Knowing your Lemon before you Dump It*

Alessandro Pavan† and Jean Tirole‡

February 2023

Abstract

In many games of interest (e.g., trade, entry, leadership, warfare, and partnership environments), one player (the leader) covertly acquires information about the state of Nature before choosing whether to engage with another player (the follower). The friendliness of the follower's reaction depends on his beliefs about what motivated the leader's choice to engage. We provide necessary and sufficient conditions for the leader's value of acquiring more information to increase with the follower's expectations. We then derive the economic implications of this characterization, focusing on three closely related topics (cognitive traps, disclosure, and cognitive styles), drawing policy implications.

Keywords: Adverse selection, expectation conformity, generalized lemons problem, endogenous information, cognitive traps.
JEL numbers: C72; C78; D82; D83; D86.

*Research support from European Research Council advanced grant (European Community's Seventh Framework Programme (FP7/2007-2013) Grant Agreement no. 249429 and European Union's Horizon 2020 research and innovation programme, Grant Agreement no. 669217) is gratefully acknowledged. For comments and useful suggestions, we thank participants at various conferences and workshops where the paper was presented.

†Department of Economics, Northwestern University, and CEPR. Email: alepavan@northwestern.edu.

‡Toulouse School of Economics (TSE) and Institute for Advanced Study in Toulouse (IAST), University of Toulouse Capitole. Email: jean.tirole@tse-fr.eu.

1 Introduction

Many strategic situations of interest can be thought of as Stackelberg games in which one player, player L (the leader, she) chooses between an "adverse-selection sensitive" action and an "adverse-selection insensitive" one. The reaction of the other player, player F (the follower, "he") to the adverse-selection sensitive action depends on his beliefs about what motivated L's choice of action. For example, player L may represent a seller choosing between offering to trade with a buyer (the adverse-selection sensitive action) and opting out of the negotiations (more generally, player F may still act following the information insensitive action, as when player L discloses what she knows). The key assumption is that information that makes player L want to engage with player F by choosing the adverse-selection sensitive action makes the latter player react in an unfriendly manner. Notable examples of such situations include, in addition to Akerlof's (1970) lemons model, many entry and partnership games that are central to the Industrial Organization, Finance, and Organization Economics literatures (we discuss a few of these situations in the next section).

We enrich this classic model by allowing player L to covertly acquire information about the state of Nature before making her engagement decision. We are interested in understanding how player L's choice of cognition depends on player F's expectations. We identify sufficient and/or necessary conditions for *expectation conformity* (EC) to hold in these games, namely for player L to find it more valuable to acquire more information when player F expects her to do so. Besides being of independent interest, EC has important economic implications. In particular, interactions or markets may switch behavior abruptly; for instance, asset markets can tip from a pattern in which the assets receive little scrutiny to one in which they are heavily scrutinized by participants. EC also shapes the benefits of disclosure of hard information and plays a key role in the possibility that the players end up in a cognitive trap where they suffer from the information they are expected to acquire.

Section 2 defines a broad class of generalized lemons environments, in which one of the players acquires information covertly and then decides whether or not to engage (choose the adverse-selection sensitive action) with the other player; it reinterprets a number of familiar games within this framework. Section 3 introduces the notion of EC; to put flesh on the characterization, we compare information structures through the mean-preserving-spread order, or the more refined rotations order (Subsection 7.2 provides similar results for more flexible information structures). The rotations ordering is for instance satisfied for any prior distribution in the case of non-directed search in which L invests in the probability of learning the state of Nature. For more general environments, it is a property of the family of distributions (we give examples with uniform, Pareto and exponential distributions). The analysis derives a sufficient (and, under further assumptions, necessary) condition for such games to satisfy EC. This condition says that more cognition aggravates adverse selection (in a well-defined sense, which makes F's reaction less friendly) and that an unfriendlier reaction by F raises L's incentive to acquire information; or that both conditions are simultaneously reversed. The condition for EC is easier to check than verifying directly that EC prevails. It shows that, in the cognition-augmented lemons game, EC holds when the gains from engagement are large, but not for low gains from engagement. To further the intuition about what drives EC in generalized lemons games, it applies the analysis to the cognition-augmented lemons game under directed and non-directed search.

The paper then derives the economic implications of this characterization in Section 4, focusing on three closely-related topics: cognitive traps, disclosure, and cognitive styles. In generalized lemons games, the information-acquiring player is worse off in a high-information-intensity equilibrium; this cognitive trap is due to the unfriendly reaction of the other player in response to the exacerbated adverse selection. We then modify the game by assuming that the information acquired is hard instead of soft and so the information-acquiring player can choose to disclose how much external resources she devoted to the issue. Such disclosure turns out to be mostly irrelevant. The intuition is related to the cognitive-trap phenomenon: Disclosure serves to demonstrate that one is knowledgeable, which in the end is not profitable. Along a similar vein, it is shown that it is optimal to pose as an informational puppy dog, i.e., to convince the other player that one is dumb or busy.

Section 5 discusses how the results change in economies in which the lemons assumption is replaced by its anti-lemon counterpart (that is, states in which the leader is most eager to engage are those in which the follower's response is most favorable to the leader).

Section 6 contains policy analysis. It identifies conditions under which subsidies/taxes to trade are welfare enhancing as well as conditions under which the endogeneity of information calls for larger policy interventions. For example, in the Akerlof model, subsidies to trade come with a double dividend under endogenous information. In addition to inducing player L to engage more often, they discourage player L from acquiring information. As a result, the optimal subsidy is typically higher when information is endogenous than when it is exogenous.

Section 7 discusses the relation to other investment games and the robustness of the key insights to alternative information-acquisition technologies. Section 8 concludes. Omitted proofs are in the Appendix at the end of the document.

Related Literature.

The paper is related to various strands of the literature. The first one is the literature on the lemons problem under alternative information structures. Kartik and Zhong (2019) consider a bilateral trading environment with interdependent values and characterize the combination of all consumer and producer surplus that can be sustained in equilibrium under any possible information structure.[1] Related are also Levin (2001), Kessler (2001), and Bar-Isaac et al. (2018).

[1]The analysis parallels the one in Bergemann, Brooks and Morris (2015) but in a setting with interdependent

These papers, as Kartik and Zhong (2019), study how payoffs, the volume of trade, and the efficiency of bargaining outcomes vary with the information structure in variants of the Akerlof model. In contrast, we study how the acquisition of information is shaped by other players' expectations, and how the latter in turn depend on the information acquisition technology and the effect of information on the severity of the adverse selection problem.

Dang (2008) and Thereze (2022) also endogenize the information structure in the Akerlof model. However, the focus of the analysis in these papers is different. Dang (2008) derives conditions under which no information is acquired in equilibrium as well as conditions under which the player acquiring information receives positive surplus despite not having bargaining power at the negotiation stage. Thereze (2022), instead, considers a competitive adverse selection market in which the buyers' information also affects the sellers' costs (as in health markets) and investigates how the elasticity of the demand and the market equilibrium are affected by a change in the cost of information.[2]

A fairly vast literature studies information acquisition in bargaining games with private values. See for example Ravid (2020), and Ravid, Roesler, and Szentes (2022) and the references therein. The first paper considers a repeated bargaining setting with a rationally-inattentive buyer. The second paper investigates the properties of the equilibrium when the cost of the buyer's information vanishes in a one-shot ultimatum-bargaining game. Our paper, instead, considers games with interdependent payoffs (as in the lemons problem) and investigates how the information acquired in equilibrium is affected by the effect of information on the severity of the adverse selection problem and how this in turn leads to cognitive traps.

Pavan and Tirole (2022a) shares with the present paper the interest in how the possibility to disclose verifiable/hard information affects equilibrium outcomes in settings with interdependent payoffs. That paper focuses on the welfare effects of mandatory disclosure laws. The present paper, instead, focuses on the effects of cognition on the severity of the adverse selection problem and on policy interventions aimed at alleviating such a severity. Expectation conformity is also studied in Pavan and Tirole (2022b). The analysis in that paper, however, is not specific to settings with adverse selection and none of the results in the present paper have counterparts in that paper.

Finally, the discussion of how benevolent governments can alleviate the adverse selection problem in Section 6 is related to Philippon and Skreta (2012) and Tirole (2012). The sellers' information in those papers is exogenous. Instead, the present paper studies how governments' programs influence the acquisition of information. See also Colombo, Femminis and Pavan (2022)

payoffs.

[2]In Thereze (2022), the buyers acquire information after seeing the prices asked by the sellers, whereas, in our model, as in Ravid, Roesler, and Szentes (2022), information acquisition takes place prior to observing the prices. See also Cremer and Khalil (1992), and Cremer, Khalil, and Rochet (1998) for the earlier work on information acquisition in contractual settings.

for how governments can incentivize information acquisition in economies with investment complementarities, and Pavan, Sundaresan, and Vives (2022) for how governments can influence information acquisition in financial markets.

2 Framework

2.1 Description

Consider the following game between two players, a "leader" (she) and a "follower" (he).

(a) Actions

Player L (the "leader") first selects an information structure. She then chooses between two actions, $a = 0$ and $a = 1$, after updating her beliefs about the state of Nature upon observing the realization of the selected information structure (equivalently, of the selected experiment). Player F (the "follower") then chooses his reaction to the leader's action, after observing a but not L's choice of an information structure or its realization. As we explain below, player F's reaction to $a = 0$ plays no role in the analysis and hence we do not formally describe it. His reaction to $a = 1$, instead, will be denoted by $r \in \mathbb{R}$.

To take a classical example, consider Akerlof's (1970) lemons model. Player L is the seller of a used car. Action $a = 1$ represents the seller's decision to put the car on sale, whereas $a = 0$ the decision to keep the car for own consumption. The follower's reaction to $a = 1$ 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 normalize player F's action so that a higher r stands for a friendlier response: player L's utility is increasing in r.

(b) Information/cognition

Prior to choosing a, player L acquires information about the state of Nature. The state of Nature, say the car's quality in the lemons model, is denoted by $\omega \in (-\infty, +\infty)$, and is commonly believed to be drawn from a distribution G with prior mean ω_0. We will assume that the two players' preferences are affine in ω, so they care only about the posterior mean m of the state. An experiment, indexed by $\rho \in \mathbb{R}_+$, will be taken to be the choice of a distribution $G(m;\rho)$ of the induced posterior mean, satisfying the martingale property $\int_{-\infty}^{+\infty} m dG(m;\rho) = \omega_0$ for all ρ.[3] We will denote by $G_\rho(m;\rho)$ the partial derivative of G with respect to ρ. The cost of acquiring information, by abuse of notation, will be denoted by $C(\rho)$; when ρ is interpreted as the intensity of information-collection, as is often the case, one can naturally assume that $C(0) = 0$, $C'(\rho) > 0$ and $C''(\rho) \geq 0$ (see below for a family of information structures indexed by a uni-dimensional parameter, ρ).

[3] Note that the support of $G(m;\rho)$ can be a strict subset of $\mathbb{R}$.

(c) Preferences

Follower. Action $a = 1$ is "adverse-selection sensitive," in the sense that player F's reaction to $a = 1$ depends on his beliefs about what information privately held by player L motivated L to engage. After observing $a = 1$, player F, anticipating cognition level $\rho^\dagger$, updates his beliefs about L's posterior mean m using the distribution $G(m; \rho^\dagger)$ and L's engagement strategy $a^*(m; \rho^\dagger)$. Given cognition $\rho^\dagger$, the function $a^*(m; \rho^\dagger)$ specifies, for each m, the probability that player L engages given her posterior mean m.[4] Hereafter, we denote by $\hat{G}(m;\, \rho^\dagger)$ the cdf of the distribution over m corresponding to F's posterior beliefs, when the latter expects L to select cognition $\rho^\dagger$ and engage according to $a^*(m; \rho^\dagger)$, after observing $a = 1$. Given $\hat{G}(\cdot;\, \rho^\dagger)$, F then maximizes his expected payoff $\mathbb{E}_{\hat{G}(\cdot;\, \rho^\dagger)}[u_F(1,\, r,\, m)]$ by means of an action $r \in \mathbb{R}$, where $u_F(1,\, r,\, m)$ is F's payoff when L engages (i.e., selects $a = 1$), F's reaction is r, and L's posterior mean is m.[5]

Let $r(\rho^\dagger)$ denote F's reaction to $a = 1$ when expecting L to choose cognition $\rho^\dagger$. The reason for expressing F's reaction only as a function of $\rho^\dagger$ (thus dropping its dependence on L's engagement strategy is that, as the analysis below reveals, the relevant reaction by F when expecting cognition $\rho^\dagger$ is the one that obtains in equilibrium in a fictitious game in which L's cognition is exogenously fixed at $\rho^\dagger$. Let $a^*(m; \rho^\dagger)$ denote the equilibrium engagement strategy in such a fictitious game. Hereafter, we ease the notation by dropping the strategy $a^*(m; \rho^\dagger)$ and denoting by $r(\rho^\dagger)$ the follower's reaction to the anticipation that L's cognition is $\rho^\dagger$. We also maintain that $r(\rho^\dagger)$ is unique, for any $\rho^\dagger$.

