote i 's optimal investment when player j expects investment ρ_i^\dagger (resp. $\widehat{\rho}_i^\dagger$).² One can show that there is a complementarity between investment and anticipation of investment: $(\widehat{\rho}_i - \rho_i)(\widehat{\rho}_i^\dagger - \rho_i^\dagger) \geq 0$ whether the second-stage game involves strategic substitutes or strategic complements.

The intuition goes as follows: Suppose that firm j expects i to invest more in capacity and

¹Eg. Bulow et al (1985) and Fudenberg and Tirole (1984).

²That is, ρ_i results from

$$\max_{\{\rho_i, a_i\}} \left\{ \left[\phi_i(a_i, \mathbf{a}_j(\rho_i^\dagger)) - \psi_i(a_i, \rho_i) \right] - C_i(\rho_i) \right\}$$

and $\{\mathbf{a}_i(\rho_i^\dagger), \mathbf{a}_j(\rho_i^\dagger)\}$ is the Nash equilibrium of the normal-form game under common knowledge that i has invested ρ_i^\dagger (i.e. under symmetric information).

therefore to produce more output. It will accordingly raise its output (*SC*) or decrease it (*SS*). Firm i is then induced in both cases to raise its output, vindicating a higher investment in the first place. It can also easily be checked that when there are two equilibria ($\rho_i = \rho_i^\dagger$ and $\hat{\rho}_i = \hat{\rho}_i^\dagger$), player i prefers the high investment one, again regardless of the strategic interaction (*SC* or *SS*).

Games such as this one are simple. “Cognitive games”, in which the investment is in a filtration of the state space, are a priori much more complex. The action “ a_i ” is then an information-contingent one, i.e. a function. To obtain results on expectation conformity or its absence, we will need to put structure on the second-stage game. While no further assumption of the nature of information acquisition may be needed (zero-sum games), in general some regularity on the family of potential information structures needs to be imposed. This is for example what we will do when studying the generalized lemons environment, for which we will draw a formal analogy with the covert investment model just analyzed.

Relationship to the literature: Several literatures, including those on search (starting with Stigler 1961), on rational inattention (e.g., Maćkowiak and Wiederholt 2009, Matejka and McKay 2012, Sims 2003), or on security design with information acquisition (e.g., Dang et al 2011, Farhi and Tirole 2015, Yang 2013) have used costly-cognition models. Our particular interest here is on strategic interactions with information acquisition.

This interaction also has been the focus of the literatures on information acquisition and aggregation in competitive markets building on Morris and Shin (2002)’s beauty-contest model (e.g., Colombo et al 2014, Hellwig and Veldkamp 2009, Llosa and Venkateswaran 2012, Myatt and Wallace 2012, Pavan 2014),³ and on information acquisition prior to an auction (e.g., Persico 2000) or to contracting (e.g. Dang 2008, Tirole 2009, Bolton and Faure-Grimaud 2010).

Technically, the paper is an application of the general theory of supermodular games (e.g. Milgrom-Shannon 1994). The added value is therefore not in the techniques, which are well-known. The contribution is three-fold: First, the paper introduces the notion of expectation conformity and its impact on self-fulfilling cognition. Second, it analyses whether the condition obtains in some familiar games (zero-sum, coordination). Finally, for an interesting class of games the paper provides a sufficient condition for *EC* that is much simpler to check than verifying directly that *EC* obtains.

³Hellwig and Veldkamp (2009) assume that players pay to receive signals of varying precisions and show that public signals, unlike private ones, create scope for equilibrium multiplicity; a signal serves both to better adjust one’s action to the state of nature, and also, if it is public, to coordinate with the other players’ actions. Myatt and Wallace (2012) demonstrate that for different information acquisition technologies, equilibrium uniqueness need not rely on private signals. In their model, players exert effort to achieve a better understanding of existing public signals (select “receiver noise”); this may naturally give rise to decreasing returns in the understanding effort. They derive a unique linear equilibrium, with interesting comparative statics. Amir and Lazzati (2010) consider general games with strategic complementarities and, for given information structures, select the equilibrium with the maximal actions. They emphasize increasing returns in information acquisition and derive existence of pure strategy Bayesian equilibria.

2 Cognitive games

2.1 Model and expectation-conformity

There are n players, $i \in I = \{1, \dots, n\}$. Some of the results, and all applications will refer to the two-player environment, though. In the “stage-2 game”, the players play an arbitrary, normal or extensive form game. Player i has action space A_i and receives gross payoff $u_i(\sigma_i, \sigma_{-i}, \omega)$ in state of nature $\omega \in \Omega$, where (σ_i, σ_{-i}) are mixed strategy profiles. The players have a common prior distribution Q on the state space Ω .⁴ Gross expected payoffs are

$$U_i(\sigma_i, \sigma_{-i}) \equiv E_\omega \left[u_i(\sigma_i, \sigma_{-i}, \omega) \right].$$

Prior to playing the stage-2 game, the players *privately* choose at stage 1 their information structure. Let Ψ_i denote the set of available information structures or sigma-fields \mathcal{F}_i for player i , and $C_i(\mathcal{F}_i)$ player i 's cost of acquiring information \mathcal{F}_i . Player i 's stage-2 strategy must be measurable with respect to the information structure \mathcal{F}_i chosen by player i at stage 1.⁵ The expected *net* payoffs are equal to the gross expected stage-2 payoffs minus the stage-1 information acquisition costs. In a number of applications only one of the players acquires information; this amounts to the other players' having infinite cost except for some partition; we will call this case “one-sided cognition”.

A special case that is prominent in applications arises when the sets Ψ_i of information structures are totally ordered. Player i 's choice of information structure is then represented by a filtration $\{\mathcal{F}_{i,\rho}\}$, where $\rho \in \mathbb{R}$ and $\mathcal{F}_{i,\rho}$ is an increasing sequence of sigma-algebras: For $\rho_1 < \rho_2$, \mathcal{F}_{i,ρ_2} is finer than \mathcal{F}_{i,ρ_1} ($\mathcal{F}_{i,\rho_1} \subset \mathcal{F}_{i,\rho_2}$). We will then assume that C_i is monotonically increasing: A finer partition is more costly.⁶

For simplicity, we will be focusing on equilibria in which players use a pure strategy at the information acquisition stage.⁷ All players share expectations as to how much cognition other players engage in even though they do not observe the actual realization of these cognitive choices. Thus, consider a common knowledge information structure for the players $\mathcal{F} = (\mathcal{F}_1, \dots, \mathcal{F}_n)$. We let $\sigma^*(\mathcal{F}) = \{\sigma_i^*(\mathcal{F})\}_{i \in I}$ denote the stage-2 equilibrium strategy profile for \mathcal{F} ; that is, we assume that either the stage-2 equilibrium is unique given the commonly known information structure or some equilibrium selection has been performed; otherwise, the date-1 choices of information acquisition are not well-defined.⁸ From now on and unless otherwise stated, “equilibria” will

⁴Che and Kartik (2009) by contrast look at incentives to acquire information in an environment with heterogeneous priors.

⁵Messages and disclosure decisions, if any, are part of the stage-2 strategies in this formulation.

⁶This need not be the case for all applications. Consider memory management, an instance of a “signal-jamming” cognitive gamely as discussed in Appendix B: Increasing the probability of forgetting some information that one has received (repression) is likely to be costly. By contrast, the case in which a player receives two pieces of information simultaneously when searching and would have to pay an extra cost to receive only one (unbundling) is not problematic in our interpersonal *covert*-information-acquisition context: the unbundled information structure is simply irrelevant and can be assumed not to belong to Ψ_i (this would not be the case with overt information acquisition since we know that a player may suffer when other players know that he has more information).

⁷If stage 2 corresponds to an extensive form game and player i 's cognition is in mixed strategy, then player i 's early actions in stage 2 might reveal something about his actual choice of cognition.

⁸Existence of a stage-2 equilibrium follows standard assumptions.

therefore refer to pure-strategy equilibria of the (stage-1) information acquisition game.

Let \mathcal{F} and $\widehat{\mathcal{F}}$ denote two arbitrary information structures, and σ and $\widehat{\sigma}$ denote the stage-2 equilibrium strategy profiles for information structures \mathcal{F} and $\widehat{\mathcal{F}}$, respectively. Let

$$V_i(\mathcal{F}'_i; \mathcal{F}) = \max_{\{\sigma_i, \mathcal{F}'_i\text{-measurable}\}} \left\{ U_i(\sigma_i, \sigma_{-i}^*(\mathcal{F})) \right\}$$

denote player i 's gross payoff from deviating to information structure \mathcal{F}'_i when he is expected to choose \mathcal{F}_i and the other players have information \mathcal{F}_{-i} .

Definition 1 (expectation conformity). *Expectation conformity for information structures \mathcal{F} and $\widehat{\mathcal{F}}$ is satisfied if for all i*

$$V_i(\widehat{\mathcal{F}}_i; \widehat{\mathcal{F}}) - V_i(\mathcal{F}_i; \widehat{\mathcal{F}}) \geq V_i(\widehat{\mathcal{F}}_i; \mathcal{F}) - V_i(\mathcal{F}_i; \mathcal{F}) \quad (EC_{\{\mathcal{F}, \widehat{\mathcal{F}}\}})$$

Expectation conformity is an increasing differences condition. While familiar from the theory of supermodular games, the economic emphasis is what distinguishes this from previous contributions. A special case of expectation conformity arises when only one of the players can acquire information, i.e. the other players are endowed with a fixed information structure. We will then label the condition “unilateral expectation conformity” if they is a need to distinguish it from expectation conformity.

2.2 Equilibrium multiplicity/uniqueness

Let

$$\Gamma_i^{EC}(\mathcal{F}, \widehat{\mathcal{F}}) \equiv \left[V_i(\widehat{\mathcal{F}}_i; \widehat{\mathcal{F}}) - V_i(\mathcal{F}_i; \widehat{\mathcal{F}}) \right] - \left[V_i(\widehat{\mathcal{F}}_i; \mathcal{F}) - V_i(\mathcal{F}_i; \mathcal{F}) \right].$$

Revealed preference implies that a necessary condition for $(\mathcal{F}, \widehat{\mathcal{F}})$ to form two equilibria is that $\Gamma_i^{EC}(\mathcal{F}, \widehat{\mathcal{F}}) \geq 0$ for all i . The following proposition says that the condition is also sufficient for both \mathcal{F} and $\widehat{\mathcal{F}}$ to be equilibria for appropriately chosen cost functions:

Proposition 1 (multiplicity and uniqueness).

(i) *If $EC_{\{\mathcal{F}, \widehat{\mathcal{F}}\}}$ is satisfied for two distinct information structures \mathcal{F} and $\widehat{\mathcal{F}}$, then there exist cost functions $\{C_i(\cdot)\}_{i=1, \dots, n}$ such that \mathcal{F} and $\widehat{\mathcal{F}}$ are both equilibrium information structures of the stage-1 game. If furthermore Ψ_i is totally ordered and $\widehat{\mathcal{F}}$ is finer than \mathcal{F} , the cost functions can be chosen to be monotonic.*

(ii) *If $EC_{\{\mathcal{F}, \widehat{\mathcal{F}}\}}$ is satisfied for no two distinct information structures $(\mathcal{F}, \widehat{\mathcal{F}})$, then there cannot exist multiple equilibria.*

Proof: (i) Assume that $EC_{\{\mathcal{F}, \widehat{\mathcal{F}}\}}$ is satisfied. For \mathcal{F} and $\widehat{\mathcal{F}}$ to be both equilibrium information structures, it is necessary that for all i

$$V_i(\widehat{\mathcal{F}}_i; \mathcal{F}) - V_i(\mathcal{F}_i; \mathcal{F}) \leq C_i(\widehat{\mathcal{F}}_i) - C_i(\mathcal{F}_i) \leq V_i(\widehat{\mathcal{F}}_i; \widehat{\mathcal{F}}) - V_i(\mathcal{F}_i; \widehat{\mathcal{F}}). \quad (1)$$

In the absence of further requirement, pick cost functions satisfying (1) as well as $C_i(\tilde{\mathcal{F}}_i) = +\infty$ if $\tilde{\mathcal{F}}_i \notin \{\mathcal{F}_i, \hat{\mathcal{F}}_i\}$.

When Ψ_i is totally ordered and $\mathcal{F}_i \subseteq \hat{\mathcal{F}}_i$ for all i , pick a cost function satisfying (1) as well as:

$$C_i(\tilde{\mathcal{F}}_i) = \begin{cases} C_i(\mathcal{F}_i) & \text{for } \tilde{\mathcal{F}}_i \subseteq \mathcal{F}_i \\ C_i(\hat{\mathcal{F}}_i) & \text{for } \mathcal{F}_j \subset \tilde{\mathcal{F}}_i \subseteq \hat{\mathcal{F}}_i \\ +\infty & \text{for } \hat{\mathcal{F}}_i \subset \tilde{\mathcal{F}}_i. \end{cases}$$

From (1) and the fact that $V_i(\hat{\mathcal{F}}_i; \mathcal{F}) - V_i(\mathcal{F}_i; \mathcal{F}) \geq 0$, $C_i(\hat{\mathcal{F}}_i) \geq C_i(\mathcal{F}_i)$ and so $C_i(\cdot)$ is indeed monotonic. Because more information cannot hurt if covertly acquired, \mathcal{F} and $\hat{\mathcal{F}}$ are indeed both equilibria for these cost functions.

(ii) Conversely, if \mathcal{F} and $\hat{\mathcal{F}}$ were two distinct equilibria, condition (1) would be satisfied, and so $EC_{\{\mathcal{F}, \hat{\mathcal{F}}\}}$ would hold, a contradiction. ■

Definition 2 (cognitive trap). *Players are exposed to a cognitive trap if there exist two information structures \mathcal{F} and $\hat{\mathcal{F}}$ such that*

- (i) \mathcal{F} and $\hat{\mathcal{F}}$ are both equilibria, and for all i such that $\hat{\mathcal{F}}_i \neq \mathcal{F}_i$:
- (ii) $\hat{\mathcal{F}}_i$ is finer than \mathcal{F}_i ,
- (iii) $V_i(\mathcal{F}_i; \mathcal{F}) - C_i(\mathcal{F}_i) > V_i(\hat{\mathcal{F}}_i; \hat{\mathcal{F}}) - C_i(\hat{\mathcal{F}}_i)$.

If the conditions in Definition 2 are fulfilled, cost functions are such that players who alter their information structure conform to expectations, choosing either $\hat{\mathcal{F}}_i$ or \mathcal{F}_i when expected to (condition (i)), and prefer the low-cognition outcome to the high-cognition one (conditions (ii) and (iii)).

2.3 Pure conflict and pure alignment: Zero-sum and coordination games

(a) Two-person zero-sum games

Before we move on to analyze classes of games that satisfy expectation conformity, it is interesting to consider an important class that does not satisfy it. Suppose that the stage-2 game is a zero-sum game (or more generally a constant-sum game) between two players; that is, the gross payoffs satisfy the zero-sum condition: for all $(\sigma_i, \sigma_j, \omega)$

$$u_i(\sigma_i, \sigma_j, \omega) + u_j(\sigma_j, \sigma_i, \omega) = k(\omega),$$

where k is an arbitrary function of the state of nature. The overall game obviously is not a zero-sum game. Any information acquisition, if costly, necessarily reduces total surplus and just amounts to pure rent-seeking.

Zero-sum games have several remarkable properties; for example, a player can only benefit from having (and being known to have) more information (Lehrer and Rosenberg 2006), a property that is well-known to be violated for general games. Another interesting property is given by the following result.⁹

Proposition 2 (two-person zero-sum games). *Two-person zero-sum games satisfy for all $(\mathcal{F}, \widehat{\mathcal{F}})$*

$$\Sigma_i \Gamma_i^{EC}(\mathcal{F}, \widehat{\mathcal{F}}) \leq 0.$$

As a consequence, if there are multiple equilibria, in none can a player have a strict preference for his equilibrium strategy (a fortiori, there cannot exist a strict equilibrium).

Proof: The zero-sum property implies that for all strategies (σ_i, σ_j) and $(\hat{\sigma}_i, \hat{\sigma}_j)$,

$$\Sigma_i \left[[U_i(\hat{\sigma}_i, \hat{\sigma}_j)] - [U_i(\sigma_i, \hat{\sigma}_j)] - [U_i(\hat{\sigma}_i, \sigma_j) - U_i(\sigma_i, \sigma_j)] \right] = 0.$$

Now consider two information structures $(\mathcal{F}, \widehat{\mathcal{F}})$ and strategies $\sigma = (\sigma_i, \sigma_j)$ \mathcal{F} -measurable and $\hat{\sigma} = (\hat{\sigma}_i, \hat{\sigma}_j)$ $\widehat{\mathcal{F}}$ -measurable. Let $R_i(\hat{\sigma}_j)$ denote player i 's best \mathcal{F}_i -measurable response to $\hat{\sigma}_j$ and $\widehat{R}_i(\sigma_j)$ denote player i 's best $\widehat{\mathcal{F}}_i$ -measurable response to σ_j . Obviously, $V_i(\mathcal{F}_i; \widehat{\mathcal{F}}) = U_i(R_i(\hat{\sigma}_j), \hat{\sigma}_j) \geq U_i(\sigma_i, \hat{\sigma}_j)$ and $V_i(\widehat{\mathcal{F}}_i; \mathcal{F}) = U_i(\widehat{R}_i(\sigma_j), \sigma_j) \geq U_i(\hat{\sigma}_i, \sigma_j)$. This implies that

$$\Sigma_i \Gamma_i^{EC}(\mathcal{F}, \widehat{\mathcal{F}}) \leq 0$$

for all $(\mathcal{F}, \widehat{\mathcal{F}})$. Because equilibrium multiplicity requires $\Gamma_i^{EC}(\mathcal{F}, \widehat{\mathcal{F}}) \geq 0$ for all i , the inequality implies indifference for both players, i.e., $\Gamma_i^{EC}(\mathcal{F}, \widehat{\mathcal{F}}) = 0$ for all i . Suppose, say, that in the $(\widehat{\mathcal{F}}_i, \widehat{\mathcal{F}}_j)$ equilibrium, player i has strictly optimal strategy $\widehat{\mathcal{F}}_i$. Then

$$V_i(\widehat{\mathcal{F}}_i; \widehat{\mathcal{F}}) - V_i(\mathcal{F}_i; \widehat{\mathcal{F}}) > C_i(\widehat{\mathcal{F}}_i) - C_i(\mathcal{F}_i) \geq V_i(\widehat{\mathcal{F}}_i; \mathcal{F}) - V_i(\mathcal{F}_i; \mathcal{F}),$$

and so $\Gamma_i^{EC}(\mathcal{F}, \widehat{\mathcal{F}}) > 0$, a contradiction.¹⁰ ■

To illustrate the possibility of double indifference, consider the zero-sum game in which i 's payoff is $(a_i - a_j)\omega$ where $a_k \in \{1, -1\}$ for all k and ω takes value 1 and -1 with equal probabilities. Each player can learn ω at cost 1. Regardless of j 's behavior, i is indifferent between acquiring the information or not. There are multiple equilibria with different levels of ex-ante payoffs.

⁹I am grateful to Gabriel Carroll for prompting me to have a look at zero-sum games and for conjecturing that they do not satisfy expectation conformity.

¹⁰In contrast, an n -player zero-sum game ($n > 2$) may admit multiple strict equilibria. To see this, take a non-zero-sum two-player game admitting multiple strict equilibria in information acquisition (such as the coordination game studied below), Have this game played twice, by players 1 and 2 and by players 3 and 4 respectively. Use players 3 and 4 (resp. players 1 and 2) as passive "budget balancers" in the game played by 1 and 2 (resp., 3 and 4). The transformed game is a zero-sum game that admits multiple strict equilibria.

(b) *Coordination games*

In stark contrast with zero-sum games, which exhibit completely antinomic interests, coordination games (which have been the focus of much recent interest in macroeconomics) involve perfectly aligned interests. With two players, a typical coordination game has payoffs¹¹

$$u_i(a_i, a_j, \omega) = -(a_i - \omega)^2 - (a_i - a_j)^2;$$

that is, each player wants to match his action both with the state of nature and with the other player's choice. The density of ω is continuous on some interval $[\omega^{\text{inf}}, \omega^{\text{sup}}]$, say.

As for the sets Ψ_i of information structures, we assume that they are totally ordered. We take the filtration to be a sequence of finer and finer (and more and more costly) information structures. Furthermore, we assume that feasible information structures are the same for both players: $\Psi_i = \Psi_j$.

Suppose that player i has chosen a (weakly) finer information structure than player j : $\mathcal{F}_j \subseteq \mathcal{F}_i$. An element of \mathcal{F}_i , say, is then characterized by a mean ω_i and a conditional variance $\sigma_i^2 \equiv E[(\omega - \omega_i)^2]$. Player i knows ω_j and is able to predict j 's choice, but the converse may not hold. Like in the coordination games literature, the parties do not communicate prior to choosing their actions. Optimal actions are then

$$a_j = \frac{E(a_i(\omega_i)|\mathcal{F}_j) + \omega_j}{2} \quad \text{and} \quad a_i(\omega_i) = \frac{a_j + \omega_i}{2},$$

and so, using the law of iterated expectations,

$$a_j = \omega_j \quad \text{and} \quad a_i = \frac{\omega_j + \omega_i}{2}.$$

Furthermore

$$U_i = -\frac{(\omega_j - \omega_i)^2}{2} - \sigma_i^2 \quad \text{and} \quad U_j = -\frac{E(\omega_j - \omega_i)^2}{4} - \sigma_j^2$$

Proposition 3 (two-player coordination games). *Two-player coordination games with totally ordered information structures satisfy expectation conformity for any two distinct information structures $(\mathcal{F}_1, \mathcal{F}_2)$ and $(\widehat{\mathcal{F}}_1, \widehat{\mathcal{F}}_2)$ such that $\mathcal{F}_1 \subseteq \widehat{\mathcal{F}}_1$ and $\mathcal{F}_2 \subseteq \widehat{\mathcal{F}}_2$.*

The proof of Proposition 3 can be found in the Appendix. The one case in which expectation conformity is only weakly satisfied is when player i is always better informed than player j ($\mathcal{F}_j \subseteq \widehat{\mathcal{F}}_j \subseteq \mathcal{F}_i \subseteq \widehat{\mathcal{F}}_i$). Then player j does not adjust his strategy to the information held by player i and so the value of information for player i is independent of player j 's expectation.

The welfare comparison among equilibria of coordination games is in general ambiguous. On the one hand, a player may not increase his cognitive intensity by fear that the other would not, while more cognition would be beneficial for both. On the other hand, the two players may be trapped by the same coordination motive into a wastefully high-cognition state.¹²

¹¹See Angeletos and Pavan (2007) for a more general version than the quadratic coordination game. Much of the macroeconomic literature analyzes the relative use of public and private signals about the state of nature. This literature often assumes Gaussian distributions; we will not need this assumption for our purposes.