By contrast, action $a = 0$ is "adverse-selection insensitive". In some applications, such as Akerlof's lemons model, action $a = 0$ involves no decision for the follower. More generally, we assume that the follower's reaction to $a = 0$ is independent of his beliefs about $\rho^\dagger$. This is the case for instance in the disclosure example discussed in the next subsection, in which action $a = 0$ corresponds to the decision to disclose hard information proving that the state (or L's posterior belief) is m, making F's conjecture about L's cognition irrelevant. This example also shows that "adverse-selection sensitivity" is different from "information sensitivity".

Leader. Player L's payoff differential between $a = 1$ and $a = 0$ depends on r and on player L's posterior mean m. Let $u_L(0,\, m)$ denote L's payoff when choosing $a = 0$. As just discussed, this payoff may depend on F's reaction. However, because the latter is invariant in F's expectations over L's cognition, we can omit it to ease the notation and interpret $u(0, m)$ as L's payoff in state m given F's reaction to $a = 0$. Similarly let $u_L(1,\, r,\, m)$ denote L's payoff when choosing $a = 1$

[4] We are interested in situations in which L engages for a positive measure set of m, in which case L's beliefs upon observing $a = 1$ are pinned down by Bayes' rule. Also, in some of the applications of interest, it may be more natural to think of L as engaging after observing F's reaction (e.g., a seller decides whether to sell an asset after observing the buyer's offer). Our results apply verbatim to these setting as well. Whether player L observes player F's reaction at the time she engages, or anticipates it, is inconsequential for our analysis.

[5] As explained above, the assumption that L's and F's payoffs are affine in ω implies that $u_F(1,\, r,\, m)$ is also F's ex-post payoff when the state is $\omega = m$.

and then denote by

$$\delta_L(r,\, m) \equiv u_L(1,\, r,\, m) - u_L(0,\, m)$$

L's payoff differential between $a = 1$ and $a = 0$, when F's reaction to $a = 1$ is r and L's posterior mean is m.

Assumption 1 (***leader's preferences***)**.** *Player L's payoff differential, $\delta_L(r,\, m)$, is Lipschitz continuous and twice continuously differentiable in each argument, strictly increasing in r, strictly decreasing in m, and such that the marginal impact of a friendlier reaction, $\partial\delta_L/\partial r$, is weakly increasing in m:*

$$\frac{\partial^2 \delta_L(r, m)}{\partial m \partial r} \geq 0. \tag{1}$$

That $\partial\delta_L/\partial r > 0$ reflects the property that a higher r represents a friendlier reaction, favoring $a = 1$. That $\partial\delta_L/\partial m < 0$ implies that a lower m favors $a = 1$. The strict monotonicity of δ_L in m in turn implies that, no matter the actual choice of ρ, L optimally plays $a = 1$ if and only if m falls below some cutoff $m^*(r)$ that depends on F's reaction r, with the cutoff $m^*(r)$ solving $\delta_L(r,\, m^*(r)) = 0$ and hence strictly increasing in r. Clearly, in any equilibrium in which L's actual cognition is ρ, the cognition $\rho^\dagger$ expected by F coincides with L's actual cognition ρ, and F's reaction is $r(\rho)$, where $r(\rho)$ is F's equilibrium reaction in a fictitious game in which cognition is exogenously fixed at ρ.

Condition (1) in Assumption 1 will be used to determine whether cognition becomes more or less attractive to player L when player F behaves in a friendlier way (see the proof of Part (iii) of Proposition 1 below).

Let player F anticipate cognition $\rho^\dagger$ by player L. Out-of-equilibrium, $\rho^\dagger$ can differ from L's actual cognition ρ, because cognition is covert. However, suppose for a moment that cognition is exogenous and restricted to $\rho^\dagger$. Because player F's payoff is quasilinear in ω, his reaction $r(\rho^\dagger)$ depends on the distribution $\hat{G}(m;\, \rho^\dagger)$ describing his beliefs over L's posterior mean m only through the mean $\mathbb{E}_{\hat{G}(\cdot;\, \rho^\dagger)}[m]$ of $\hat{G}(m;\, \rho^\dagger)$. Furthermore, as explained above, when L's cognition is exogenously fixed at $\rho^\dagger$, in equilibrium, player L's engagement strategy $a^*(m; \rho^\dagger)$ takes the form of a cutoff rule, i.e., L engages if and only if $m \leq m^*$, in which case $\mathbb{E}_{\hat{G}(\cdot;\, \rho^\dagger)}[m] = M^-(m^*;\, \rho^\dagger)$, where

$$M^-(m^*;\, \rho^\dagger) \equiv \mathbb{E}_{G(\cdot;\, \rho^\dagger)}[m | m \leq m^*] = m^* - \frac{\int_{-\infty}^{m^*} G(m; \rho^\dagger) dm}{G(m^*; \rho^\dagger)}$$

denotes the truncated mean of the distribution of m under the distribution $G(\cdot;\, \rho^\dagger)$ corresponding to cognition $\rho^\dagger$. An increase in M^- can then be viewed as a reduction of the adverse selection problem.

Assumption 2 (***lemons***)**.** *The friendliness of player F's reaction to an increase in player L's cognition is determined by the effect of L's cognition on adverse selection:*[6]

[6]Note that $\frac{\partial}{\partial \rho^\dagger} M^-(m^*(r(\rho^\dagger));\, \rho^\dagger)$ is the partial derivative of $M^-(m^*;\, \rho^\dagger)$ with respect to $\rho^\dagger$, holding m^*

$$\frac{dr(\rho^\dagger)}{d\rho^\dagger} \overset{sgn}{=} \frac{\partial}{\partial\rho^\dagger} M^-\big(m^*(r(\rho^\dagger));\, \rho^\dagger\big). \tag{2}$$

Remark [**Relative adverse selection sensitivity**]. We assume that action $a = 0$ is "adverse-selection insensitive". However, we expect most of the results to extend to settings in which F's reaction to $a = 0$ also depends on F's beliefs about ρ and m, but with a lower sensitivity to these variables than F's reaction to $a = 1$. The following example illustrates the type of applications that this more general setting can capture. Player L is an employee who can choose between a high- and a low-powered incentive scheme (for brevity, HPIS and LPIS). Action $a = 0$ correspond to the decision to choose HPIS, whereas $a = 1$ the decision to choose LPIS. Let y_a denote the employee's "skin in the game," e.g., the amount of shares held, with $1 \geq y_0 > y_1 \geq 0$. Player F is an employer whose payoff is $\kappa + (1 - y_a)(e_a + m) - r_a$, where κ is a constant, e_a is the effort exerted by the employee (at increasing and convex private cost $\psi(e)$) after choosing action $a \in \{0, 1\}$, and r_a is a fixed wage paid by F to L on top of the money paid through the incentive payment y_a. Hence, in this application, there are two reactions by player F, r_1 and r_0, and each may depend on $\rho^\dagger$. Let $U_L(a, r_a, m)$ and $U_F(a, r_a, m)$ denote the two players' payoffs when the leader takes action a, the follower reacts with action r_a, and L's posterior mean is m. Then $U_L(a, r_a, m) \equiv \max_e\{r_a + y_a(e + m) - \psi(e)\}$ and $U_F(a, r_a, m) \equiv \kappa + (1 - y_a)(e_a + m) - r_a$. Let $r \equiv r_1 - r_0$. Then, up to an additive constant, $\delta_L(r, m) = r - (y_0 - y_1)m$, and, up to additive and multiplicative constants, $r = M^-(m^*(r); \rho^\dagger) - zM^+(m^*(r); \rho^\dagger)$, where $z \equiv (1 - y_0)/(1 - y_1) < 1$. In our model, $z = 0$. We expect our results to extend to this type of settings provided that δ_L continues to depend only on r and m and satisfies Assumption 1 above, z is small (action $a = 0$ is relatively less "adverse-selection sensitive" than $a = 1$) and the property in Assumption 2 applies to $M^- - zM^+$ instead of M^- (which is the case, for example, when m is drawn from a Uniform or a Pareto distribution).

2.2 Examples

The Stackelberg game described above (and its key assumptions, 1 and 2) may look somewhat abstract. Let us illustrate how a number of games of interest fit into this structure. The first four examples are variants of Akerlof's lemons model, augmented by the seller's endogenous covert information acquisition. In the classic variant, the seller holds exogenous soft information about the quality of the good and must decide whether to sell the good or keep it for own usage or future sale. In the second example, a government engages in asset repurchases so as to jumpstart a frozen market. In the third example, the good is divisible (a share in a project); the owner benefits from synergies associated with taking an associate in the project, but is hesitant about

fixed at $m^* = m^*(r(\rho^\dagger))$, where $m^*(r(\rho^\dagger))$ is the equilibrium threshold describing L's engagement strategy in the fictitious game in which L's cognition is exogenously fixed at $\rho^\dagger$.

sharing the proceeds if she knows the project is highly profitable. In the fourth variant, the seller may have hard information about the quality of the good and chooses whether to disclose it to the buyer. The fifth example describes herding with interdependent payoffs; for example, by entering a market, a firm may encourage a rival to follow suit. The last example is a marriage game, in which covenants may smooth the hardship of a subsequent divorce, but may also signal bad prospects about the marriage.

(a) *Akerlof.* In Akerlof's model, player L is a seller. She can sell her good in the market ($a = 1$) or keep it for herself ($a = 0$). Player F is a representative of a set of competitive buyers who choose a price r equal to the expected value of the good conditional on the good being put in the market (the general model above also admits as a special case a different version of the Akerlof model in which the buyer, instead of being competitive, has full bargaining power; this version is the common-value counterpart of the game considered in Ravid, Roesler, and Szentes (2022)). Suppose that the players' utilities from the good are m for the seller and $m + \Delta$ for the representative buyer, with $\Delta \geq 0$ representing the gains from trade. Then, $r(\rho^\dagger)$ is the price offered by the competitive buyer when the seller's cognition is exogenously fixed at $\rho^\dagger$ and is given by the solution to the following equation

$$r = \mathbb{E}_{G(\cdot;\rho^\dagger)}\left[m + \Delta | m \leq r\right] = M^-(r;\rho^\dagger) + \Delta \tag{3}$$

reflecting the fact that the cutoff $m^*(r)$ for L's equilibrium engagement strategy $a^*(m;\rho^\dagger)$ is equal to r. We assume that the solution to (3) is unique, which is indeed the case if the inverse hazard rate of the distribution of m for cognition $\rho^\dagger$, $G(m;\rho^\dagger)/g(m;\rho^\dagger)$, is increasing in m.[7] Assumption 2 is then satisfied. So is Assumption 1, as $\delta_L(r,m) = r - m$.
Turning to the case in which the seller's cognition is endogenous, we then have that L's optimal choice of ρ when L anticipates a reaction r by F is given by

$$\max_\rho \{G(r;\rho)r + \int_r^\infty m dG(m;\rho) - C(\rho)\}.$$

When C and G are differentiable in ρ and the above objective function for player L satisfies the appropriate concavity conditions (we will maintain these assumptions throughout when referring to this example), the optimal level of ρ is then given by the following first-order condition[8]

$$-\int_r^{+\infty} G_\rho(m;\rho)dm = C'(\rho). \tag{4}$$

(b) *Jumpstarting frozen markets.* The suboptimal volume of trade in Akerlof's model motivates

[7] Then $\partial M^-(r;\rho^\dagger)/\partial m^* \in (0,1)$. See An (1998).