¹²Suppose, e.g., that $\omega = -a$ with probability 1/2 and $+a$ with probability 1/2. The no-cognition equilibrium

3 Generalized lemons environment

This section considers a class of Stackelberg games in which one player, player i , chooses whether to play “in” (i.e. interact with player j) or “out”, and then, if “in” has been selected, the other player, player j , chooses an action that is more or less “friendly” to player i . Player i covertly acquires information about the state of nature before making her decision; her set of possible information structures is indexed by a rotation parameter. In the lemons (anti-lemons) environment, information that makes player i eager to engage with player j would, if it were known, trigger a hostile (friendly) action from player j . We provide sufficient and/or necessary conditions for EC to hold in this class of games, which is shown to encompass a number of economic games of interest.

3.1 Description

One player, say player i (the “leader”), first chooses between “out” (“outside option”) and “in” (“interact with player j ”, the “follower”). When player i picks “in”, player j chooses an action $a_j \in \mathbb{R}$.¹³ Cognition is one-sided; prior to choosing between “in” and “out”, player i selects an information structure. Player j , when choosing a_j , by contrast knows only that player i chose “in”.

To take an example to which we will later return, player i might be the seller of a used car; $a_i = \text{“in”}$ if the seller decides to put the car in the market, and $a_i = \text{“out”}$ if he keeps the car. Then a_j is the price that a competitive buyer offers for the car (i.e. its expected value conditionally on the car being in the market). More generally, we will normalize player j ’s action so as to be a friendly one: player i ’s utility is increasing in a_j (conditional on playing “in”).

Prior to choosing between “in” and “out”, player i acquires information about the state of nature. We will assume that news that make player i want to interact with player j also make the latter choose an unfriendly action (a friendly action when later on we turn to the anti-lemons environment). In that sense, the game generalizes the lemons game as in that game the seller is more keen on parting with a low-quality car.

We assume that the two players’ preferences are quasi-linear in the state of nature, so they care only about the posterior mean of the state.

Information.

The state of nature is $\omega \in (-\infty, +\infty)$, with prior mean ω_0 and distribution $Q(\cdot)$. An experiment, indexed by $\rho \in \mathbb{R}$,¹⁴ will be taken to be the choice of a distribution $F(m; \rho)$ in a differentiable family of distributions over the posterior mean m , satisfying the martingale property $\int_{-\infty}^{+\infty} m dF(m; \rho) = \omega_0$ for all ρ .

exists if $-a^2 \geq -(a^2/2) - c$; the high-cognition equilibrium exists if $-c \geq -2a^2$. So the two equilibria co-exist whenever $(a^2/2) \leq c \leq 2a^2$. The no-cognition equilibrium dominates for $c > a^2$ and is dominated for $c < a^2$.

¹³More generally, player i may pick an action after picking “in”. Because a_i will be a best reaction to a_j , the envelope theorem implies that what matters is the impact of a_j on player i ’s payoff.

¹⁴We will not need to index by i here, since only player i acquires information.

Assumption 1 (rotations). *Player i 's set of possible information structures is indexed by the parameter ρ in the sense of “rotations” (or “simple mean-preserving spreads” or “single-crossing property”); that is, there exists m_ρ such that $F_\rho(m; \rho) \geq 0$ for $-\infty < m \leq m_\rho$ and $F_\rho(m; \rho) \leq 0$ for $m_\rho \leq m < +\infty$ (with some strict inequalities).¹⁵*

Preferences. Player i 's payoff difference between “in” and “out” depends on a_j , and on player i 's posterior beliefs only through the posterior mean m . This difference will be labeled $\delta_i(m, a_j)$, with $\partial\delta_i/\partial m > 0$ and $\partial\delta_i/\partial a_j > 0$ from our previous sign convention. The strict monotonicity of δ_i in m implies that player i plays “in” if and only if m exceeds some cutoff $m^*(a_j)$. The cutoff is decreasing in a_j .

Assumption 2 (leader's preferences). *Player i 's net payoff from playing “in”, $\delta_i(m, a_j)$, depends on i 's posterior mean m about ω and on j 's action a_j . $\delta_i(m, a_j)$ is twice differentiable, is increasing in m and a_j , and is such that the marginal impact of a friendly action ($\partial\delta_i/\partial a_j$) is weakly decreasing and weakly concave in the expected state of nature:*

$$\frac{\partial^2\delta_i}{\partial a_j\partial m} \leq 0 \quad \text{and} \quad \frac{\partial^3\delta_i}{\partial a_j\partial m^2} \leq 0.$$

In the four illustrations described in Section 3.2¹⁶, δ_i will take the specific functional form:

$$\frac{\partial\delta_i}{\partial a_j} \equiv \gamma - \tau m$$

with $\gamma > 0$ and $\tau \geq 0$, and so Assumption 2 is satisfied. For example, in the classic lemons game, the benefit of selling depends only on a_j and the utility from keeping the car depends only on m . So $\partial^2\delta_i/\partial a_j\partial m = 0$.

Let player j anticipate cognition ρ^\dagger by player i (ρ^\dagger out of equilibrium can differ from actual cognition ρ). Let $a_j(\rho^\dagger)$ denote the resulting equilibrium choice (at this stage player i 's decision depends only on a_j and m and no longer on the stage-1 choice of ρ). As stated above, we also assume that player j cares only about the posterior mean of m . And so, due to player i 's use of a cutoff rule, a_j depends only on $M^+(m^*(a_j), \rho^\dagger)$, where $M^+(m^*(a_j), \rho) \equiv E(m|m \geq m^*(a_j), \rho)$ denotes the truncated mean for information structure ρ .

¹⁵See e.g. Diamond and Stiglitz (1974) and Johnston and Myatt (2006). A simple mean-preserving spread (MPS) is a MPS, but the converse does not hold as a MPS may have multiple points of intersections (densities that cross more than twice). A combination of two rotations need not be a rotation, unless of course they have the same rotation point. But as is well-known, any MPS can be obtained through a sequence of simple MPS. Examples of mean-preserving spreads with a rotation include: the case of a normally distributed state of nature ω together with a signal that is normally distributed around the true state (ρ is then the precision of this signal); the class of triangular distributions on $[0, 1]$ with uniformly distributed underlying state (so $\rho = +\infty$ corresponds to $F(m; \rho) = m$ on $[0, 1]$). We will provide other examples of rotations later on. In these examples (Pareto, exponential, directed and non-directed search) as well as the two just discussed, information structures are ordered.

¹⁶This will also hold for the warfare game of Section 3.4. By contrast, this functional form holds, but with $\gamma = 0$, $\tau = -1$ for the leadership game considered in that section.

Assumption 3 (lemons).

$$\text{sign} \left(\frac{da_j}{d\rho^\dagger} \right) = -\text{sign} \left(\frac{\partial}{\partial \rho^\dagger} \left(M^+(m^*(a_j), \rho^\dagger) \right) \right).$$

3.2 Examples

Assumptions 2 and 3 are for instance satisfied by the following games:

(a) *Akerlof's lemons game.* Let player i be, say, the seller. The seller can sell his good in the market (“in”) or not sell it (“out”). Player j is then a set of competitive buyers who choose a price a_j equal to the value for a buyer conditional on the good being put in the market. Suppose that the players' utilities from the good are $v_i - m$ for the seller and $v_j - m$ for the buyer, with $v_i < v_j$ (gains from trade). Then, a_j is the price offered by competitive buyers

$$a_j = E\left(v_j - m \mid v_i - m \leq a_j, \rho^\dagger\right) = v_j - M^+(v_i - a_j, \rho^\dagger)$$

as the cutoff $m^*(a_j)$ is equal to $v_i - a_j$; and $\delta_i(m, a_j) = a_j - (v_i - m)$ (thus, $\gamma = 1$ and $\tau = 0$). We assume that the solution a_j is unique, which is indeed the case if the hazard rate of the distribution of m for parameter ρ^\dagger is monotonic.¹⁷ Furthermore Assumption 3 is satisfied.

(b) *Team formation.* We generalize example (a) by allowing player j to take a continuous share in $[0, 1]$ (and not only 0 or 1) in player i 's ownership. Player i has a project. He can associate player j to the project or do it alone. Bringing player j on board creates synergies (lowers the cost of implementation), but forces i to share the gains, which he does not want to do if the project is a good one. Player i 's payoff is $(v - \omega) - d_i$ if he does it alone and $a_j(v - \omega) - c_i$ if it is a joint project, where a_j is the value share left to i by (competitive) player j and $c_i < d_i < v$ is player i 's reduced cost of project implementation. Let c_j denote player j 's cost (with $c_i + c_j < d_i$). Player i chooses “in” if and only if

$$\delta_i(m, a_j) \equiv d_i - c_i - (1 - a_j)(v - m) \geq 0.$$

And so $\gamma = v$ and $\tau = 1$. Finally, $a_j = a_j(\rho^\dagger)$ solves¹⁸:

$$(1 - a_j) \left(v - M^+ \left(v - \frac{d_i - c_i}{1 - a_j}, \rho^\dagger \right) \right) = c_j.$$

Provided that $c_j - \frac{\partial M^+ \left(v - \frac{d_i - c_i}{1 - a_j}, \rho^\dagger \right)}{\partial m^*} (d_i - c_i) > 0$ (e.g. $2c_j > d_i - c_i$ for a uniform distribution), which guarantees a unique solution if it exists (otherwise the team does not form), then Assumption 3 is satisfied.

(c) *Interdependent herding game.* Player i decides on whether to enter a market. Player j , a rival, then decides whether to follow suit. Player j uses the information revealed by player i 's

¹⁷Then $\partial M^+ / \partial m^* \in (0, 1)$. See An (1998).

¹⁸Assuming that the solution to this equation satisfies $a_j(v - \omega^{\max}) \geq c_i$ (here we must assume a finite support).

decision, but in contrast with most herding models, payoffs are interdependent and so externalities are not purely informational. Suppose for instance that i and j are rivals, with per-customer profit π^m under monopoly and $\pi^d < \pi^m$ under duopoly.¹⁹ The state of nature ω here indexes (minus) the fixed cost of entry or opportunity cost of firms i and j . Let a_j denote the probability of non-entry by firm j . Then

$$\delta_i(m, a_j) \equiv \left[a_j \pi^m + (1 - a_j) \pi^d \right] - (k_i - m),$$

where $k_i - m$ is firm i 's entry cost. So $m^*(a_j) = k_i - a_j(\pi^m - \pi^d) - \pi^d$, $\gamma = \pi^m - \pi^d > 0$ and $\tau = 0$; and so Assumption 2 is satisfied.

Firm j has entry cost $k_j - m$, where, say, $k_j \in (0, +\infty)$ with distribution $G(k_j)$. The realization of k_j is unknown to player i . Then $a_j(\rho^\dagger)$ is the solution to

$$a_j = 1 - G\left(\pi^d + M^+(m^*(a_j), \rho^\dagger)\right).$$

Assumption 3 is satisfied whenever the solution to this equation is unique (which is the case if the density g satisfies, at least at the cutoff, $g(\pi^m - \pi^d) < 1$ and the distribution of m satisfies the monotone-hazard-rate condition, which implies that $\partial M^+ / \partial m^* < 1$).

(d) *Marriage game.* Consider the following variant of Spier (1992)'s model, augmented with cognition. Players i and j decide whether to get married. Getting married has value v_i and v_j if all goes well; with probability ω distributed on $[0, 1]$, things will go wrong (divorce), generating utility $v_k - L_k$ for $k = i, j$. The divorce can however be made less painful (utility $v_k - \ell_k$) through a covenant spelling out the outcome in case of divorce, where the losses satisfy $0 < \ell_k < L_k < v_k$ for all k . Adding the covenant costs a fixed $c_k < L_k - \ell_k$ to player k (so it is efficient to add the covenant if the parties want to marry but are certain to divorce). Player i has a v_i high enough that (s)he wants to marry regardless ($v_i \geq L_i$), while player j 's value v_j is distributed on $[\underline{v}_j, +\infty)$ according to c.d.f. G and is private information. Player i may acquire information about ω and then chooses between a contract with ("in") and without ("out") covenant. Player j then decides whether to accept to marry. Assume that $\underline{v}_j - \omega_0 L_j \geq 0$ (so in the absence of any information, player j always accept to marry). Let a_j denote the probability that player j accepts to marry when the proposed contract includes the covenant. This game also satisfies Assumptions 2 and 3. Note first that

$$\delta_i(m, a_j) = a_j[v_i - m\ell_i - c_i] + (1 - a_j) \cdot 0 - (v_i - mL_i)$$

is increasing in m and satisfies $\gamma = v_i - c_i$ and $\tau = \ell_i$.²⁰ Furthermore, $m^* = m^*(a_j(\rho^\dagger))$ is given

¹⁹One can also perform the analysis for complementors, with $\pi^d > \pi^m$.

²⁰The absence of covenant is "good news" about m . And so, $v_j - \omega_0 L_j \geq 0$ for all v_j implies that the contract without covenant is always accepted.

by $m^* [L_i - a_j(\rho^\dagger)\ell_i] = [1 - a_j(\rho^\dagger)] v_i + c_i$.²¹ To see if Assumption 3 holds, note that

$$a_j = 1 - G(M^+(m^*(a_j), \rho^\dagger)\ell_j + c_j);$$

and so Assumption 3 holds under a condition analogous to that found in Example (c).

3.3 Expectation conformity

The monotonicity of δ_i in m implies that player i 's relative payoff of choosing “in” rather than “out” when his information is indexed by ρ and player j expects a choice of ρ^\dagger is

$$V_i(\rho, \rho^\dagger) \equiv \int_{m^*(a_j(\rho^\dagger))}^{+\infty} \delta_i(m, a_j(\rho^\dagger)) dF(m; \rho).$$

(Strict) expectation conformity holds locally provided that

$$\frac{\partial^2 V_i}{\partial \rho \partial \rho^\dagger} > 0 \text{ at } \rho = \rho^\dagger.$$

Proposition 4 (generalized lemons environments). *Under Assumptions 1 through 3:*

(i) *Player j reacts in an unfriendly manner to an increase in anticipated cognition ($da_j/d\rho^\dagger < 0$) if and only if, for the relevant cutoff $m^* = m^*(a_j(\rho^\dagger))$ defined by $\delta_i(m^*, a_j(\rho^\dagger)) = 0$,*

$$A \equiv [M^+(m^*, \rho^\dagger) - m^*] F_\rho(m^*; \rho^\dagger) - \int_{m^*}^{+\infty} F_\rho(m; \rho^\dagger) dm > 0;$$

in particular, it holds whenever $F_\rho(m^; \rho^\dagger) > 0$ at $m^* = m^*(a_j(\rho^\dagger))$.*

(ii) *Let B denote the sensitivity of i 's marginal benefit of cognition to an unfriendly action:*

$$\begin{aligned} B &\equiv - \int_{m^*(a_j(\rho^\dagger))}^{+\infty} \frac{\partial \delta_i}{\partial a_j}(m, a_j(\rho^\dagger)) dF_\rho(m; \rho) \\ &\equiv \frac{\partial \delta_i}{\partial a_j}(m^*(a_j(\rho^\dagger)), a_j(\rho^\dagger)) F_\rho(m^*(a_j(\rho^\dagger)); \rho) + \int_{m^*(a_j(\rho^\dagger))}^{+\infty} \frac{\partial^2 \delta_i}{\partial a_j \partial m}(m, a_j(\rho^\dagger)) F_\rho(m; \rho) dm. \end{aligned}$$

Then EC holds if $AB > 0$. In particular, it holds whenever $F_\rho(m^; \rho) > 0$ at $m^* = m^*(a_j(\rho^\dagger))$.*

(iii) *A sufficient condition for local expectation conformity is therefore that the cutoff lie to the left of the rotation point: $m^*(a_j(\rho^\dagger)) < m_\rho$.*

Proof. Using the condition that $\delta_i(m^*(a_j(\rho^\dagger)), a_j(\rho^\dagger)) = 0$,

$$\frac{\partial^2 V_i}{\partial \rho \partial \rho^\dagger} = \left[\int_{m^*(a_j(\rho^\dagger))}^{+\infty} \frac{\partial \delta_i}{\partial a_j}(m, a_j(\rho^\dagger)) dF_\rho(m; \rho) \right] \frac{da_j}{d\rho^\dagger}.$$

²¹Such a cut-off is guaranteed to exist if $[1 - G(\ell_j + c_j)][v_i - \ell_i - c_i] > v_i - L_i$, since the contract with covenant is always accepted if $v_j \geq \ell_j + c_j$.

(i) From Assumption 3 ,

$$\text{sign} \left(\frac{da_j}{d\rho^\dagger} \right) = -\text{sign} \left[[M^+(m^*, \rho^\dagger) - m^*] F_\rho(m^*; \rho^\dagger) - \int_{m^*}^{+\infty} F_\rho(m; \rho^\dagger) dm \right]$$

Because ρ indexes a mean-preserving spread, $\int_{m^*}^{+\infty} F_\rho(m; \rho^\dagger) dm \leq 0$. And thus $da_j/d\rho^\dagger < 0$ whenever $F_\rho(m^*(a_j(\rho^\dagger)), \rho^\dagger) > 0$.

(ii) Next

$$\begin{aligned} \int_{m^*(a_j(\rho^\dagger))}^{+\infty} \frac{\partial \delta_i}{\partial a_j}(m, a_j(\rho^\dagger)) dF_\rho(m; \rho) &= -\frac{\partial \delta_i}{\partial a_j}(m^*(a_j(\rho^\dagger)), a_j(\rho^\dagger)) F_\rho(m^*(a_j(\rho^\dagger)); \rho) \\ &\quad - \int_{m^*(a_j(\rho^\dagger))}^{+\infty} \frac{\partial^2 \delta_i}{\partial a_j \partial m}(m, a_j(\rho^\dagger)) F_\rho(m; \rho) dm. \end{aligned}$$

The latter term is negative as ρ is an index of mean-preserving spread and $\partial^2 \delta_i / \partial a_j \partial m$ is negative and weakly decreasing in m . The former term is negative provided that $F_\rho > 0$ at $m^*(a_j(\rho^\dagger))$. ■

Corollary 1 (cognitive trap). *Consider two equilibria $\{\rho_1, a_j(\rho_1)\}$ and $\{\rho_2, a_j(\rho_2)\}$ with $\rho_1 < \rho_2$ and assume that $A > 0$ (which is guaranteed if $F_\rho(m^*(a_j(\rho)); \rho) > 0$) for $\rho \in [\rho_1, \rho_2]$. Then player i is better off in the low-cognition equilibrium $\{\rho_1, a_j(\rho_1)\}$.*

Proof: $F_\rho(m^*(a_j(\rho)); \rho) > 0$ on $[\rho_1, \rho_2]$ and part (i) of Proposition 4 ensure that $da_j/d\rho \leq 0$ for $\rho \in [\rho_1, \rho_2]$. For a given a_j , player i 's welfare is $\mathcal{V}(a_j)$ given by

$$\mathcal{V}(a_j) = \max_{\{\rho\}} \left\{ \int_{m^*(a_j)}^{+\infty} \delta_i(m, a_j) dF(m; \rho) - C_i(\rho) \right\}.$$

The envelope theorem and the property that a_j is a friendly action imply that $d\mathcal{V}/da_j > 0$. Because $a_j(\rho_2) \leq a_j(\rho_1)$, player i is better off in the $\{\rho_1, a_j(\rho_1)\}$ equilibrium. ■

3.4 Discussion

(i) *Necessity.* Proposition 4 provides only a sufficient condition for local EC. The condition that the cutoff be to the left of the rotation point ($F_\rho > 0$) is by no means necessary for $A > 0$, that is for an increase in cognition to exacerbate the adverse selection problem. For a number of distributions indeed, $\frac{\partial M^+}{\partial \rho}(m^*, \rho) > 0$ regardless of the value of m^* ; these include:

- the *uniform distribution*: m is uniformly drawn from $[\omega_0 - \rho, \omega_0 + \rho]$. Then for $|m - \omega_0| \leq \rho$, $F(m; \rho) = \frac{1}{2} + \frac{m - \omega_0}{2\rho}$. The rotation point is $m_\rho = \omega_0$ for all ρ and $M^+(m^*, \rho) = \frac{m^* + (\omega_0 + \rho)}{2}$ is increasing in ρ for all ρ .

- the *Pareto distribution*: m is distributed on $[1/\rho, +\infty)$ according to the survival function $1 - F(m; \rho) = (1/\rho m)^{\alpha(\rho)}$ where $\omega_0 = \frac{\alpha(\rho)}{\alpha(\rho)-1} \frac{1}{\rho}$ (changes in ρ are mean preserving). One can check that increases in ρ induce a rotation, with rotation point

$$m_\rho = \frac{e^{\omega_0 \rho - 1}}{\rho} > \omega_0.$$

- the *exponential distribution*: m is distributed on $[1/\rho, +\infty)$ according to the survival function $1 - F(m; \rho) = e^{-\lambda(\rho)(m - (1/\rho))}$ where $\frac{1}{\rho} + \frac{1}{\lambda(\rho)} = \omega_0$ (mean preservation). Then

$$M^+(m^*; \rho) = m^* + \frac{1}{\lambda(\rho)} \quad \Rightarrow \quad \frac{\partial M^+}{\partial \rho} = -\frac{\lambda'(\rho)}{\lambda^2(\rho)} = \frac{1}{\rho^2} > 0,$$

and the rotation point is $m_\rho = \omega_0$.

Corollary 2 (necessary and sufficient condition). *Suppose that $M^+(m^*; \rho)$ is always increasing in ρ (as is the case for the uniform, Pareto and exponential distributions). If furthermore $\partial^2 \delta_i / \partial a_j \partial m = 0$ (as is the case for examples a) and c)), then $m^*(a_j(\rho)) < m_\rho$ is a necessary and sufficient condition for local EC.*

(ii) *Role of gains from trade.* It is interesting to note that EC is more likely to be satisfied when gains from interaction are large.

Corollary 3 (gains from interaction). *Suppose that player i 's gain from playing in is $\delta_i(m, a_j) + \theta$.²² Then if $F_\rho(m^*(a_j(\rho^\dagger), \theta); \rho^\dagger) > 0$ (and so $A(m^*(a_j(\rho^\dagger), \theta), \rho^\dagger) > 0$ and $B(m^*(a_j(\rho^\dagger), \theta), \rho) > 0$ at $\rho = \rho^\dagger$), a fortiori $F_\rho(m^*(a_j(\rho^\dagger), \theta'); \rho^\dagger) \geq 0$ for $\theta' > \theta$ and so local EC prevails as well. So for all ρ , there exists $\theta^*(\rho)$ such that for all $\theta \geq \theta^*(\rho)$, EC prevails locally at ρ .*

(iii) *Cognitive styles.* Another focus of comparative statics concerns the player i 's “cognitive style”. We provide only an informal account. In the generalized lemons model, suppose that the cost-of-cognition function depends on a parameter, interpreted as cognitive ability. A higher ability player i has a lower marginal cost of cognition. As player i 's ability increases, the equilibrium cognition increases (if unique, or in case of multiple equilibria, in the sense of monotone comparative statics: the minimum and maximum of this set both increase). Put differently, player i 's ability, while directly beneficial, indirectly hurts him as player j becomes more wary of adverse selection. This suggests that if player i has side opportunities to signal cognitive ability, he will want to adopt a dumbed-down profile.