[8] Note that the FOC for ρ can also be written as $\int_{-\infty}^r G_\rho(m;\rho)dm = C'(\rho)$. This is because $\int_{-\infty}^{+\infty} m dG(m;\rho)$ is invariant in ρ, implying that $\int_{-\infty}^{+\infty} G_\rho(m;\rho)dm = 0$.

policy interventions (see also the analysis in Section 6). Consider a government that, in the context of Example (a) above, maximizes total welfare. The government can neither coerce the agents to trade nor prevent the existence of a free private market. However, it can influence the market outcome by purchasing some of the assets. The shadow cost of public funds used for such purchases is $1+\lambda$, with $\lambda > 0$. The government, just like private buyers, does not know ω and values the assets at $\omega+\Delta$. Philippon-Skreta (2012) and Tirole (2012) show that, when the sellers' information is exogenous, the optimal policy has a simple form: The government purchases the lowest-quality assets, the market the intermediate-quality ones, and the best-quality assets are kept by the sellers. Furthermore, the optimum can be implemented by the government setting a price r so as to maximize

$$\max_r \left\{ G(r;\rho^\dagger) \left[\Delta - \lambda\left(r - M^-(r;\rho^\dagger) - \Delta\right)\right]\right\}, \tag{5}$$

where $\rho^\dagger$ is the sellers' cognition.[9] The first-term in the square bracket is the gain from trade, whereas the second term in the square brackets represents the deadweight loss on the deficit, reflecting the fact that, when the competitive private market breaks even, it is *as if* the government purchases all tendered assets itself.[10] The first-order condition for the optimal choice of r is[11]

$$\frac{g(r;\rho^\dagger)}{G(r;\rho^\dagger)} = \frac{\lambda}{(1+\lambda)\Delta}. \tag{6}$$

Under the same differentiability and convexity assumptions as in example (a), we then have that the sellers' acquisition of information continues to be given by (4). The only difference is that r is now chosen by the government instead of being determined by the market.

(c) *Partnerships.* Player L has a project. She can associate player F to it or do it alone. Bringing player F on board creates synergies (lowers the cost of implementation), but forces L to share the gains, which she does not want to do if the project is a good one. Player L's payoff is $\omega - d_L$ if she does it alone and $r\omega - c_L$ if it is a joint project, where r is the value share left to

[9]To see this, let m^g denote the critical threshold below which a seller of an asset of quality m sells to the government. The social cost of the government's program is then equal to $\lambda\left\{r - \mathbb{E}_{G(\cdot;\rho^\dagger)}[m|m \le m^g] - \Delta\right\} G(m^g;\rho^\dagger)$, which accounts for the fact that the government can resell the assets at price $\mathbb{E}_{G(\cdot;\rho^\dagger)}[m|m \le m^g] + \Delta$, and the proceeding from the sale can be used to reduce the distortions associated with future needs to collect money from taxpayers. The non-arbitrage condition between the government's program and the private market, along with the fact that buyers are competitive, then implies that $r = \mathbb{E}_{G(\cdot;\rho^\dagger)}[m|m \in [m^g, r]] + \Delta$. Combining the two conditions above and using the law of iterated expectations, we have that

$$M^-(r;\rho^\dagger) = \frac{G(m^g;\rho^\dagger)}{G(r;\rho^\dagger)} \mathbb{E}_{G(\cdot;\rho^\dagger)}[m|m \le m^g] + \frac{G(r;\rho^\dagger) - G(m^g;\rho^\dagger)}{G(r;\rho^\dagger)} \mathbb{E}_{G(\cdot;\rho^\dagger)}[m|m \in [m^g, r]]$$

which gives the formula in (5).

[10]It is important, though, that the government does not buy all these assets and lets the market rebound. Otherwise, the market would nonetheless rebound, and the sellers' anticipation of this rebound would force the planner to buy assets at an even higher price: See the papers mentioned above for details.

[11]The condition uses the fact that $\partial M^-(r;\rho^\dagger)/\partial m^* = g(r;\rho^\dagger)[r - M^-(r;\rho^\dagger)]/G(r;\rho^\dagger)$.

L by (competitive) player F and c_L is player L's reduced cost of project implementation. Let c_F denote player F's cost, with $c_L + c_F < d_L$. Given her posterior mean m, player L chooses $a = 1$ if and only if $\delta_L(r, m) = d_L - c_L - (1-r)m \geq 0$. Provided that $m > 0$, Assumption 1 is then satisfied. Finally, $r = r(\rho^\dagger)$ solves

$$(1-r)M^{-}\left(\frac{d_L - c_L}{1-r};\, \rho^\dagger\right) = c_F. \tag{7}$$

Provided that (7) has one and only one solution, then Assumption 2 is also satisfied.[12]

(d) *Disclosure of Hard Information.* In another important variant of Akerlof's game, the seller wants to sell for sure (she has no value for the good, say), and either has no information about ω (with probability $1-\rho$) or knows ω (with probability ρ), as in Dye's (1985) model. So $G(m;\rho) = \rho G(m)$ for $m < \omega_0$ and $G(m;\rho) = \rho G(m) + 1 - \rho$ for $m \geq \omega_0$, where ω_0 is the prior mean of G. In contrast with Akerlof's soft-information lemons game, the seller's decision is not whether to put the good in the market (a foregone conclusion), but whether to reveal the state of Nature when knowing it. A well-established literature, surveyed by Milgrom (2008), has studied such an incentive to disclose. A natural extension of the disclosure model consists in thinking of ρ (the precision of information) as endogenous.

In order to apply the general results, we must define the actions and the corresponding δ_L function. Let $a = 1$ stand for the decision of non disclosing and $a = 0$ for the decision of disclosing the state of Nature. The rationale for this choice is that player F's beliefs about player L's cognition matter only if there is no disclosure. As in Examples (a) and (b), let $\omega + \Delta$ denote the buyer's utility. Then let r denote the price offered by the buyer in the absence of disclosure. The seller thus obtains $m + \Delta$ if she discloses, and r if she does not disclose, so that $\delta_L(r, m) = r - (m + \Delta)$, so that Assumption 1 is satisfied.

To compute $r(\rho^\dagger)$, note that, when cognition is exogenously fixed at $\rho^\dagger$, the seller discloses if and only if she is informed and $m > m^*(r) = r - \Delta$. Hence $r(\rho^\dagger)$ solves

$$r = \frac{(1-\rho^\dagger)\omega_0 + \rho^\dagger \displaystyle\int_{-\infty}^{r-\Delta} m dG(m)}{1 - \rho^\dagger + \rho^\dagger G(r-\Delta)} + \Delta. \tag{8}$$

Our results below apply to this environment as well. However, in this setting, the expected value of m conditional on player L engaging (i.e., not disclosing) does not coincide with $M^{-}\left(m^*(r(\rho^\dagger));\, \rho^\dagger\right)$. This is because player L, when not receiving any information, has no choice but to engage, ir-

[12]Note that for (7) to admit one and only one solution it must be that $M^{-}\left(d_L - c_L;\, \rho^\dagger\right) \geq c_F$. When this condition holds, (7) admits at least one solution. Such a solution is unique if, and only if, in addition to the condition above, $c_F - \frac{\partial}{\partial m^*} M^{-}\left(\frac{d_L - c_L}{1-r};\, \rho^\dagger\right)(d_L - c_L) > 0$ for any r that solves (7). The last property always holds when $G(m;\rho^\dagger)/g(m;\rho^\dagger)$ is increasing in m and $c_F > d_L - c_L$ for, in this case, $\frac{\partial}{\partial m^*} M^{-}\left(\frac{d_L - c_L}{1-r};\, \rho^\dagger\right) \in (0, 1)$.

respective of whether her posterior expected value of ω (which is equal to ω_0) is below or above m^*.[13] The analog of Assumption 2 in this setting is that the sign of $dr(\rho^\dagger)/d\rho^\dagger$ coincides with the sign of $\partial \hat{M}\left(r(\rho^\dagger);\, \rho^\dagger\right)/\partial\rho^\dagger$ where, for any $(r,\rho^\dagger)$

$$\hat{M}\left(r;\, \rho^\dagger\right) \equiv \frac{(1-\rho^\dagger)\omega_0 + \rho^\dagger \int_{-\infty}^{r-\Delta} m dG(m)}{1-\rho^\dagger + \rho^\dagger G(r-\Delta)}$$

denotes the expected value of m conditional on L engaging optimally against a reaction of r by F, under cognition $\rho^\dagger$. It is easy to see that this is the case whenever the solution to (8) is unique, which is always the case when $r - \hat{M}\left(r;\, \rho^\dagger\right)$ is increasing in r.

(e) *(Interdependent herding) entry games.* Firm L decides whether to enter a market. Firm F then decides whether to follow suit. Firm F uses the information revealed by firm L's decision, but, in contrast with most herding models, payoffs are interdependent and so externalities are not purely informational. Suppose for instance that L and F are rivals, with per-customer profit π^m under monopoly and $\pi^d < \pi^m$ under duopoly.[14] The state of Nature ω here represents information correlated with the two firms' entry costs. Specifically, assume that firm L's entry cost is ω whereas firm F's entry cost is $\omega + \varepsilon$, where ε is drawn from $\mathbb{R}$, according to the distribution $H(\varepsilon)$ with density $h(\varepsilon)$, independently from ω. Importantly, the realization of ε is unknown to firm L when it decides whether to enter. Let r denote the probability of non-entry by firm F and let m denote firm L's posterior expected value of ω. We then have that

$$\delta_L(r,m) \equiv \left[r\pi^m + (1-r)\pi^d\right] - m,$$

implying that $m^*(r) = r\left(\pi^m - \pi^d\right) + \pi^d$. Assumption 1 is thus satisfied. In this application, $r(\rho^\dagger)$ is then the solution to $r = 1 - H(\pi^d - M^-\left(m^*(r);\, \rho^\dagger\right))$. Assumption 2 is satisfied whenever the solution to this equation is unique, which is the case if the density h of H satisfies $h(\pi^d - M^-\left(m^*(r);\, \rho^\dagger\right)) < 1$, the distribution $G(m;\rho^\dagger)$ of m is such that $G(m;\rho^\dagger)/g(m;\rho^\dagger)$ is increasing in m, which implies that $\partial M^-\left(m^*(r);\, \rho^\dagger\right)/\partial m^* < 1$, and $\pi^m - \pi^d \in (0,1)$.

(f) *Marriage.* Consider the following variant of Spier (1992)'s model, augmented with cognition. Players L and F decide whether to get married. Getting married has value v_L and v_F for L and F, respectively, provided that all goes well, which has probability ω distributed on $[0,1]$. With probability $1-\omega$, instead, things go wrong in which case the players divorce, obtaining utility $v_i - \mathscr{L}_i$, $i = L, F$. The divorce can, however, be made less painful (raising the utility to $v_i - \ell_i$, with $0 < \ell_i < \mathscr{L}_i$, $i = L, F$) through a covenant spelling out the outcome in case of divorce. Adding the covenant costs a fixed amount $c_i < \mathscr{L}_i - \ell_i$ to player $i = L, F$, implying

[13] Furthermore, it is easy to see that any solution to (8) is such that $\omega_0 > r - \Delta = m^*(r)$.

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

that it is efficient to add the covenant if the parties want to marry but expect to divorce with a sufficiently high probability. The value of v_L is large enough that player L wants to marry regardless of whether the covenant is introduced ($v_L \geq \mathscr{L}_L$). In contrast, player F's value v_F is distributed on $[(1-\omega_0)\mathscr{L}_F, +\infty)$ according to the c.d.f. H and is F's private information. Player L may acquire information about ω and then chooses between a contract with ($a = 1$) and without ($a = 0$) covenant.[15] Player F then decides whether to accept to marry. Because $v_F - (1-\omega_0)\mathscr{L}_F \geq 0$, in the absence of any information and any covenant, player F always accepts to marry, no matter the realization of v_F. Let r denote the probability that player F accepts to marry when the proposed contract includes the covenant (i.e., when player L engages). This game also satisfies Assumptions 1 and 2. To see this, first note that[16]

$$\delta_L(r, m) = r[v_L - (1-m)\ell_L - c_L] + (1-r) \cdot 0 - [v_L - (1-m)\mathscr{L}_L].$$

Hence, $\delta_L(r, m)$ satisfies Assumption 1. Next, note that, in this example, $m^*(r)$ is given by

$$m^*(r) = \max\left\{\frac{r(v_L - \ell_L - c_L) - (v_L - \mathscr{L}_L)}{\mathscr{L}_L - r\ell_L}; 0\right\}.$$

Hence $r(\rho^\dagger)$ is given by the solution to $r = 1 - H(\ell_F + c_F - M^-(m^*(r); \rho^\dagger)\ell_F)$. Provided that the above equation admits a unique solution (which is the case when $r + H(\ell_F + c_F - M^-(m^*(r); \rho^\dagger)\ell_F)$ is increasing in r) Assumption 2 holds in this example too.