Suppose indeed that player i can be bright or dumb. A bright person can demonstrate she is bright (and can always mimick a dumb one), but the reverse is impossible. Then if we add a disclosure game prior to the cognitive game in which player i can disclose she is bright if this is indeed the case, the equilibrium is a pooling one, in which the bright player i does not disclose her IQ. Conversely, player i will disclose, if she can, that she is overloaded with work, and therefore that her marginal cost of investigation is high.

²²For instance, an increase in θ corresponds to a decrease in v_i (example (a)), $c_i - d_i$ (example (b)), $\pi^m - \pi^d$ (example (c)), ℓ_i or $-L_i$ (example (d)).

(iv) *Anti-lemons*. Finally, note that the results can be applied with slight modifications to environments that do not satisfy the assumptions above. Suppose that an increase in anticipated cognition generates a friendly reaction, i.e., Assumption 3 is reversed:

Assumption 3' (anti-lemons).

$$\text{sign} \left(\frac{da_j}{d\rho^\dagger} \right) = \text{sign} \left(\frac{\partial M^+}{\partial \rho^\dagger} \right).$$

Under Assumption 3', local EC requires that $AB < 0$. Suppose that as in the examples above (uniform, Pareto, exponential distributions), $\partial M^+/\partial \rho^\dagger$ is always strictly positive and so $A > 0$. Then under Assumption 2, $B \geq 0$ as long as $F_\rho \geq 0$ and so local EC cannot hold. An illustration satisfying Assumption 2 and Assumption 3' is the warfare game:

(e) *Warfare*. Country i is a potential invader and must decide whether to engage in a fight ("in" action). The state of nature ω here represents the probability that country i wins in case of a fight. Let a_j denote the probability that country j surrenders without fighting back, 1 the payoff in case of victory and c_i the cost in case of defeat:

$$\delta_i(m, a_j) = a_j + (1 - a_j)[m - (1 - m)c_i]$$

and so

$$\gamma = \tau = 1 + c_i \quad \text{and} \quad m^*(a_j) = \frac{c_i - (1 + c_i)a_j}{(1 - a_j)(1 + c_i)}.$$

Letting, similarly, country j 's payoff from victory be equal to 1 and its loss in case of defeat be equal to c_j , country j fights back if and only if

$$[1 - M^+(m^*(a_j(\rho^\dagger)); \rho^\dagger)] - M^+(m^*(a_j(\rho^\dagger)); \rho^\dagger)c_j \geq 0.$$

Assuming that c_j is drawn from some cumulative distribution H ,

$$a_j(\rho^\dagger) = 1 - H \left(-1 + \frac{1}{M^+(m^*(a_j(\rho^\dagger)); \rho^\dagger)} \right).$$

Next, consider games in which $\partial^2 \delta_i / \partial a_j \partial m \geq 0$. Then, assuming again that A is always strictly positive, local EC is satisfied if the cutoff lies to the *right* of the rotation point: $m^* = m^*(a_j(\rho^\dagger)) > m_{\rho^\dagger}$. To see this, recall that (omitting the arguments), $B \equiv \frac{\partial \delta_i}{\partial a_j} F_\rho + \int_{m^*}^{+\infty} \frac{\partial^2 \delta_i}{\partial a_j \partial m} F_\rho dm$. Suppose that $m^* = m^*(a_j(\rho^\dagger)) > m_{\rho^\dagger}$. Then $F_\rho \leq 0$ for all $m \geq m^*$, and so $B < 0$. An illustration satisfying these assumption is the leadership game:

(f) *Leadership game*. Like in Hermalin (1998)'s theory of leadership, a leader shares in the team's output and has information about the return to effort. The information structure is here endogenous. Player i decides to undertake a costly project or not. Her gain from the project depends on whether player j gets on board and on the quality m of the project:

$$\delta_i(m, a_j) = a_j m - c_i.$$

Note that $\partial^2 \delta_i / \partial a_j \partial m = 1$. Player j observes whether i undertakes the project and decides whether to get on board, yielding payoff $m - c_j$, where c_j is drawn from some cumulative distribution H , and so

$$a_j = H(M^+(c_i/a_j))$$

is increasing in M^+ .

Proposition 4' (generalized anti-lemons environment). *Suppose that Assumptions 1 and 3' hold and that $\partial M^+ / \partial \rho^\dagger$ is always strictly positive.*

(i) *If Assumption 2 holds (like in the warfare game), local EC cannot hold if the cutoff is to the left of the rotation point: $m^*(a_j(\rho^\dagger)) \leq m_{\rho^\dagger}$.*

(ii) *If $\partial^2 \delta_i / \partial a_j \partial m \geq 0$ (like in the leadership game), then local EC is satisfied if the cutoff is to the right of the rotation point: $m^*(a_j(\rho^\dagger)) \geq m_{\rho^\dagger}$.*

(v) *Relation to the covert investment game.* Let us draw a formal analogy between the generalized lemons game and the covert investment game described in the introduction. The investment ρ_i in the latter is the cognitive investment in the generalized lemons game. Let us define the “second-stage action” a_i as being just matching this investment: $a_i \equiv \rho_i$.²³ Then one can define

$$\phi_i(a_i, a_j) \equiv \int_{m^*(a_j)}^{+\infty} \delta_i(m, a_j) dF(m; a_i)$$

and so $\partial^2 \phi_i / \partial a_i \partial a_j < 0$ whenever condition $B > 0$ in Proposition 4 is satisfied.

Finally, the exact expression of $\phi_j(a_i, a_j)$ is application-specific, but the condition $A > 0$ in Proposition 4 expresses the condition that $da_j/da_i < 0$, and so strategic substitutability (SS) prevails.

Similarly, part (ii) of Proposition 4' shows that some environments, like under some conditions the leadership game, satisfy the strategic complementarity assumption (SC) of the covert investment model of the introduction.

3.5 Directed and non-directed search in Akerlof's game

We apply these results to Akerlof's lemons game for two further and familiar forms of information acquisition: directed (D) and non-directed (ND) search.²⁴ The objective is not only to study two further information acquisition technologies, but also to provide some intuition about the role of rotations.

(a) *Non directed search.* Assume that information collection follows the standard general or non-directed search technology:

²³Say, $\psi(a_i, \rho_i) = 0$ if $a_i = \rho_i$, $= -\infty$ otherwise, a discontinuous version of the complementarity relationship $\partial^2 \psi_i / \partial a_i \partial \rho_i < 0$ of complementarity between action and cognition.

²⁴We confine attention to soft information, but a similar result holds for hard information. Under hard (i.e., verifiable) information, the seller uses acquired information to disclose to the buyer that the good for sale has a high value; under soft information (the case considered by Akerlof), the seller acquires information to withdraw from the market if the good is very valuable.

$$F(m; \rho) = \begin{cases} \rho Q(m) & \text{for } m < \omega_0 \\ \rho Q(m) + 1 - \rho & \text{for } m \geq \omega_0. \end{cases}$$

That is, the seller learns the true state of nature with probability ρ and nothing with probability $1 - \rho$. Note that the rotation point is equal to the prior mean ω_0 . It is then natural to posit an increasing information acquisition cost $C_i^{ND}(\rho)$.

Proposition 4 states that local EC holds whenever the cutoff m^* is to the left of the prior mean. Expectation conformity thus arises when gains from trade $v_j - v_i$ are large and so $m^* \leq \omega_0$, but not when they are small.²⁵ To grasp intuition about the rotation condition, consider Figure 1(a) and note that in the standard lemons game, the seller learns the value of the car with probability ρ . The seller puts his car for sale when the state is above some m^* : When m^* is smaller than the rotation point, that is the mean ω_0 , then the seller enters the market both when he is uninformed and when she learns that the state m is above m^* . Hence, as ρ^\dagger increases, the expected quality conditional on entry goes down, so the price paid by buyers goes down as well. But then it becomes even more important for the buyer to learn the value of the car, that is, to increase ρ : So local EC is satisfied. When $\omega_0 < m^*$, the seller entering the market only if she learns that $m > m^*$ implies that the value of the car conditional on entry is the same independent of ρ : So a higher ρ^\dagger does not affect the price obtained, and hence does not increase the incentives to search.

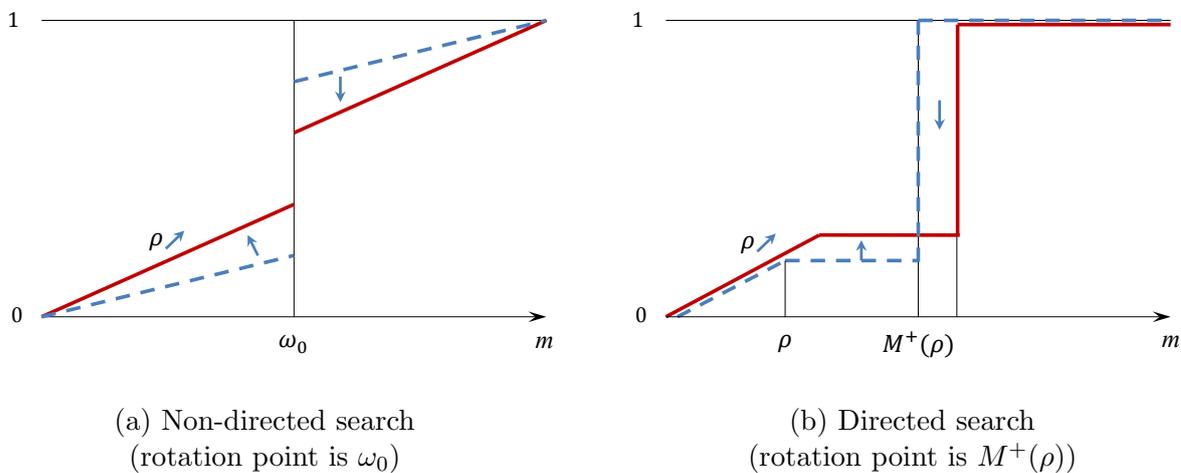


Figure 1: (cumulative distribution $F(m; \rho)$)

²⁵Let

$$H(m^*, \rho^\dagger) \equiv F_\rho(m^*; \rho^\dagger)[M^+(m^*, \rho^\dagger) - m^*] - \int_{m^*}^{+\infty} F_\rho(m, \rho^\dagger) dm \quad \begin{cases} > 0 & \text{if } m^* \leq \omega_0 \\ = 0 & \text{if } m^* > \omega_0. \end{cases}$$

To show that $H = 0$ on $[\omega_0, +\infty)$, note that $H(+\infty) = 0$ and that $dH/dm^* = 0$ on this domain. Also, $m^* - M^+(m^*, \rho^\dagger) = -(v_j - v_i)$. Because $\partial M^+ / \partial m^* \in (0, 1)$ under a monotone hazard rate, m^* decreases with $(v_j - v_i)$, i.e., with the gains from trade. For $m^* > \omega_0$, $M^+(m^*, \rho^\dagger)$ is invariant to m^* .

(b) *Directed search.* Next, suppose that search is directed. At cost $C_i^D \left(\int_{-\infty}^{+\infty} e_i^\omega dQ(\omega) \right)$, a strictly increasing and convex cost function the seller can learn ω with probability $e_i^\omega \in \{0, 1\}$; with probability $1 - e_i^\omega$, the seller receives no signal. It is easy to show that it is optimal for the seller to adopt a cutoff strategy:

$$e_i^\omega = \begin{cases} 1 & \text{if } \omega \leq \rho \\ 0 & \text{if } \omega > \rho \end{cases}$$

for some ρ , which is a measure of player i 's cognitive effort. In this optimal class,

$$Q(m; \rho) = \begin{cases} Q(m) & \text{for } m \leq \rho \\ Q(\rho) & \text{for } \rho < m < M^+(\rho) \\ 1 & \text{for } m \geq M^+(\rho). \end{cases}$$

The rotation point is $M^+(\rho)$. In equilibrium $v_j - a_j = M^+(\rho^\dagger)$. Because cognition is costly, a necessary condition to learn about the state is that the trade be disadvantageous for the seller and so $v_i - \rho > v_j - M^+(\rho^\dagger)$. One can take $m^* = m^*(a_j(\rho^\dagger)) = \rho^\dagger$ and so $m^* < M^+(\rho^\dagger) - (v_j - v_i) < M^+(\rho^\dagger)$. So m^* is below the rotation point:

$$F_\rho(m^*; \rho^\dagger) > 0.$$

Thus the lemons game always satisfies expectation conformity under directed search.²⁶

Under directed search, illustrated in Figure 1(b), the seller focuses attention on the high-value states so as to refrain from putting his car on the market in those states and puts his car for sale otherwise. If the cutoff belief exceeded the rotation point, the seller at that cutoff would be better strictly off selling as there are gains from trade; so it is not rational to acquire information in such states: The seller might as well sell without acquiring information, a contradiction.

Proposition 5 (lemons game). *The cognition-augmented lemons game always satisfies expectation conformity under directed search, and under non-directed search satisfies it if and only if the gains from trade are sufficiently large.*

The general intuition why the lemons game often satisfies expectation conformity goes as follows: Suppose that the seller is expected to engage in a high level of cognition; then adverse

²⁶One can combine directed and non-directed search in the following way: Player i 's information collection proceeds in two sub-stages in which search is first non-directed and then directed. First, the player performs some preliminary, *general-purpose search* to try to apprehend the context; this search costs $C_i^G(\rho_i^G)$ and succeeds with probability ρ_i^G . If the first stage is unsuccessful, the search process stops. If the general-purpose search is successful, the player can engage in *directed search* and pick the probability $e_i^\omega \in [0, 1]$ of learning that the state is ω ; this latter search costs $C_i^D(\rho_i^D; \rho_i^G)$ where $\rho_i^D \equiv \int_{-\infty}^{+\infty} e_i^\omega dQ(\omega)$. In general, $\partial C_i^D / \partial \rho_i^D$ can be independent of ρ_i^G , decrease with ρ_i^G (benefits of acquired knowledge) or increase with ρ_i^G (fatigue, time constraints). Information structures can easily be ranked whenever directed and non-directed search efforts are "weak complements", i.e., if the cross-partial derivative of C_i^D is non-positive (an increase in general search does not discourage directed search).

selection is a serious concern for the buyers, who are therefore willing to pay only a low price. A low price in turn makes it particularly costly for the seller to part with a valuable item, raising his incentives to acquire information.

A similar insight holds when it is the *buyer* who engages in cognition (is player i). Suppose that there is a single buyer, who may acquire information. If there is a single, price-setting seller, the model is not the mirror image of the one considered so far. The treatment is slightly different as price setting involves market power. Because our objective here is merely to show that similar considerations apply to buyer information acquisition, suppose instead that sellers (player j now) are competitive. The price offered by the sellers is $-a_j$. The buyer has utility 0 when not buying and $\delta_i = v_i + m + a_j$ when buying. Think of m as (minus) the scope for ex-post holdup by the selected seller, where the holdup comes from a renegotiation following the discovery that the initial specification is not optimal for the buyer. So $\gamma = 1$, $\tau = 0$ and $m^* = -v_i - a_j$. A representative seller's utility is $v_j - m - a_j$, and so the competitive outcome is $a_j = v_j - M^+(m^*; \rho^\dagger)$. The results apply to buyer information acquisition as well.

The intuition for expectation conformity is again straightforward. Suppose that the seller anticipates more cognition; he then raises price to reflect the fact that the buyer has ruled out more bad news for himself. Facing a higher price, the buyer then finds it more costly to enter disadvantageous deals and thus is incentivized to find out about possible bad news.

4 Concluding remarks

Economic agents manage their information in multiple ways: allocation of scarce cognitive resources, brainstorming, search and experimentation, hiring of engineering, financial or legal experts. Such “cognitive activities” determine information structures and are often the essence of adverse selection; they thereby condition the functioning of contracts and markets, and more broadly of social interactions. This motivates the study of “cognitive games”, defined as games in which a normal- or extensive-form game is preceded by players’ selecting their or their rivals’ information structures.

Expectation conformity arises when players have an incentive to comply with the level of cognition they are expected to engage in. We showed that games of pure conflict (zero-sum games) never give rise to self-fulfilling cognition while games of pure alignment (coordination games) always do. We then considered a generalized lemons (or anti-lemons) environment, which comprises many games of interest to economists such as the cognition-augmented lemons model. A characterization of the expectation conformity property in terms of rotation points was obtained for this class of games.

Because of their importance for economics, cognitive games need to be better understood and there are multiple alleys for future research. For instance, we have assumed that cognition is unobservable; one would like to investigate how its equilibrium level is affected by the ability, if any, to disclose its intensity to other players. Relatedly, cognition may occur in multiple stages as an extensive form game unfolds. In multi-stage cognition, players may learn with an endogenous lag about their rivals’ choice of cognitive strategies.

Cognitive strategies are particularly relevant to contracting environments. In such environments, parties may acquire information not only to reach more efficient agreements, but also to either design self-advantageous covenants or, as in the lemons game, eschew undesirable trades. Contracts may then be too complete or too incomplete from a social welfare viewpoint, where incompleteness refers to the need for ex-post adjustments or more generally to inefficiencies generated by imperfect contracting. The relationship between EC, cognitive traps and excess cognition in contracting environments is developed in a follow-up paper (Tirole 2015).

We have assumed that players choose their own information structure, Appendix B considers “signal-jamming cognition games”, in which players choose their rival’s information structure rather than their own. Its main purpose is to adapt the definition of expectation conformity to this context. It also provides various economic environments satisfying expectation conformity.

Finally, and as developed further in Appendix C, expectation conformity results from a combination of effects when multiple players acquire information: increasing differences (the standard form of strategic complementarity in information structure choices, when these are publicly observable), unilateral expectation conformity (a player’s incentive to conform to the other players’ expectation of her cognitive activities) and the impact of the rivals’ information on a player’s preferred perception of his information by the rivals. While some of these effects but not others are at play in each environment, a better understanding of their relative importance and of when they are likely to hold would bring a deeper understanding of cognitive games and cognitive traps. We leave these and the many other questions related to cognitive games to future research.

References

- Amir, R., and N. Lazzati (2010) “Endogenous Information Acquisition on Bayesian Games with Strategic Complementarities,” mimeo, University of Arizona.
- An, M.Y. (1998) “Logconcavity versus Logconvexity: A Complete Characterization,” *Journal of Economic Theory*, 80(2): 350–369.
- Angeletos, G.M., and A. Pavan (2007), “Efficient Use of Information and Social Value of Information,” *Econometrica* 75: 1103–1142.
- Bénabou, R. (2013) “Groupthink: Collective Delusions in Organizations and Markets,” *Review of Economics Studies*, 80: 429–462.
- Bénabou, R., and J. Tirole (2002) “Self-Confidence and Personal Motivation,” *Quarterly Journal of Economics*, 117(3): 871–915.
- Bénabou, R., and J. Tirole (2006) “Belief in a Just World and Redistributive Politics,” *Quarterly Journal of Economics*, 121: 699–746.
- Bolton, P., and A. Faure-Grimaud (2010) “Satisficing Contracts,” *Review of Economic Studies*, 77(3): 937–971.
- Bulow, J., Geanakoplos, J. and P. Klemperer (1985) “Multimarket Oligopoly: Strategic Substitutes and Complements,” *Journal of Political Economy*, 93(3): 488–511.
- Che, Y.K., and N. Kartik (2009) “Opinions as Incentives,” *Journal of Political Economy*, 117(5): 815–860.
- Colombo, I., Femminis, G., and A. Pavan (2014) “Information Acquisition and Welfare,” *Review of Economic Studies*. DOI: 10.1093/restud/rdu015 First published online: May 9, 2014.
- Dang, T.V. (2008) “Bargaining with Endogenous Information,” *Journal of Economic Theory*, 140: 339–354.
- Dang, T.V., Gorton, G. and B. Holmström (2011) “The Information Acquisition Sensitivity of a Security,” mimeo.
- Dessi, R. (2008) “Collective Memory, Cultural Transmission and Investments,” *American Economic Review*, 98(1): 534–560.
- Dewatripont, M., Jewitt, I., and J. Tirole (1999) “The Economics of Career Concerns. I: Comparison of Information Structures,” *Review of Economic Studies*, 66 (1): 183–198.
- Dewatripont, M., and J. Tirole (2005) “Modes of Communication,” *Journal of Political Economy*, 113: 1217–1238.
- Diamond, P. and J. Stiglitz (1974) “Increases in Risk and Risk Aversion,” *Journal of Economic Theory*, 8: 337–360.

- Farhi, E., and J. Tirole (2015) “Liquid Bundles,” *Journal of Economic Theory*, 158(B): 634–655.
- Fudenberg, D., and J. Tirole (1984) “The Fat-Cat Effect, the Puppy-Dog Ploy, and the Lean and Hungry Look,” *American Economic Review*, 74(2): 361–366.
- Gottlieb, D. (2014a) “Imperfect Memory and Choice under Risk,” *Games and Economic Behavior*, 85: 127–185.
- Gottlieb, D. (2014b) “Will you Never Learn? Self-Deception and Biases in Information Processing,” mimeo, Wharton School, University of Pennsylvania.
- Hellwig, C., and L. Veldkamp (2009) “Knowing What Others Know: Coordination Motives in Information Acquisition,” *Review of Economic Studies*, 76(1): 223–251.
- Hermalin, B. (1998) “Toward an Economic Theory of Leadership: Leading by Example,” *American Economic Review*, 88: 1188–1206.
- Holmström B. (1999) “Managerial Incentive Problems: A Dynamic Perspective,” *Review of Economic Studies*, 66(1): 169–182.
- Johnston, J., and D. Myatt (2006) “On the Simple Economics of Advertising, Marketing, and Product Design,” *American Economic Review*, 96(3): 756–784.
- Lehrer, E., and D. Rosenberg (2006) “What Restrictions do Bayesian Games Impose on the Value of Information?,” *Journal of Mathematical Economics*, 42: 343–357.
- Lester, B., Postlewaite, A., and R. Wright (2012) “Information, Liquidity, Asset Prices, and Monetary Policy,” *Review of Economic Studies*, 79 (3): 1209–1238.
- Llosa, L. G., and V. Venkateswaran (2012) “Efficiency under Endogenous Information Choice,” mimeo Penn State.
- Maćkowiak, B., and M. Wiederholt (2009) “Optimal Sticky Prices under Rational Inattention,” *American Economic Review*, 99(3): 769–803.
- Matejka, F., and A. McKay (2012) “Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model,” *American Economic Review*, 105(1): 272–298.
- Milgrom, P. and C. Shannon (1994) “Monotone Comparative Statics,” *Econometrica*, 62: 157–180.
- Morris, S. and H. Shin (2002) “Social Value of Public Information,” *American Economic Review*, 92: 1521–1534.
- Myatt, D. P., and C. Wallace (2012) “Endogenous Information Acquisition in Coordination Games,” *Review of Economic Studies*, 79: 340–374.
- Pavan, A. (2014) “Attention, Coordination, and Bounded Recall,” mimeo, Northwestern University.