3 Expectation conformity

We now investigate how L's choice of cognition (i.e., of an information structure) is influenced by F's expectations and how the latter in turn depend on whether adverse selection aggravates with L's cognition. Adverse selection is here captured by the truncated mean $M^-(m^*;\, \rho^\dagger)$:

Definition 1 (impact of cognition on adverse selection). Starting from cognition $\rho^\dagger$, an increase in cognition by player L

- aggravates adverse selection if $\frac{\partial}{\partial \rho^\dagger} M^-(m^*(r(\rho^\dagger));\, \rho^\dagger) < 0$,
- alleviates adverse selection if $\frac{\partial}{\partial \rho^\dagger} M^-(m^*(r(\rho^\dagger));\, \rho^\dagger) > 0$.

Simple computations show that, for any cognition $\rho^\dagger$ and truncation point m^*,

$$\frac{\partial}{\partial \rho^\dagger} M^-(m^*;\, \rho^\dagger) \overset{\text{sgn}}{=} A(m^*; \rho^\dagger) \tag{9}$$

[15]As anticipated above, in this application, both actions are adverse-selection sensitive, but $a = 0$ is less so than $a = 1$.

[16]Observe that the absence of a covenant is "good news" about m.

where

$$A(m^*;\rho^\dagger) \equiv \left[m^* - M^-(m^*;\,\rho^\dagger)\right]G_\rho(m^*;\rho^\dagger) - \int_{-\infty}^{m^*} G_\rho(m;\rho^\dagger)dm.$$

The first term of A captures the direct effect of a change in the probability that player L engages on player F's expectation of the state. Because $m^* \geq M^-(m^*;\,\rho^\dagger)$, an increase in cognition alleviates adverse selection when it increases the chances that player L engages (i.e., when $G_\rho(m^*;\rho^\dagger) > 0$), whereas it aggravates it when it reduces the probability of such an event (i.e., when $G_\rho(m^*;\rho^\dagger) < 0$). The second term, of A, $\int_{-\infty}^{m^*} G_\rho(m;\rho^\dagger)dm$, in turn is related to the effect of cognition on the dispersion of L's posterior mean m. When more cognition induces more dispersion in the sense of second-order stochastic dominance (which is the case when information structures can be ranked according to the mean-preserving-spread order), this second effect unambiguously contributes to an aggravation of the adverse selection problem. Hereafter, we will refer to

$$\mathrm{A}(\rho^\dagger) \equiv A(m^*(r(\rho^\dagger));\rho^\dagger)$$

as the "*adverse-selection effect*" of an increase of cognition at $\rho^\dagger$. Note that, under Assumption 2, when $A > 0$ (alternatively, $A < 0$) a higher level of anticipated cognition triggers a friendlier (alternatively, an unfriendlier) reaction by player F.

Now recall that L's ex-ante payoff when choosing cognition ρ and expecting a reaction r by F to her decision to engage is equal to

$$\Pi(\rho;r) \quad \equiv \quad \sup_{a(\cdot)} \left\{U_L(0) + \textstyle\int_{-\infty}^{+\infty} a(m)\,\delta_L(r,m)dG(m;\rho)\right\}$$

where $U_L(0) \equiv \int_{-\infty}^{+\infty} u_L(0,m)dG(m)$ is L's ex-ante expected payoff when she never engages, and $a(m)$ represents the probability that L engages when her posterior mean is m.[17]

Then let

$$B(\rho;\rho^\dagger) \equiv -\frac{\partial^2\Pi(\rho;r(\rho^\dagger))}{\partial\rho\partial r}$$

denote the effect of a reduction in the friendliness of F's reaction, starting from $r = r(\rho^\dagger)$, on L's marginal value of information, evaluated at cognition ρ. Hereafter, we will refer to $B(\rho;\rho^\dagger)$ as the "*benefit of friendlier reactions effect*".

Definition 2 (cognition incentive effect of unfriendly reactions). Given $(\rho,\rho^\dagger)$, a reduction in the friendliness of player F's reaction starting from $r = r(\rho^\dagger)$, raises (alternatively, lowers) player L's incentive to invest in cognition at ρ if $B(\rho;\rho^\dagger) > 0$ (alternatively, if $B(\rho;\rho^\dagger) < 0$).

Using the envelope theorem along with the fact that, for any ρ, the optimal engagement strategy for L when F anticipates cognition $\rho^\dagger$, is to engage if and only if $m \leq m^*(r(\rho^\dagger))$, and

[17] Note that, because $u_L(0,m)$ is affine in m, $\int_{-\infty}^{+\infty} u_L(0,m)dG(m;\rho) = \int_{-\infty}^{+\infty} u_L(0,m)dG(m)$ for any ρ.

integrating by parts, we have that

$$\begin{aligned} B(\rho;\rho^{\dagger}) &= -\int_{-\infty}^{m^*(r(\rho^{\dagger}))} \frac{\partial \delta_L}{\partial r}(r(\rho^{\dagger}),m)dG_{\rho}(m;\rho) \\ &= -\frac{\partial \delta_L(r(\rho^{\dagger}),m^*(r(\rho^{\dagger})))}{\partial r} G_{\rho}\Big(m^*(r(\rho^{\dagger}));\rho\Big) + \int_{-\infty}^{m^*(r(\rho^{\dagger}))} \frac{\partial^2 \delta_L(r(\rho^{\dagger}),m)}{\partial r \partial m} G_{\rho}(m;\rho)dm. \end{aligned}$$

Because δ_L is increasing in r, the sign of the first term of $B(\rho;\rho^{\dagger})$ is determined by whether an increase in cognition increases or reduces the chances that player L engages. Under Assumption 1, the benefit of a friendlier reaction by player F are increasing in the state. As a result, the second term of $B(\rho;\rho^{\dagger})$ is always positive when higher cognition by player L indexes a mean preserving spread of the induced posterior mean.

Next, let $V_L(\rho;\rho^{\dagger}) \equiv \Pi(\rho; r(\rho^{\dagger}))$ denote the maximal payoff that player L can obtain by engaging in cognition ρ when player F expects cognition $\rho^{\dagger}$.

Definition 3 (expectation conformity). Expectation conformity holds at $(\rho,\rho^{\dagger})$ if and only if

$$\frac{\partial^2 V_L(\rho;\rho^{\dagger})}{\partial \rho \partial \rho^{\dagger}} > 0.$$

Hence, expectation conformity is a local property that says that, starting from $(\rho,\rho^{\dagger})$, the marginal value to player L from expanding her cognition is higher when player F expects a higher cognition. When such a property holds for all $\rho,\rho^{\dagger} \in [\rho_1,\rho_2]$, the gross value to player L from changing cognition from ρ_1 to ρ_2 is higher when player F expects cognition ρ_2 than when he expects cognition ρ_1: $V_L(\rho_2;\rho_2) - V_L(\rho_1;\rho_2) > V_L(\rho_2;\rho_1) - V_L(\rho_1;\rho_1)$. Below we relate this property to the determinacy of equilibria and a few other phenomena of interest.

In a number of applications, information structures are ordered by their informativeness. Say that the distribution $G(\cdot;\rho)$ is obtained by observing the realization $z \in Z$ of some experiment $q : \Omega \to \Delta(Z)$, where Z is a measurable space of signal realizations. Then if higher ρ index distributions corresponding to Blackwell more informative experiments, the family of distributions $G(\cdot;\rho)$ must be consistent with the *mean-preserving-spread* (MPS) order.[18]

Assumption 3 (*MPS*). *Player L's set of feasible information structures is consistent with the MPS order if, for any ρ and $\rho' > \rho$, any $m^* \in \mathbb{R}$, $\int_{-\infty}^{m^*} G(m;\rho')dm \geq \int_{-\infty}^{m^*} G(m;\rho)dm$, with $\int_{-\infty}^{+\infty} G(m;\rho')dm = \int_{-\infty}^{+\infty} G(m;\rho)dm$.*

Consistently with what assumed above, when invoking Assumption 3, we will maintain that the set of information structures is rich in that $G(m;\rho)$ is differentiable in ρ, for any $m \in \mathbb{R}$. Assumption 3 then boils down to the requirement that, for any $m^* \in \mathbb{R}$ and ρ, $\int_{-\infty}^{m^*} G_{\rho}(m;\rho)dm \geq$

[18] An experiment q'' is Blackwell more informative than another experiment q' if observing the realization of q'' is equivalent to observing the realization of q' along with the realization of some other experiment $t : \Omega \to \Delta(Z)$.

0, with $\int_{-\infty}^{+\infty} G_\rho(m;\rho)dm = 0$. Our key results below assume that information structures are consistent with the MPS order. However, some of the results assume a special subset of such an order in which the spreads correspond to "rotations."

Definition 4 (rotations). Player L's set of possible information structures are "rotations" (or "simple mean-preserving spreads" or experiments consistent with the "single-crossing property") if, for any ρ, there exists a rotation point m_ρ such that $G_\rho(m;\rho) \geq 0$ for $-\infty < m \leq m_\rho$ and $G_\rho(m;\rho) \leq 0$ for $m_\rho \leq m < +\infty$ (with some inequalities strict).

A simple mean-preserving spread is a mean-preserving spread, but the converse is not true. For example, a combination of two rotations need not be a rotation, unless they have the same rotation point. As is well known, however, any mean-preserving spread can be obtained through a sequence of simple mean-preserving spreads.

Non-directed search example. Assume that information collection follows the standard non-directed search technology:

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

That is, L learns the true state with probability $\rho \in [0,1]$ and nothing with probability $1-\rho$. In this example, the rotation point is thus equal to the prior mean ω_0. Figure 1 below illustrates the idea for the special case in which G is uniform.

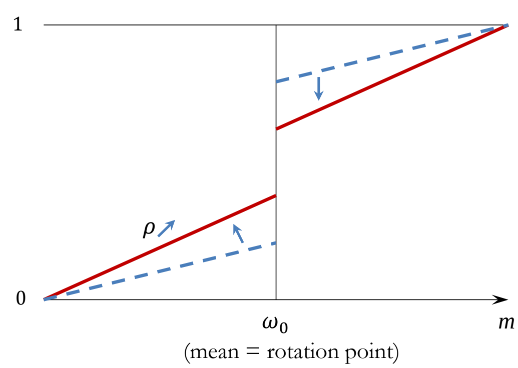

Figure 1: (cumulative distribution function $G(m;\rho)$ for non-directed search)

Other examples of rotations include a normally distributed state ω together with a signal that is normally distributed around the true state (ρ is then the precision of this signal), and the family of Pareto, Exponential, and Uniform distributions in Proposition 1 below. See Diamond and Stiglitz (1974) and Johnston and Myatt (2006) for a broader discussion of rotations and their properties.

Proposition 1 (***expectation conformity***). *Suppose that Assumptions 1, 2, and 3 hold.*
(i) Expectation conformity holds at $(\rho, \rho^\dagger)$ *if and only if the adverse selection effect and the benefit of a friendlier reaction effect are of opposite sign:* $\mathrm{A}(\rho^\dagger)B(\rho;\rho^\dagger) < 0$.
(ii) Cognition always aggravates adverse selection at $\rho^\dagger$ *(i.e.,* $\mathrm{A}(\rho^\dagger) < 0$*) when the distribution* $G(\cdot;\rho)$ *from which* m *is drawn is Uniform, Pareto, or Exponential. For other distributions, a sufficient condition for cognition to aggravate adverse selection at* $\rho^\dagger$ *is that* $G_\rho(m^*(r(\rho^\dagger));\rho^\dagger) < 0$.
(iii) Starting from $r(\rho^\dagger)$*, a reduction in the friendliness of player* F*'s reaction raises player* L*'s incentive for cognition at* ρ *(i.e.,* $B(\rho;\rho^\dagger) > 0$*) if* $G_\rho(m^*(r(\rho^\dagger));\rho) < 0$.
(iv) Therefore a sufficient condition for expectation conformity at $(\rho, \rho^\dagger)$ *is that*

$$\max\left\{G_\rho(m^*(r(\rho^\dagger));\rho^\dagger), G_\rho(m^*(r(\rho^\dagger));\rho)\right\} < 0. \tag{10}$$

(v) Suppose that, for any m^**,* $M^-(m^*;\rho)$ *is decreasing in* ρ *(as for the Uniform, Pareto, and Exponential distributions), implying that, for any* $\rho^\dagger$*,* $\mathrm{A}(\rho^\dagger) < 0$*. If furthermore* $\partial^2\delta_L(r,m)/\partial r\partial m = 0$ *(as is examples (a), (b) and (d) and (e) in the previous section), then* $G_\rho(m^*(r(\rho^\dagger));\rho) < 0$ *is a necessary and sufficient condition for expectation conformity at* $(\rho,\rho^\dagger)$*. When the distributions* $G(m;\rho)$ *are rotations, in the sense of Definition 4, this is the case if and only if* $m^*(r(\rho^\dagger))$ *is to the right of the rotation point* m_ρ.

Proof. (i) By the chain rule and the definitions of V_L and B, we have that

$$\frac{\partial^2 V_L(\rho;\rho^\dagger)}{\partial\rho\partial\rho^\dagger} = -B(\rho;\rho^\dagger)\frac{dr(\rho^\dagger)}{d\rho^\dagger}.$$

By virtue of Assumption 2, $dr(\rho^\dagger)/d\rho^\dagger$ is of the same sign as $\mathrm{A}(\rho^\dagger)$. We thus have that expectation conformity holds if, and only if, $\mathrm{A}(\rho^\dagger)$ and $B(\rho;\rho^\dagger)$ are of opposite sign.
(ii) Using Condition (9), we have that, for any $m^* \in \mathbb{R}$ and $\rho^\dagger$,

$$\frac{\partial M^-(m^*;\rho^\dagger)}{\partial\rho^\dagger} \stackrel{\text{sgn}}{=} \left[m^* - M^-(m^*;\rho^\dagger)\right]G_\rho(m^*;\rho^\dagger) - \int_{-\infty}^{m^*} G_\rho(m;\rho^\dagger)dm.$$

Because a higher ρ indexes a mean-preserving spread, $\displaystyle\int_{-\infty}^{m^*} G_\rho(m;\rho^\dagger)dm \geq 0$. Hence, starting from $\rho^\dagger$, cognition always aggravates adverse selection when $G_\rho(m^*(r(\rho^\dagger));\rho^\dagger) < 0$. Note, however, that this condition is sufficient but not necessary for $\mathrm{A}(\rho^\dagger) < 0$. For a number of distributions, $\partial M^-(m^*(r(\rho^\dagger));\rho^\dagger)/\partial\rho < 0$ regardless of the sign of $G_\rho(m^*(r(\rho^\dagger));\rho^\dagger)$. These distributions include the Uniform, Pareto, and Exponential distributions, as shown below.

- *Uniform distribution*: m is drawn uniformly from $[\underline{m}(\rho), \bar{m}(\rho)]$, with $\underline{m}(\rho)$ decreasing in ρ and satisfying $\underline{m}(\rho) \leq \omega_0$ for all ρ, and $\bar{m}(\rho) = 2\omega_0 - \underline{m}(\rho)$ for all ρ (mean preservation). Then for any $m \in [\underline{m}(\rho), \bar{m}(\rho)]$, $G(m;\rho) = \frac{m-\underline{m}(\rho)}{2(\omega_0-\underline{m}(\rho))}$. This family of distributions is thus consistent with the rotation order of Definition 4, with rotation point $m_\rho = \omega_0$ for all ρ. Furthermore, for any $m^* \in [\underline{m}(\rho), \bar{m}(\rho)]$,

$$M^-(m^*; \rho) = \frac{m^* + \underline{m}(\rho)}{2}$$

which is decreasing in ρ.

- *Pareto distribution*: m is drawn from $[\underline{m}(\rho), +\infty)$ according to the survival function $1 - G(m; \rho) = (\underline{m}(\rho)/m)^{\alpha(\rho)}$, with $\underline{m}(\rho)$ decreasing in ρ and $\alpha(\rho) = \omega_0/(\omega_0 - \underline{m}(\rho))$ for all ρ.[19] This family of distributions too is consistent with the rotation order of Definition 4. For each ρ, the rotation point is $m_\rho = \underline{m}(\rho) \exp\left(\frac{\omega_0 - \underline{m}(\rho)}{\underline{m}(\rho)}\right)$. Furthermore, for any $m^* > \underline{m}(\rho)$,

$$M^-(m^*; \rho) = \omega_0 \frac{1 - \left(\frac{\underline{m}(\rho)}{m^*}\right)^{\alpha(\rho)-1}}{1 - \left(\frac{\underline{m}(\rho)}{m^*}\right)^{\alpha(\rho)}}$$

which is decreasing in ρ.

- *Exponential distribution*: m is drawn from $[\underline{m}(\rho), +\infty)$ according to the survival function $1 - G(m; \rho) = e^{-\lambda(\rho)(m - \underline{m}(\rho))}$, with $\underline{m}(\rho)$ decreasing in ρ and $\lambda(\rho) = 1/(\omega_0 - \underline{m}(\rho))$ for all ρ.[20] One can verify that an increase in ρ induces a rotation of $G(m; \rho)$ in the sense of Definition 4, with rotation point $m_\rho = \omega_0$ for all ρ. Furthermore

$$M^-(m^*; \rho) = \omega_0 - \frac{(m^* - \underline{m}(\rho))\, e^{-\lambda(\rho)(m^* - \underline{m}(\rho))}}{1 - e^{-\lambda(\rho)(m^* - \underline{m}(\rho))}}$$

which is decreasing in ρ.

(iii) Next, recall that, starting from $r = r(\rho^\dagger)$, a reduction in the friendliness of F's reaction raises the incentive for cognition at ρ if and only if $B(\rho; \rho^\dagger) > 0$, where

$$\begin{aligned} B(\rho; \rho^\dagger) &= -\int_{-\infty}^{m^*(r(\rho^\dagger))} \frac{\partial \delta_L}{\partial r}(r(\rho^\dagger), m) dG_\rho(m; \rho) \\ &= -\frac{\partial \delta_L}{\partial r}\Big(r(\rho^\dagger), m^*(r(\rho^\dagger))\Big) G_\rho\Big(m^*\big(r(\rho^\dagger)\big); \rho\Big) \qquad (11) \\ &+ \int_{-\infty}^{m^*(r(\rho^\dagger))} \frac{\partial^2}{\partial r \partial m} \delta_L\big(r(\rho^\dagger), m\big) G_\rho(m; \rho) dm. \end{aligned}$$

[19] Note that the function $\alpha(\rho)$ is constructed so that, for any ρ, given $\underline{m}(\rho)$, $\mathbb{E}_{G(\cdot;\rho)}[m; \rho] = \frac{\alpha(\rho)\underline{m}(\rho)}{\alpha(\rho)-1} = \omega_0$ (mean preservation).

[20] Again, the function $\lambda(\rho)$ is constructed so that, for any ρ, given $\underline{m}(\rho)$, $\mathbb{E}_{G(\cdot;\rho)}[m; \rho] = \underline{m}(\rho) + \frac{1}{\lambda(\rho)} = \omega_0$ (mean preservation).

The latter term in (11) is positive as, by virtue of Assumption 3, ρ is an index of mean-preserving spread and $\partial^2\delta_L/\partial r\partial m$ is positive and constant in m by virtue of Assumption 1. Because δ_L is increasing in r by virtue of Assumption 1, the former term is positive provided that $G_\rho\Big(m^*\big(r(\rho^\dagger)\big);\rho\Big) < 0$. Hence, starting from $r = r(\rho^\dagger)$, a reduction in the friendliness of F's reaction raises the incentive for cognition at ρ if $G_\rho\Big(m^*\big(r(\rho^\dagger)\big);\rho\Big) < 0$.

(iv) The result follow from parts (i)-(iii) in the proposition.

(v) The result follows from parts (i)-(iii) in the proposition along with the fact that, in this case,

$$B(\rho;\rho^\dagger) = -\frac{\partial\delta_L}{\partial r}\Big(r(\rho^\dagger), m^*(r(\rho^\dagger))\Big)G_\rho\Big(m^*\big(r(\rho^\dagger)\big);\rho\Big)$$

which is positive if and only if $G_\rho\Big(m^*\big(r(\rho^\dagger)\big);\rho\Big) < 0$. ■

Hence, expectation conformity holds at $(\rho,\rho^\dagger)$ when, fixing player F's reaction at $r(\rho^\dagger)$, an increase in cognition by player L decreases the probability that L engages, both when such an increase is evaluated from player L's perspective (i.e., starting from cognition ρ) and when evaluated from player F's perspective (i.e., starting from cognition $\rho^\dagger$). This is because, from F's perspective, that player L engages less often (formally, $G_\rho(m^*(r(\rho^\dagger));\rho^\dagger) < 0$) implies an aggravation in the perceived adverse selection problem, which induces player F to respond in an unfriendlier manner (part (ii) in the proposition). That player F responds in an unfriendlier manner, together with the fact that $G_\rho(m^*(r(\rho^\dagger));\rho) < 0$, in turn implies a higher value for player L to increase her cognition starting from ρ (part (iii) in the proposition). Jointly, the above two properties imply that player L finds it more profitable to expand her cognition when player F expects more cognition.

As mentioned above, Condition (10) in the proposition is sufficient for expectation conformity but not necessary. Furthermore, the sufficiency of such a condition hinges on the information structures being consistent with the MPS order. The result is thus perhaps less obvious than what it may look like.

Finally, the last part of the proposition establishes that, in markets in which cognition always aggravates adverse selection and L's payoff is separable in m and r, condition $G_\rho(m^*(r(\rho^\dagger));\rho) < 0$ is not only sufficient but also necessary for expectation conformity to hold at $(\rho,\rho^\dagger)$.

As we discuss in the next section, expectation conformity has various economic implications. Before doing so, we illustrate how expectation conformity naturally emerges in the lemons problem with non-directed search introduced above.

3.1 Example: Cognition-augmented lemons model under non-directed search

We have shown that, under non-directed search, the rotation point is the prior mean. Proposition 1, when applied to the Akerlof's model of Subsection 2.2, thus says that expectation conformity holds at $(\rho, \rho^\dagger)$ whenever the engagement threshold $m^* = r(\rho^\dagger)$ is to the right of the prior mean, that is, when the price offered by the competitive buyer is sufficiently high. In other words, expectation conformity arises when the gains from trade (in the example parametrized by Δ) are large, and it never occurs when they are small.

To gather some intuition, recall that, in the lemons problem, the seller puts her car up for sale when her value for the car is small (i.e., when the posterior mean is below a threshold m^* that coincides with the price $r(\rho^\dagger)$ offered by the buyer). Naturally, when the gains from trade Δ are large, the price offered by the buyer is also large, in which case $r(\rho^\dagger)$ exceeds the rotation point, which coincides with the prior mean ω_0 of the car's value for the seller. Economically, what this implies is that the seller finds it optimal to enter the market both when she is uninformed and when she learns that the state ω is below $r(\rho^\dagger)$. Starting from such a situation, an increase in the cognition expected from the seller by the buyer reduces the quality of the car perceived by the buyer after seeing that the car is on sale. Faced with an exacerbated adverse selection problem, the buyer then reduces the price offered. But then it becomes even more important for the seller to learn the value of the car, that is, to increase her cognition starting from ρ. So expectation conformity naturally holds for $(\rho, \rho^\dagger)$ in this case.[21]

While the mechanism just described is fairly natural, it is important to appreciate that it need not always be in place. In fact, expectation conformity fails to obtain in this model when the gains from trade are positive but small. To see this, note that, when Δ is small, because of the adverse selection problem, the price offered by the buyer may well be lower than the ex-ante prior mean of the asset, meaning that $r(\rho^\dagger) < \omega_0$. Anticipating such a low price, the seller enters the market only if she receives information that reveals that $\omega \leq r(\rho^\dagger)$. The buyer then understands that the expected value of the car conditional on the seller putting it in on the market is the same independently of the seller's cognition: $M^-(r(\rho^\dagger); \rho^\dagger) = \int_{-\infty}^{r(\rho^\dagger)} \omega dG(\omega)/G(r(\rho^\dagger))$, which is invariant in $\rho^\dagger$. When this is this case, an increase in the cognition $\rho^\dagger$ expected from the seller by the buyer does not affect the price offered by the buyer, and hence does not increase L's incentives to search. We thus have the following result:

Corollary 1 (***lemons under non-direct search***)**.** *Consider the cognition-augmented lemons game under non-directed search described above. Expectation conformity holds in this game if and only if the gains from trade Δ are sufficiently large.*

[21]Consistently with the result in Proposition 1, when $r(\rho^\dagger) > \omega_0$, $G_\rho\Big(m^*\big(r(\rho^\dagger)\big); \rho^\dagger\Big) < 0$, meaning that an increase in cognition reduces the probability that the seller engages (see Figure 1).