- Persico, N. (2000) "Information Acquisition in Auctions," *Econometrica*, 68(1): 135–148.
- Sims, C. (2003) "Implications of Rational Inattention," *Journal of Monetary Economics*, 50(3): 665–690.
- Spier, K (1992) "Incomplete Contracts and Signaling," *Rand Journal of Economics*, 23: 432–443.
- Stigler, G. (1961) "The Economics of Information," *Journal of Political Economy*, 69: 213–225.
- Tirole, J. (2009) "Cognition and Incomplete Contracts," *American Economic Review*, 99(1): 265–294.
- Tirole, J. (2015) "Cognition-intensive Contracting," mimeo, Toulouse School of Economics.
- Yang, M. (2013) "Optimality of Debt under Flexible Information Acquisition," working papers, Duke University.

Appendix A. Proof of Proposition 3

Define $\omega_i \equiv E(\omega|\mathcal{F}_i)$ and $\sigma_i^2 \equiv E((\omega - \omega_i)^2|\mathcal{F}_i)$.

For any $\mathcal{F} = (\mathcal{F}_i, \mathcal{F}_j)$ such that $\mathcal{F}_j \subseteq \mathcal{F}_i$, best responses of the stage-2 game are:

$$a_j \equiv a_j(\mathcal{F}) = \frac{E(a_i(\mathcal{F})|\mathcal{F}_j) + \omega_j}{2} \quad \text{and} \quad a_i \equiv a_i(\mathcal{F}) = \frac{a_j + \omega_i}{2}$$

hence

$$a_j = \omega_j \quad \text{and} \quad a_i = \frac{\omega_j + \omega_i}{2}$$

Expected gross payoffs conditional on \mathcal{F}_i , resp. \mathcal{F}_j , are:

$$\frac{-(\omega_j - \omega_i)^2}{2} - \sigma_i^2 \quad \text{and} \quad U_j(\sigma_i, \sigma_j, \mathcal{F}_j) = -\frac{E((\omega_j - \omega_i)^2|\mathcal{F}_j)}{4} - \sigma_j^2$$

while ex-ante gross payoffs, defined as in the general model, are:

$$-E_\omega \left[\frac{(\omega_j - \omega_i)^2}{2} + \sigma_i^2 \right] \quad \text{for player } i \quad \text{and} \quad -E_\omega \left[\frac{(\omega_j - \omega_i)^2}{4} + \sigma_j^2 \right] \quad \text{for player } j.$$

Consider two information structures \mathcal{F} and $\widehat{\mathcal{F}}$. Let

$$\Delta_i \equiv V_i(\widehat{\mathcal{F}}_i; \mathcal{F}_i, \mathcal{F}_j) - V_i(\mathcal{F}_i; \mathcal{F}_i, \mathcal{F}_j)$$

and

$$\widehat{\Delta}_i \equiv V_i(\widehat{\mathcal{F}}_i; \widehat{\mathcal{F}}_i, \widehat{\mathcal{F}}_j) - V_i(\mathcal{F}_i; \widehat{\mathcal{F}}_i, \widehat{\mathcal{F}}_j).$$

Case 1 : $\mathcal{F}_1 \subseteq \widehat{\mathcal{F}}_1 \subseteq \mathcal{F}_2 \subseteq \widehat{\mathcal{F}}_2$

$$\Delta_2 = E \left[\sigma_2^2 - \widehat{\sigma}_2^2 - \frac{1}{2} [(\omega_1 - \widehat{\omega}_2)^2 - (\omega_1 - \omega_2)^2] \right] = E \left[\sigma_2^2 - \widehat{\sigma}_2^2 - \frac{(\widehat{\omega}_2 - \omega_2)^2}{2} \right]$$

$$\widehat{\Delta}_2 = E \left[\sigma_2^2 - \widehat{\sigma}_2^2 - \frac{1}{2} [(\widehat{\omega}_1 - \widehat{\omega}_2)^2 - (\widehat{\omega}_1 - \omega_2)^2] \right]$$

$$\widehat{\Delta}_2 - \Delta_2 = 0$$

Intuitively, the coordination ability is the same for player 2 regardless of whether his information structure is \mathcal{F}_2 or $\widehat{\mathcal{F}}_2$. The only gain from being better informed comes from a better adjustment to the state of nature and is independent of player 1's information structure.

$$\Delta_1 = E \left[\sigma_1^2 - \widehat{\sigma}_1^2 - \frac{1}{4} [(\omega_1 + \omega_2 - 2\widehat{\omega}_1)^2 - (\omega_1 - \omega_2)^2] \right] = E [\sigma_1^2 - \widehat{\sigma}_1^2]$$

$$\widehat{\Delta}_1 = E \left[\sigma_1^2 - \widehat{\sigma}_1^2 - \frac{1}{4} [(\widehat{\omega}_1 - \widehat{\omega}_2)^2 - (\widehat{\omega}_1 + \widehat{\omega}_2 - 2\omega_1)^2] \right] = E [\sigma_1^2 - \widehat{\sigma}_1^2 + (\widehat{\omega}_1 - \omega_1)^2]$$

$$\widehat{\Delta}_1 - \Delta_1 = E(\omega_1 - \widehat{\omega}_1)^2$$

Case 2: $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \widehat{\mathcal{F}}_1 \subseteq \widehat{\mathcal{F}}_2$

$$\begin{aligned}\Delta_2 &= E \left[\sigma_2^2 - \hat{\sigma}_2^2 - \frac{(\widehat{\omega}_2 - \omega_2)^2}{2} \right] \\ \widehat{\Delta}_2 &= E \left[\sigma_2^2 - \hat{\sigma}_2^2 - \frac{1}{2} [(\widehat{\omega}_2 - \widehat{\omega}_1)^2 - 2(\omega_2 - \widehat{\omega}_1)^2] \right] \\ \widehat{\Delta}_2 - \Delta_2 &= \frac{3}{2} E(\omega_2 - \widehat{\omega}_1)^2 \\ \Delta_1 &= E \left[\sigma_1^2 - \hat{\sigma}_1^2 + \frac{1}{4}(\omega_1 - \omega_2)^2 - \frac{1}{2} \left(\widehat{\omega}_1 - \frac{\omega_1 + \omega_2}{2} \right)^2 \right] \\ \widehat{\Delta}_1 &= E [\sigma_1^2 - \hat{\sigma}_1^2] + E(\widehat{\omega}_1 - \omega_1)^2 \\ \widehat{\Delta}_1 - \Delta_1 &= \frac{3}{4} E(\omega_1 - \omega_2)^2 + \frac{1}{2} E \left(\widehat{\omega}_1 - \frac{\omega_1 + \omega_2}{2} \right)^2 + E(\widehat{\omega}_1 - \omega_2)^2\end{aligned}$$

Case 3: $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \widehat{\mathcal{F}}_2 \subseteq \widehat{\mathcal{F}}_1$

$$\begin{aligned}\Delta_2 &= E \left[\sigma_2^2 - \hat{\sigma}_2^2 - \frac{(\widehat{\omega}_2 - \omega_2)^2}{2} \right] \\ \widehat{\Delta}_2 &= E [\sigma_2^2 - \hat{\sigma}_2^2 + (\widehat{\omega}_2 - \omega_2)^2] \\ \widehat{\Delta}_2 - \Delta_2 &= \frac{3}{2} E(\widehat{\omega}_2 - \omega_2)^2 \\ \Delta_1 &= E \left[\sigma_1^2 - \hat{\sigma}_1^2 + \frac{1}{4}(\omega_1 - \omega_2)^2 - \frac{1}{2} \left(\widehat{\omega}_1 - \frac{\omega_1 + \omega_2}{2} \right)^2 \right] \\ \widehat{\Delta}_1 &= E \left[\sigma_1^2 - \hat{\sigma}_1^2 + (\omega_1 - \widehat{\omega}_2)^2 - \frac{1}{2} (\widehat{\omega}_1 - \widehat{\omega}_2)^2 \right] \\ \widehat{\Delta}_1 - \Delta_1 &= E(\widehat{\omega}_2 - \omega_2)^2 + \frac{3}{4} E(\omega_2 - \omega_1)^2 + \frac{1}{2} E \left(\widehat{\omega}_2 - \frac{\omega_1 + \omega_2}{2} \right)^2\end{aligned}$$

Appendix B. Signal-jamming cognitive games

The paper has assumed that players choose their *own* information structure. In a number of economic games, though, players choose *their opponents'* information structure. Such signal jamming has been studied for example in industrial organization, as when a firm secretly cuts its price so as to convince its rivals that demand is low and induce their exit. Furthermore, cognitive traps are common in such games as well as we will shortly observe.

Defining expectation conformity in signal-jamming cognitive games Consider a two-player game. In a signal-jamming cognitive game, player i chooses player j 's information structure \mathcal{F}_j at cost $C_i(\mathcal{F}_j)$. We have to be a bit careful with regards to measurability, as a deviation from \mathcal{F}_j to $\widehat{\mathcal{F}}_j$ is not observed by player j . Thus, think of \mathcal{F}_j as a conditional distribution $q_j(s_j|\omega)$ over the signal s_j received by player j in state of nature ω . Player j then plays a stage-2 (mixed)

strategy $\alpha_j(s_j)$. Player j 's stage-2 strategy under \mathcal{F}_j is then an “ \mathcal{F}_j -measurable” strategy $\sigma_j^{\mathcal{F}_j}$, defined by:

$$\sigma_j^{\mathcal{F}_j}(\omega) = \sum_{s_j} q_j(s_j|\omega) \alpha_j(s_j).$$

Let $\{\mathcal{F}_i, \mathcal{F}_j\}$ denote a common-knowledge choice of information structures with signal distributions $q_i(s_i|\omega)$ and $q_j(s_j|\omega)$ and $\{\alpha_i, \alpha_j\}$ denote the corresponding equilibrium strategies. Suppose that player i deviates and picks information structure $\widehat{\mathcal{F}}_j$ (corresponding to contingent signal distribution $\widehat{q}_j(s_j|\omega)$). Let

$$V_i(\widehat{\mathcal{F}}_j; \mathcal{F}_i, \mathcal{F}_j) \equiv \max_{\{\alpha'_i(\cdot)\}} \left\{ \sum_{\omega, s_i, s_j} q(\omega) q_i(s_i|\omega) \widehat{q}_j(s_j|\omega) u_i(\alpha'_i(s_i), \alpha_j(s_j), \omega) \right\}.$$

Definition 3 (expectation conformity under signal jamming). $EC_{\{\mathcal{F}, \widehat{\mathcal{F}}\}}$ is satisfied if for all i ,

$$V_i(\widehat{\mathcal{F}}_j; \mathcal{F}_i, \mathcal{F}_j) - V_i(\mathcal{F}_j; \mathcal{F}_i, \mathcal{F}_j) \leq V_i(\widehat{\mathcal{F}}_j; \widehat{\mathcal{F}}_i, \widehat{\mathcal{F}}_j) - V_i(\mathcal{F}_j; \widehat{\mathcal{F}}_i, \widehat{\mathcal{F}}_j). \quad (2)$$

Examples of signal-jamming games satisfying expectation conformity One-sided²⁷ signal-jamming environments (described rather informally below) exhibiting expectation conformity include:

(a) *Imperfect persuasion.* Consider the trading game when the seller with strictly positive probability knows the buyer's willingness to pay. For simplicity, suppose that the state of nature can take one of two values $\{\omega, \widehat{\omega}\}$ and is the buyer's utility, with $\omega > \widehat{\omega}$ and that the seller does not value the good. By exerting more effort, the seller can increase the probability that the buyer understands the argument and thereby learns the true state of nature: information is “semi-hard” in that it can be disclosed, but the amount of disclosure depends on the seller's communication effort.²⁸ The seller's effort (understood as the effort incurred prior to actual communication with the buyer) increases the probability that the buyer identifies the true state and is unobserved by the buyer. Clearly, the seller exerts no effort if the state is $\widehat{\omega}$. By contrast, convincing the buyer that the state is ω is profitable and so in general elicits effort.