3.2 Gains from engagement

The example in the previous subsection suggests that expectation conformity is more likely to be satisfied when the gains from engagement for player L are large. The next result shows that this is true more generally.

Proposition 2 (***gains from engagement***)**.** *Suppose that Assumptions 1 and 2 hold and that information structures take the form of rotations, as in Definition 4. Further assume that player L's payoff gain from playing $a = 1$ instead of $a = 0$ is $\delta_L(m, r) + \theta$, with $\theta \in \mathbb{R}$.*[22] *For any (r, θ), let $m^*(r; \theta)$ denote the optimal cut-off below which player L engages when the gains from engagement are parametrized by θ and F's reaction is r. For any $(\rho^\dagger, \theta)$, then let $r(\rho^\dagger; \theta)$ denote player F's response when player L's cognition is exogenously fixed at $\rho^\dagger$ and the gains from engagement are parametrized by θ. Take any $(\rho^\dagger, \rho, \theta')$ such that*

$$\max\left\{G_\rho(m^*(r(\rho^\dagger;\theta'),\theta');\rho^\dagger), G_\rho(m^*(r(\rho^\dagger;\theta'),\theta');\rho)\right\} < 0. \tag{12}$$

Then, for any $\theta'' > \theta'$,

$$\max\left\{G_\rho(m^*(r(\rho^\dagger;\theta''),\theta'');\rho^\dagger), G_\rho(m^*(r(\rho^\dagger;\theta''),\theta'');\rho)\right\} < 0. \tag{13}$$

Therefore, if expectation conformity holds at $(\rho,\rho^\dagger, \theta')$, it also holds at $(\rho,\rho^\dagger, \theta'')$. So for all $(\rho,\rho^\dagger)$, there exists $\theta^*(\rho, \rho^\dagger)$ such that, for all $\theta \geq \theta^*(\rho, \rho^\dagger)$, expectation conformity prevails at $(\rho,\rho^\dagger, \theta)$: expectation conformity is more likely, the larger the gains from engagement.

Proof. Note that, under Assumption 1, given r, the engagement threshold $m^*(r, \theta)$, which is implicitly defined by $\delta_L(r, m) + \theta = 0$, is strictly increasing in θ. Also observe that, under Assumption 2, given $\rho^\dagger$, $r(\rho^\dagger; \theta)$ is increasing in θ; this is because, fixing r and $\rho^\dagger$, a higher θ implies a higher engagement point $m^*(r, \theta)$, and hence a higher truncated mean $M^-(m^*(r, \theta); \rho^\dagger)$ which in turn implies a higher equilibrium response $r(\rho^\dagger; \theta)$ by virtue of Assumption 2. Because, for any θ, $m^*(r; \theta)$ is also increasing in r, we conclude that, for any $\rho^\dagger$, $m^*(r(\rho^\dagger; \theta''); \theta'') \geq m^*(r(\rho^\dagger; \theta'); \theta')$. The result in the proposition then follows from the fact that, when information structures are rotations, Condition (12) implies that $\max\{m_\rho, m_{\rho^\dagger}\} \leq m^*(r(\rho^\dagger; \theta'), \theta')$. The above properties in turn imply that $\max\{m_\rho, m_{\rho^\dagger}\} \leq m^*(r(\rho^\dagger; \theta''), \theta'')$ which means that Condition (13) also holds. ∎

Proposition 2 says that higher gains from engagement reinforce expectation conformity. On the other hand, holding player F's reaction fixed, larger gains from engagement reduce the

[22] For instance, an increase in θ corresponds to an increase in $d_L - c_L$ in example (c), a reduction in Δ in example (d), an increase in π^d and/or in $\pi^m - \pi^d$ in example (e), and an increase in $\mathscr{L}_L - \ell_L$, or a reduction in c_L, in example (f), in Subsection 2.2.

marginal benefit from cognition under the sufficient condition for expectation conformity identified in Proposition 1:

$$\frac{\partial^2}{\partial\theta\partial\rho}\left[\int_{-\infty}^{m^*(r(\rho^\dagger;\theta),\theta)}[\delta_L(r(\rho^\dagger;\theta),m)+\theta]dG(m;\rho)\right]=G_\rho(m^*(r(\rho^\dagger;\theta),\theta);\rho)\leq 0.$$

The reason for this last result is the following: Holding player F's reaction fixed, cognition reduces the probability that player L engages, which is costly when the gains from engagement are large. This property helps clarify that it is only because of the adverse selection problem that larger gains from engagement contribute to expectation conformity. They make player F respond to an increase in the anticipated cognition by player L by reducing r more sharply, which in turn raises player L's value of cognition.

4 Cognitive Traps, Disclosure, and Cognitive Styles

We now turn to three phenomena that are intrinsically related to expectation conformity in the type of situations described above, cognitive traps, disclosure, and cognitive styles.

4.1 Cognitive traps

Assume that ρ captures the intensity of cognition, so that $C(\rho)$ is increasing in ρ.

Proposition 3 (***cognitive traps***)**.** *Suppose that Assumptions 1 and 2 hold and that ρ_1 and ρ_2 are both equilibrium cognitive levels, with $\rho_1<\rho_2$. If, for any $\rho^\dagger\in[\rho_1,\rho_2]$, $\mathrm{A}(\rho^\dagger)<0$ (which is the case when Assumption 3 also holds and $G_\rho(m^*(r(\rho^\dagger));\rho^\dagger)<0$ for all $\rho^\dagger\in[\rho_1,\rho_2]$, or when the distributions are Uniform, Pareto, or Exponential), then player L is better off in the low-cognition equilibrium ρ_1. Conversely, when for any $\rho^\dagger\in[\rho_1,\rho_2]$, $\mathrm{A}(\rho^\dagger)>0$, then player L is better off in the high-cognition equilibrium ρ_2.*

Proof: Under Assumptions 1 and 2, we have that, for any $\rho^\dagger\in[\rho_1,\rho_2]$, $dr(\rho^\dagger)/d\rho^\dagger\overset{\text{sgn}}{=}\mathrm{A}(\rho^\dagger)$. For any given r, player L's welfare is given by

$$\mathcal{V}(r)=\sup_\rho\left\{U_L(0)+\int_{-\infty}^{m^*(r)}\delta_L(r,m)dG(m;\rho)-C(\rho)\right\}.$$

The envelope theorem, along with the property that $\delta_L(r,m)$ is increasing in r under assumption 1, imply that $d\mathcal{V}(r)/dr>0$. The result then follows from the fact that $r(\rho_2)<r(\rho_1)$ when $\mathrm{A}(\rho^\dagger)<0$ for all $\rho^\dagger\in[\rho_1,\rho_2]$, whereas $r(\rho_2)>r(\rho_1)$ when $\mathrm{A}(\rho^\dagger)>0$ for all $\rho^\dagger\in[\rho_1,\rho_2]$. ■

Cognitive traps do not result just from the fact that, when $C(\rho)$ is increasing, in a high-cognition equilibrium, player L spends more resources in cognition. In fact, at the margin, player L's gain from a more informative structure is equal to the increase in the cost of information acquisition.

Rather, cognitive traps occur because player F, anticipating an exacerbated adverse selection problem when expecting player L to invest more in cognition, reacts in an unfriendlier way, which not only forces player L to acquire more information, vindicating player F's expectation, but hurts player L.

The result in the previous proposition contrasts with what one obtains in markets without adverse selection. To see this, consider a setting in which player F is a monopolistic seller maximizing expected profits $p - c(\omega)$ by means of a take-it-or-leave-it offer p, whereas player L is a monopsonistic buyer choosing how much information ρ to acquire about her gross value ω for the seller's product and whether or not to accept the seller's offer of $r = -p$ so as to maximize her net payoff $\omega - p - C(\rho)$. When the seller's cost c is invariant in ω, this model corresponds to the private-value setting of Ravid, Roesler, and Szentes (2022). In such a setting, when information is free and the buyer can choose any mean-preserving contraction $G(\cdot;\rho)$ of the prior distribution G at no cost (so that $C(\rho) = 0$ for all ρ, with higher ρ denoting mean-preserving spreads), there are multiple equilibria. All equilibria are Pareto ranked, with each player's payoff maximized in the equilibrium in which the buyer fully learns the state. When, instead, payoffs are interdependent with c increasing with ω, (e.g., $c(\omega) = \omega - \Delta$ for all ω), the result in Proposition 3 suggests that, when more cognition by the buyer aggravates the adverse selection problem, which is consistent with the distributions $G(\cdot;\rho)$ being mean-preserving contractions of G, then equilibria in which the buyer acquires more information are equilibria in which the buyer is necessarily worse off, no matter the cost of information, and hence a fortiori when $C(\rho) = 0$ for all ρ. The reason for the difference is precisely the negative effect that more cognition exerts on the severity of the adverse selection problem, which induces the seller to ask for a higher price.

Next, consider markets in which player F is competitive, as in Akerlof's original lemons model. The result in the previous proposition implies that, when cognition aggravates adverse selection, in the presence of multiple equilibria, the equilibria are Pareto ranked, with the buyer's payoff constant across equilibria and with the seller's payoff decreasing in the equilibrium cognitive level.

The result in Proposition 3 calls for government interventions aimed at discouraging the players from acquiring information. We discuss some of these interventions in Section 6. Here, instead, we want to emphasize that cognitive traps are intrinsically related to expectation conformity. Recall that expectation conformity relates to the benefit of cognition in strategic settings. It does not depend on the cost of information. When the sufficient conditions for expectation conformity of Proposition 1 hold, however, one can identify cost functionals for which multiple equilibria arise.[23]

[23]Namely, suppose that Assumptions 1, 2, and 3 hold and that there exist ρ_1 and ρ_2, with $\rho_2 > \rho_1$, such that $G_\rho(m^*(r(\rho^\dagger));\rho) < 0$ for all $\rho^\dagger, \rho \in [\rho_1, \rho_2]$. There exist monotone cost functionals $C(\cdot)$ such that ρ_1 and ρ_2 are both equilibrium cognitive levels. Furthermore, under any such cost functionals, player L is better off in the low-cognition equilibrium ρ_1 than in the high-cognition equilibrium ρ_2. See also Pavan and Tirole (2022b) for a

4.2 Disclosure and Cognitive Style

We have so far assumed that information acquisition is covert. Suppose that it is indeed covert, but that some form of disclosure prior to F's action is feasible: player L can disclose how much attention or external resources she devoted to the issue (but not what she actually learnt).[24] Namely, given her actual cognition ρ, she can prove that her choice of cognition is above any level $\hat{\rho} \leq \rho$. The disclosed information is hard. For an arbitrary disclosure $\hat{\rho}$, one can consider the "$\hat{\rho}$-constrained cognition game," namely the no-disclosure game with modified cost function $D(\rho, \hat{\rho}) = C(\rho)$ if $\rho \geq \hat{\rho}$ and $D(\rho, \hat{\rho}) = +\infty$ if $\rho < \hat{\rho}$. Let $E(\hat{\rho})$ denote the set of equilibrium cognitive levels of the $\hat{\rho}$-constrained cognition game.

Definition 5 (**monotone selections**)**.** For any $\hat{\rho}$, let $e(\hat{\rho}) \in E(\hat{\rho})$ be an equilibrium of the "$\hat{\rho}$-constrained cognition game". The equilibrium selection $e(\cdot)$ is *monotone* if for all $\hat{\rho}$ and $\hat{\rho}\,'$, with $\hat{\rho} < \hat{\rho}\,'$, $e(\hat{\rho}) \leq e(\hat{\rho}\,')$.

In words, when L can choose only among higher cognitive levels, in equilibrium, she selects a higher cognitive level. Note that, because $E(\hat{\rho}) \cap \{\rho \mid \rho \geq \hat{\rho'}\} \subseteq E(\hat{\rho'})$, it is always possible to construct monotone selections.

Definition 6 (**regularity**)**.** Take any equilibrium of the primitive game with disclosure. The equilibrium is *regular* if the selection $e(\cdot)$ associated with such an equilibrium is monotone and $e(0)$ is an equilibrium of the no-disclosure game.

That is, an equilibrium of the game with disclosure is regular when, after disclosing a higher cognitive level, player L's actual cognition is higher. Clearly, in any pure-strategy equilibrium of the game with disclosure, L selects a unique level of cognition on path. Let such an equilibrium level be ρ^* and denote by $\hat{\rho}(\rho^*)$ the information L discloses in the pure-strategy equilibrium supporting ρ^*. The equilibrium being regular means that the strategy that L follows in such an equilibrium is such that, if she were to disclose any $\hat{\rho} < \hat{\rho}(\rho^*)$ (alternatively, any $\hat{\rho} > \hat{\rho}(\rho^*)$), in the continuation game, she would then select a cognition below ρ^* (alternatively, above ρ^*).

Let $\bar{\rho}$ be the highest equilibrium cognitive level in the game without disclosure. It is easy to see that, without the above refinement, the game with disclosure may admit equilibria supporting cognitive levels ρ^* strictly above $\bar{\rho}$. These equilibria can be sustained by a strategy for player L according to which, on path, L discloses $\hat{\rho}(\rho^*) = \rho^* > \bar{\rho}$. Off path, after disclosing any $\hat{\rho} \leq \rho^*$, L select a cognitive level above ρ^* anticipating a low reaction by player F, supported by the

similar point.

[24]Shishkin (2022) studies an evidence acquisition game. He shows that, when the probability of obtaining information is small, the Sender's optimal policy has a pass/fail structure and reveals only whether the quality is above or below a threshold. The game considered in this section differs in two respects: First, the acquired information is soft; second, the acquisition is either overt or "semi-overt" in that the intensity of information acquisition can be disclosed but not the actual information obtained.

expectation of a large cognition by player L. In other words, without the refinement, there is not enough connection between the equilibrium cognitive levels of the game with and without disclosure.

Proposition 4 (***disclosure***). *Assume that Assumptions 1 and 2 hold and that* $\mathrm{A}(\rho^\dagger) < 0$ *for all* $\rho^\dagger$*, implying that cognition always aggravates adverse selection.*

- *Any equilibrium cognitive level of the game in which disclosure is not feasible is also an equilibrium cognitive level in the disclosure game.*

- *Conversely, the largest and smallest cognitive levels in the regular equilibrium set of the disclosure game are also equilibrium cognitive levels in the game without disclosure.*

Proof. (i) The logic behind Proposition 4 is similar to the one behind Proposition 3. Consider an equilibrium of the game without disclosure in which cognition is equal to ρ^*. To see that ρ^* can also be supported in the game with disclosure, let $e(\hat{\rho})$ denote the level of cognition selected by L when disclosing $\hat{\rho}$. Consider the following strategy for L in the game with disclosure. For any $\hat{\rho} \leq \rho^*$, $e(\hat{\rho}) = \rho^*$, whereas for any $\hat{\rho} > \rho^*$, $e(\hat{\rho}) \geq \hat{\rho}$ (the precise value is not important). Under Assumption 2, that $\mathrm{A}(\rho^\dagger) < 0$ for all $\rho^\dagger$ implies F's reaction $r(\rho^\dagger)$ is non-increasing in $\rho^\dagger$. Hence, for any $\hat{\rho} > \rho^*$, F's reaction is $r(e(\hat{\rho})) \leq r^* \equiv r(\rho^*)$, whereas, for any $\hat{\rho} \leq \rho^*$, F's reaction is $r(e(\hat{\rho})) = r^*$. These properties imply that

$$\sup_{\{\rho,\hat{\rho}\leq\rho\}} \left\{ \int_{-\infty}^{m^*(r(e(\hat{\rho})))} \delta_L(r(e(\hat{\rho})), m) dG(m;\rho) - C(\rho) \right\} = \int_{-\infty}^{m^*(r^*)} \delta_L(r^*, m) dG(m;\rho^*) - C(\rho^*),$$

where the equality follow from the fact that ρ^* is an equilibrium of the no-disclosure game along with the fact that L's payoff is non-decreasing in F's reaction by Assumption 1.
(ii) Conversely, let ρ^* be a cognitive level supported by a regular equilibrium of the disclosure game (with associated disclosure $\hat{\rho}(\rho^*) \leq \rho^*$). Suppose that $\rho^* < \underline{\rho}$, where $\underline{\rho}$ is the lowest equilibrium cognitive level of the no-disclosure game. That the equilibrium supporting ρ^* is regular implies that the payoff that L obtains by selecting ρ^* and disclosing $\hat{\rho}(\rho^*)$ is equal to the payoff that she obtains by selecting ρ^* and disclosing any $\hat{\rho} \leq \hat{\rho}(\rho^*)$ (recall that F's reaction is decreasing in $\hat{\rho}$ in any regular equilibrium) and is higher than the payoff that she obtains by selecting any other cognitive level ρ' and disclosing any information $\hat{\rho}$. But this means that cognition ρ^* can also be sustained in the no-disclosure game, a contradiction. Similar arguments imply that the highest cognition level that can be sustained in any regular equilibrium of the disclosure game is $\bar{\rho}$, where $\bar{\rho}$ is the largest cognitive equilibrium level in the no-disclosure game.
■

Once again, the irrelevance of disclosure in this class of games is related to expectation conformity. When, fixing F's reaction, higher cognition reduces the chances that L engages (i.e., when, for any $\rho^\dagger$, $G_\rho(m^*(r(\rho^\dagger));\rho^\dagger) < 0$), cognition aggravates adverse selection making L's value for cognition increase with the level $\rho^\dagger$ anticipated by F (Proposition 1). The same condition implies that player L never gains from proving that her cognition is large, when higher disclosures are interpreted as informative of higher actual cognition.

In the same vein, one can consider the possibility of *transparency*, namely a commitment to reveal the exact amount of cognitive resources selected. In this case, $\hat{\rho} = \rho$ for any ρ (overt information acquisition). Clearly, player L is better off committing ex ante to transparency than retaining the possibility to disclose information voluntary ex post (the case just studied). She is also better off under transparency than in the game with complete absence of any disclosure. More interestingly, when F's reaction is non-increasing in $\rho^\dagger$ (which is the case when $\mathrm{A}(\rho^\dagger) < 0$ for all $\rho^\dagger$), under transparency, player L will choose a cognitive level $\rho^* \leq \underline{\rho}$ that is lower than the lowest equilibrium cognitive level in the no-disclosure game.

Similar conclusions obtain when player L cannot reveal her cognition perfectly, but can prove that it is below some level $\hat{\rho}$ of her choosing, for example by proving that she is unable to undertake more than a certain number of informative tests. In such situations, player L always proves that her cognition is below the lowest equilibrium level of the no disclosure game.

Another focus of comparative statics concerns player L's *cognitive style*. We provide here only an informal account. In the generalized lemons model, suppose that the cost-of-cognition function $C(\rho,\xi)$ depends on a parameter ξ, interpreted as cognitive ability. A higher-ability player L has a lower marginal cost of cognition: for any (ρ,ξ), $C_{\rho\xi}(\rho,\xi) < 0$, $C(0,\xi) = 0$, $C_\rho(0,\xi) = 0$, $C_\rho(\rho,\xi) > 0$, and $C_{\rho\rho}(\rho,\xi) > 0$. As player L's ability increases, the equilibrium cognition increases (in case of multiple equilibria, in the sense of monotone comparative statics: the lowest and highest cognitive levels of the equilibrium set corresponding to ability ξ increase with ξ). Put it differently, player L's ability, while directly beneficial, indirectly hurts her as player F becomes more wary of adverse selection. This suggests that, if player L has side opportunities to signal cognitive ability, she will want to adopt a dumbed-down profile.

Suppose indeed that player L can be bright (ξ_H) or dumb (ξ_L). A bright person can demonstrate she is bright (and can always mimic a dumb one), but the reverse is impossible. The set of equilibrium cognitive levels is monotonically increasing in the posterior probability that $\xi = \xi_H$. Let us assume a monotone selection in this equilibrium set: Player F's action r is decreasing in the probability that she assigns to $\xi = \xi_H$ (a property automatically satisfied if the equilibrium is unique, for any possible belief). Then if we add a disclosure game prior to the cognitive game in which player L can disclose she is bright if this is indeed the case, the equilibrium is a pooling one, in which the bright player L does not disclose her brightness. Conversely, player L will disclose, if she can, that she is overloaded with work (assume that she cannot prove that she has a low workload), and therefore that her marginal cost of cognition is high. In either case, player L poses as an informational puppy dog (in the sense of Fudenberg and Tirole (1984)).

5 Anti-lemons

The results in the previous section can be applied, with appropriate modifications, to environments that do not satisfy Assumption 2 above. To see this, suppose that an increase in anticipated cognition increases the friendliness of F's reaction, instead of reducing it, when it aggravates the adverse selection problem, i.e., Assumption 2 is reversed and replaced by the following assumption:

Assumption 2′ (anti lemons). The friendliness of player F's reaction to an increase in player L's cognition depends negatively on the impact of L's cognition on the reduction of the adverse selection problem, as measured by the truncated conditional mean of the state of Nature:

$$\frac{dr(\rho^\dagger)}{d\rho^\dagger} \overset{\text{sgn}}{=} -\frac{\partial}{\partial\rho^\dagger}\,\mathrm{M}^-\big(\mathrm{m}^*(\mathrm{r}(\rho^\dagger));\ \rho^\dagger\big).$$

Hence, under Assumption 2', an increase in cognition by player L reduces the friendliness of player F's reaction when it increases the truncated mean $M^-\big(m^*(r(\rho^\dagger));\ \rho^\dagger\big)$; it increases it when it reduces it (the opposite of what assumed in the analysis above).

Under Assumption 2′, expectation conformity at $(\rho, \rho^\dagger)$ requires that $\mathrm{A}(\rho^\dagger)\mathrm{B}(\rho, \rho^\dagger) > 0$ (the opposite of the condition in Proposition 1). Because $\mathrm{A}(\rho^\dagger) < 0 <\mathrm{B}(\rho, \rho^\dagger)$ when

$$G_\rho(m^*(r(\rho^\dagger)); \rho^\dagger), G_\rho(m^*(r(\rho^\dagger)); \rho) \leq 0,$$

that is, when more cognition reduces the probability that player L engages no matter whether evaluated from L's or F's perspective, we have that, under the key condition in Proposition 1, expectation conformity never arises. This is because, when $G_\rho(m^*(r(\rho^\dagger)), \rho^\dagger) \leq 0$, an increase in the cognition expected from L by F always aggravates the adverse selection problem, but, in the anti-lemons case, this triggers a friendlier reaction by player F. In turn, because the marginal value of cognition decreases with the friendliness of player F's reaction when $G_\rho(m^*(r(\rho^\dagger)), \rho) \leq 0$, an increase in the cognition $\rho^\dagger$ anticipated by player F reduces the value for L to expand her cognition starting from ρ.

5.1 Examples

The following examples satisfy, under appropriate conditions, Assumptions 1 and 2′:

(g) *Start-up followed by liquidation.* An entrepreneur (player L) must decide whether to start a new business. Starting the project costs the entrepreneur $c_L > 0$ and generates cash flows equal to $1 - \omega$. Before being able to collect the project's cash flows, the entrepreneur may need to liquidate the project (for example, because of a preference shock that makes consumption at the time the projects payoffs off no longer valuable to L, as in Diamond and Dybvig (1983)).

Early liquidation occurs with probability p and results in the entrepreneur collecting a price r for the assets from a pool of risk-neutral competitive investors (player F). The entrepreneur's value from starting the project (i.e., the engagement decision in this application) is equal to $\delta_L = (1-p)(1-m) + pr - c_L$.
The entrepreneur thus starts the project if and only if $m < m^*(r) = (1-p+pr-c_L)/(1-p)$. Assuming, for simplicity, that the value of the project in the investors' hands is also equal to $1-m$, we then have that $r = 1 - M^-(m^*(r);\rho)$.

(h) *Spencian signaling.* An agent (player L) has disutility of effort ω for studying which is negatively correlated with the agent's productivity $\theta = a - b\omega$ from working on the relevant jobs after leaving school. The labor market is populated by competitive employers (player F) offering the agent a wage r equal to the agent's expected productivity. So, in this application, $\delta_L = r - m - p$, where p is the cost of enrolling in the school program under consideration (say, an MBA). Hence, the agent enrolls if and only if $m < m^*(r) = r - p$, with r satisfying $r = a - bM^-(m^*(r))$.