It is easy to check that cognitive traps quite similar to those for the lemons game arise naturally: If the seller is expected by the buyer to exert substantial effort to communicate that the state is ω , the price p in the absence of persuasion is low (the state of nature is unlikely to be ω), and then it is particularly profitable for the seller to convince the buyer that the state is ω . In case of multiplicity, the seller is better off in a lower-effort equilibrium.

(b) *Career concerns.* In Holmström (1999)'s celebrated career-concerns model, an agent's current performance depends on talent, effort and noise. The agent does not know her talent and tries to convince future employers that she is talented by secretly exerting more effort to boost current performance. The signal jamming cost is here the cost of effort in the current task. When talent and effort are complements, such signal jamming often generates information conformity and traps (e.g., Dewatripont et al 1999). Indeed, suppose that the labor market expects a higher effort; then employers put more weight on performance when updating their beliefs about talent, as performance is more informative about talent. The increased performance sensitivity of future compensation then boosts the agent's incentive to exert effort. Again, in case of multiplicity, the agent is better off in the low-effort equilibrium.

(c) *Memory management game.* Another class of signal-jamming games giving rise to expect-

²⁷That is, only one player, player i , manipulates the other player's information structure: $\widehat{\mathcal{F}}_i = \mathcal{F}_i$ in condition (2). Again, multilateral and unilateral expectation conformity coincide in such environments.

²⁸In this simplified model, only the seller exerts effort; in general communication involves moral hazard in team (see Dewatripont and Tirole 2005).

tation conformity is the class of memory-management games.²⁹ This class of games describes situations in which a player receives information that he may try to remember or repress. The individual may find himself in a self-trap, in which repression or cognitive discipline are possible self-equilibria with distinct welfare implications.

Appendix C. Expectation conformity, increasing differences and equilibrium multiplicity

The expectation-conformity condition, while itself an increasing differences conditions, should not be mistaken for the standard increasing differences condition that plays a fundamental role in monotone comparative statics. The essential difference between the two concepts is that player i 's payoff in EC depends on his information and on player j 's *anticipation* of his information. More formally, the increasing differences condition, applied to ordered information structures, takes the following form (we use the two-player version for illustrative purposes): If $\widehat{\mathcal{F}}_k$ is finer than \mathcal{F}_k for all k , then

$$V_i(\widehat{\mathcal{F}}_i; \widehat{\mathcal{F}}_i, \mathcal{F}_j) - V_i(\mathcal{F}_i; \mathcal{F}_i, \mathcal{F}_j) \leq V_i(\widehat{\mathcal{F}}_i; \widehat{\mathcal{F}}_i, \widehat{\mathcal{F}}_j) - V_i(\mathcal{F}_i; \mathcal{F}_i, \widehat{\mathcal{F}}_j). \quad (ID)$$

That is, expectation conformity reflects the fact that the players do not observe each other's choice of information structure while the increasing differences condition posits that information structures are common knowledge at stage 2 *on and off* the equilibrium path. Put differently, expectation conformity and increasing differences capture strategic complementarities in information acquisition, under *covert* acquisition for the former and *overt* acquisition for the latter.

Let us investigate the difference between unilateral and multilateral expectation conformity and their relationship with increasing differences a bit further. Let us decompose the difference between the RHS and the LHS of $EC_{\{\mathcal{F}, \widehat{\mathcal{F}}\}}$,

$$\Gamma_i^{EC}(\mathcal{F}, \widehat{\mathcal{F}}) \equiv \left[V_i(\widehat{\mathcal{F}}_i; \widehat{\mathcal{F}}_i, \widehat{\mathcal{F}}_j) - V_i(\mathcal{F}_i; \widehat{\mathcal{F}}_i, \widehat{\mathcal{F}}_j) \right] - \left[V_i(\widehat{\mathcal{F}}_i; \mathcal{F}_i, \mathcal{F}_j) - V_i(\mathcal{F}_i; \mathcal{F}_i, \mathcal{F}_j) \right]$$

into three terms:³⁰

$$\Gamma_i^{UEC}(\mathcal{F}, \widehat{\mathcal{F}}) \equiv \left[V_i(\widehat{\mathcal{F}}_i; \widehat{\mathcal{F}}_i, \mathcal{F}_j) - V_i(\mathcal{F}_i; \widehat{\mathcal{F}}_i, \mathcal{F}_j) \right] - \left[V_i(\widehat{\mathcal{F}}_i; \mathcal{F}_i, \mathcal{F}_j) - V_i(\mathcal{F}_i; \mathcal{F}_i, \mathcal{F}_j) \right]$$

$$\Gamma_i^{ID}(\mathcal{F}, \widehat{\mathcal{F}}) \equiv \left[V_i(\widehat{\mathcal{F}}_i; \widehat{\mathcal{F}}_i, \widehat{\mathcal{F}}_j) - V_i(\mathcal{F}_i; \mathcal{F}_i, \widehat{\mathcal{F}}_j) \right] - \left[V_i(\widehat{\mathcal{F}}_i; \widehat{\mathcal{F}}_i, \mathcal{F}_j) - V_i(\mathcal{F}_i; \mathcal{F}_i, \mathcal{F}_j) \right]$$

and

$$\Gamma_i^P(\mathcal{F}, \widehat{\mathcal{F}}) \equiv \left[V_i(\mathcal{F}_i; \widehat{\mathcal{F}}_i, \mathcal{F}_j) - V_i(\mathcal{F}_i; \mathcal{F}_i, \mathcal{F}_j) \right] - \left[V_i(\mathcal{F}_i; \widehat{\mathcal{F}}_i, \widehat{\mathcal{F}}_j) - V_i(\mathcal{F}_i; \mathcal{F}_i, \widehat{\mathcal{F}}_j) \right],$$

and so

$$\Gamma_i^{EC}(\mathcal{F}, \widehat{\mathcal{F}}) = \Gamma_i^{UEC}(\mathcal{F}, \widehat{\mathcal{F}}) + \Gamma_i^{ID}(\mathcal{F}, \widehat{\mathcal{F}}) + \Gamma_i^P(\mathcal{F}, \widehat{\mathcal{F}}).$$

The term Γ_i^{EC} takes player j 's information structure \mathcal{F}_j as given and measures the increase in i 's incentive to acquire information $\widehat{\mathcal{F}}_i$ rather than \mathcal{F}_i when player j anticipates this change in i 's information acquisition strategy. The difference Γ_i^P measures a pure perception effect: fixing player i 's actual cognition \mathcal{F}_i , it represents how i 's gain (or loss) of being perceived as

²⁹Introduced in Bénabou-Tirole (2002). See also Gottlieb (2014a,b). Dessi (2008) applies similar ideas in the context of cultural transmission with multiple agents. Bénabou (2013) and Bénabou-Tirole (2006) show how memory management and collective decisions interact to produce collective delusions.

³⁰"UEC" stands for "unilateral expectation conformity".

acquiring information $\widehat{\mathcal{F}}_i$ rather than \mathcal{F}_i is affected by player j 's actual cognition (\mathcal{F}_j vs. $\widehat{\mathcal{F}}_j$). To compare multilateral and unilateral expectation conformity, one can aggregate Γ_i^{ID} and Γ_i^P into a “strategic interaction” term

$$\Gamma_i^{SI}(\mathcal{F}, \widehat{\mathcal{F}}) \equiv \left[V_i(\widehat{\mathcal{F}}_i; \widehat{\mathcal{F}}_i, \widehat{\mathcal{F}}_j) - V_i(\mathcal{F}_i; \widehat{\mathcal{F}}_i, \widehat{\mathcal{F}}_j) \right] - \left[V_i(\widehat{\mathcal{F}}_i; \widehat{\mathcal{F}}_i, \mathcal{F}_j) - V_i(\mathcal{F}_i; \widehat{\mathcal{F}}_i, \mathcal{F}_j) \right].$$

That is, keeping j 's expectation about i 's information constant, $\Gamma_i^{SI}(\mathcal{F}, \widehat{\mathcal{F}})$ depicts the increase in i 's incentive to acquire $\widehat{\mathcal{F}}_i$ rather than \mathcal{F}_i when j 's information changes from \mathcal{F}_j to $\widehat{\mathcal{F}}_j$. There is *positive strategic interaction* for information structures $(\mathcal{F}, \widehat{\mathcal{F}})$ if $\Gamma_i^{SI}(\mathcal{F}, \widehat{\mathcal{F}}) \geq 0$.

Note that for the prominent class of *one-sided cognitive games* (only player i , say, acquires information, i.e., Ψ_j is a singleton), $\Gamma_i^{ID} = \Gamma_i^P = 0$ and so

$$\Gamma_i^{EC}(\mathcal{F}, \widehat{\mathcal{F}}) = \Gamma_i^{UEC}(\mathcal{F}, \widehat{\mathcal{F}}).$$

To illustrate the fact that EC can arise from ID rather than from unilateral EC, consider a *matching model* (reminiscent of Lester et al's 2012 model of asset liquidity based on recognizability) in which players may invest in recognizing what's in it for them in a given partnership; that is, each potential match is characterized by a surplus for player i , say 1 if the partner is adequate and a highly negative number otherwise, and so a match occurs only if both can ascertain it is a good one for them. An information structure for player i here is the probability $\rho_i \in [0, 1]$ that player i gets informed (at some effort cost $C_i(\rho_i)$) about what he will derive from the match: $\mathcal{F} \equiv (\rho_i, \rho_j)$ and $\widehat{\mathcal{F}} \equiv (\widehat{\rho}_i, \widehat{\rho}_j)$. At stage 2, players 1 and 2 each have a veto right on the two players' matching. Each player's stage-2 behavior is independent of his expectation about the other player's information: A player who knows he receives 1 from the match accepts to match; one who either is uninformed or knows he receives a negative surplus does not accept the match. In this matching game, $\Gamma_i^{UEC}(\mathcal{F}, \widehat{\mathcal{F}}) = [\widehat{\rho}_i \rho_j - \rho_i \rho_j] - [\widehat{\rho}_i \rho_j - \rho_i \rho_j] = 0$, and so expectation conformity is not strictly satisfied. Similarly, $\Gamma_i^P(\mathcal{F}, \widehat{\mathcal{F}}) = 0$, and so

$$\Gamma_i^{EC}(\mathcal{F}, \widehat{\mathcal{F}}) = \Gamma_i^{ID}(\mathcal{F}, \widehat{\mathcal{F}}).$$

By contrast, $\Gamma_i^{ID}(\mathcal{F}, \widehat{\mathcal{F}}) = (\widehat{\rho}_i - \rho_i)(\widehat{\rho}_j - \rho_j)$, capturing the standard strategic complementarities that are conducive to equilibrium multiplicity.