(i) *Warfare.* Country L is a potential invader and must decide whether to engage in a fight ($a = 1$) or abstain from doing so ($a = 0$). The state of nature ω represents the probability that country F wins in case of a fight. Let r denote the probability that country F surrenders without fighting back. The payoff that L obtains in case of victory is 1, whereas the cost of a defeat is c_L, implying that $\delta_L(r,m) = r + (1-r)(1-m-mc_L)$. Furthermore, in this game, L engages if and only if $m \leq m^*(r)$ where

$$m^*(r) = \frac{1}{(1-r)(1+c_L)}.$$

Similarly, letting country F's payoff from victory be equal to 1 and its loss in case of defeat be equal to c_F, we have that country F concedes if and only if

$$M^-(m^*(r(\rho^\dagger));\rho^\dagger) - \left(1 - M^-(m^*(r(\rho^\dagger));\rho^\dagger)\right) c_F \leq 0.$$

Assuming that c_F is drawn from some cumulative distribution H, we then have that $r(\rho^\dagger)$ is given by the solution to

$$r(\rho^\dagger) = 1 - H\left(\frac{M^-(m^*(r(\rho^\dagger));\rho^\dagger)}{1 - M^-(m^*(r(\rho^\dagger));\rho^\dagger)}\right). \tag{14}$$

Hence, this is an anti-lemon problem, in that the decision by player L to engage carries information that the state is one in which, if player F were to fight back, he would likely lose, thus making F play in a friendlier way towards player L. Whenever equation (14) admits a unique solution, Assumption A2' then holds: an increase in cognition by player L, when it leads to a reduction in $M^-(m^*(r(\rho^\dagger));\rho^\dagger)$, induces player F to respond with an action that is friendlier to

L (i.e., he surrounds more often).

(j) *Leadership.* Like in Hermalin (1998)'s theory of leadership, consider a setting in which a leader has information about the profitability of a project and benefits from binging on board a partner. Contrary to Hermalin (1998), however, assume that the leader's information is endogenous. Specifically, suppose that player L's gain from starting the project is $\delta_L(r, m) = 1 - m + r - c_L$, where $1 - m$ is the probability that the project succeeds, r is the probability that player F joins the venture, and c_L is L's cost of initiating the project. Player F, after observing L's decision to initiate the project, decides whether to join. If he does, his payoff is equal to $1 - m + 1 - c_F$, whereas, if he does not, it is equal to zero. Again, this is an anti-lemon problem, in that the decision by L to engage (here to start a project) is good news for player F, instead of bad news. Assuming that c_F is drawn from some cumulative distribution H, we then have that the probability that F joins is given by the solution to

$$r(\rho^\dagger) = H\left(2 - M^-\left(1 + r(\rho^\dagger) - c_L; \rho^\dagger\right)\right). \tag{15}$$

Hence, Assumption 2' holds whenever equation (15) admits a unique solution.

5.2 Expectation conformity

The following result summarizes the relationship between expectations and incentives for higher cognition in the anti-lemons case (the proof follows from the arguments above):

Proposition 5 (*expectation conformity – anti-lemons*)**.** *Suppose that Assumptions 1, 2', and 3 hold and that* $\mathrm{A}(\rho^\dagger) < 0$ *for all* $\rho^\dagger$.[25] *Then expectation conformity holds at* $(\rho, \rho^\dagger)$ *only if* $G_\rho\Big(m^*\big(r(\rho^\dagger)\big); \rho\Big) > 0$*, which, in the case of rotations, holds if and only if the cutoff* $m^*\big(r(\rho^\dagger)\big)$ *is to the left of the rotation point* m_ρ *(the opposite of the lemons case). Furthermore,* $G_\rho\Big(m^*\big(r(\rho^\dagger)\big); \rho\Big) > 0$ *is both necessary and sufficient for expectation conformity at* $(\rho, \rho^\dagger)$ *if* $\partial^2\delta(m, r)/\partial m \partial r = 0$ *for all* m *and* r *(as in examples (g), (h), and (j) above).*

Naturally, many of the results in the previous sections are thus reversed when Assumption 2 (lemons) is replaced with Assumption 2' (anti-lemons). For example, disclosure can be effective, and player L may want to appear an inoffensive cognitive fat cat in the anti-lemons case.

6 Policy Interventions

We now investigate how a benevolent government can improve over the laissez-faire equilibrium by subsidizing (alternatively, taxing) trade. We start with a fairly general analysis geared at shedding light on (1) what forces contribute to the optimality of subsidies (alternatively, taxes),

[25] *Recall that this is the case for example when information structures are Uniform, Pareto, or Exponential.*

and (2) how the endogeneity of information calls for larger (alternatively, smaller) interventions. We then apply the insights to the Akerlof model of Example (a) in Subsection 2.2 and to the environment of Example (b) where the government designs an asset purchase scheme to increase the efficiency of trade in a market affected by adverse selection.

Let $\delta_F(r, m) \equiv u_F(1, r, m) - u_F(0, m)$ denote the follower's payoff from responding with a reaction r to the leader's choice of engaging, when the leader's posterior mean (equivalently, the state) is m. Let s denote the subsidy (tax if $s < 0$) the planner promises to pay to player L in case of engagement.[26] Abusing notation, for any r and s, we let $m^*(r, s)$ denote the optimal engagement threshold for the leader when the follower's reaction is r and the subsidy is s, with m^* implicitly defined by the solution to $\delta_L(r, m^*)+s = 0$. Let $\rho^*(s)$ and $r^*(s)$ denote, respectively, the leader's equilibrium cognition and the follower's equilibrium response in the continuation game that starts after the planner announces a subsidy equal of s. Throughout, we assume that, for any s, $\rho^*(s)$ and $r^*(s)$ are unique, Lipschitz continuous, and differentiable. Likewise, we assume that the payoff functions $\delta_L(r, m)$ and $\delta_F(r, m)$ and the distributions $G(m; \rho)$ are differentiable and Lipschitz-continuous. In addition to facilitating the description of the relevant optimality conditions, these properties validate a certain envelope theorem we use in the Appendix to establish the results in this section.

For simplicity, we assume that player F is a representative of a competitive market in which case, for any s, given the leader's cognition $\rho^*(s)$, the follower's reaction $r^*(s)$ satisfies

$$\int_{-\infty}^{m^*(r^*(s),s)} \delta_F(r^*(s), m) dG(m; \rho^*(s)) = 0.$$

Many of the insights below extend to settings in which player F's expected payoff is different from zero and the planner cares about F's payoff.

For any s, total welfare is given by (up to scalars that are irrelevant for the analysis)

$$W(s) \equiv \int_{-\infty}^{m^*(r^*(s),s)} \left(\delta_L(r^*(s), m) + s\right) dG(m; \rho(s)) - C(\rho^*(s)) - (1 + \lambda) s G(m^*(r(s), s); \rho^*(s)),$$

where $\lambda \geq 0$ is the unit cost of public funds. The first two terms represent the leader's payoff, whereas the last term represents the cost of the program to the government. Hereafter, we assume

[26] More generally, both the decision to engage and that of not engaging can be subject to taxes and subsidies. For example, in the Akerlof model of Example (a) in Subsection 2.2, the decision to hold on a car or a security can be taxed. Likewise, in the herding entry game of Example (e) of Subsection 2.2, firm L could be taxed or subsidized regardless of the particular entry decision. Hence, in the analysis below, s should be interpreted as the differential in the subsidy/tax when player L engages relative to when she does not engage.

that W is strictly quasi-concave and that, in the laissez-faire equilibrium,

$$\frac{d}{dr}\delta_F(r, M^-(m^*(r,0);\rho^*(0)))\Big|_{r=r^*(0)} < 0. \tag{16}$$

The condition simply says that, starting from the laissez-faire equilibrium reaction $r^*(0)$, a small increase in the friendliness of the follower's response reduces the follower's payoff. In the Akerlof model, for example, this condition always holds when $G(m;\rho)/g(m;\rho)$ is increasing in m, which, as discussed above, implies that, for any ρ, $\frac{\partial}{\partial m^*}M^-(m^*;\rho) < 1$ which in turn guarantees that, given ρ, the equilibrium's response of the follower is unique and that Assumption 2 is satisfied.[27]

Proposition 6 (***social value of subsidizing/taxing trade***)**.** *Suppose Assumption 1 holds. In the lemons case (i.e., when Assumption 2 also holds), there exists a constant $K > 0$ such that a strictly positive subsidy is optimal if $\frac{d}{ds}M^-(m^*(r^*(0),s);\rho^*(s))\big|_{s=0} > K$, whereas a tax on engagement is optimal when the above inequality is reversed. When, instead, Assumption 2' holds (anti-lemons), there exists a constant $K < 0$ such that a strictly positive subsidy is optimal if $\frac{d}{ds}M^-(m^*(r^*(0),s);\rho^*(s))\big|_{s=0} < K$, whereas a tax on engagement is optimal when $\frac{d}{ds}M^-(m^*(r^*(0),s);\rho^*(s))\big|_{s=0} > K$.*

Hence, whether subsiding trade is preferable to taxing it depends on whether the government faces a lemon or an anti-lemon problem and whether subsidizing trade aggravates or alleviates adverse selection. Note that

$$\frac{d}{ds}M^-(m^*(r^*(0),s);\rho^*(s))\big|_{s=0} = \frac{\partial}{\partial m^*}M^-(m^*(r^*(0),0);\rho^*(0))\ \frac{\partial m^*(r^*(0),s)}{\partial s}\Big|_{s=0}$$

$$+\frac{\partial}{\partial \rho}M^-(m^*(r^*(0),0);\rho^*(0))\frac{d\rho^*(0)}{ds}.$$

When information is endogenous, there are two channels through which a subsidy alleviates (alternatively, aggravates) the adverse selection problem. The first one is through its effect on the leader's engagement, as captured by the threshold m^*. The second one is through its effect on the leader's cognition. A higher subsidy always increases the engagement threshold m^*. Because, for any ρ, M^- is increasing in m^*, the first effect always contributes to alleviating adverse selection. Hence, through this channel, the planner induces a friendlier reaction by player F when Assumption 2 holds (lemons), and a more adversarial one when, instead, Assumption 2' holds (anti-lemons). The second effect, instead, can be either positive or negative, depending on whether cognition aggravates or alleviates adverse selection and whether a positive subsidy increases or decreases the leader's equilibrium cognitive level.

[27] The case in which the condition is violated is not particularly interesting. For example, in the version of the Akerlof model in which the seller observes r before choosing whether or not to engage, when the condition is violated, there exists $r' > r^*(0)$ such that any of the buyers, by deviating to r', would induce the seller to trade with them with certainty making a strictly positive profit.

The threshold K, whose formula is in the Appendix, depends on the primitives of the environment. In the Akerlof model, for example, $K = \lambda$, i.e., the threshold coincides with the unit cost of public funds, and the following conditions jointly imply that subsidizing trade is optimal:

1. $\frac{\partial}{\partial m^*} M^-(m^*(r^*(0), 0); \rho^*(0)) > \lambda$;
2. $\frac{\partial}{\partial \rho} M^-(m^*(r^*(0), 0); \rho^*(0)) < 0$;
3. $d\rho^*(0)/ds < 0$.

The first condition is satisfied when the unit cost of public funds, λ, is small. When information is exogenous, this condition is jointly necessary and sufficient for a positive subsidy on trade to be optimal. When, instead, information is endogenous, the other two conditions also play a key role. The second condition says that cognition aggravates adverse selection. From Proposition 1, we know that this condition always holds when information structures are consistent with the MPS order and $G_\rho(m^*(r^*(0), 0); \rho^*(0)) < 0$, i.e., cognition reduces the probability the seller engages by putting the asset on sale.[28] Condition 3, in turn holds when, in addition to $G_\rho(m^*(r^*(0), 0); \rho^*(0)) < 0$ (which, as shown in Proposition 1, also implies that the benefit of expanding cognition decreases with the friendliness of the follower's reaction), the comparative statics of the equilibrium have the same monotonicity as those of the best responses.[29]

Next we turn to the effects of the endogeneity of information on the optimal policy. Let s^* denote the optimal policy when information is endogenous. Now suppose information is exogenous and equal to $\rho = \rho^*(s^*)$, where $\rho^*(s^*)$ is the equilibrium cognition when information is endogenous and the subsidy is equal to s^*. For any s, let $\hat{r}(s)$ denote the follower's equilibrium reaction when the subsidy is equal to s and information is exogenous and equal to $\rho^*(s^*)$. Clearly, for $s = s^*$, $\hat{r}(s^*) = r^*(s^*)$, where $r^*(s^*)$ is the equilibrium reaction when information is endogenous. Let $W^{\#}(s)$ denote welfare when information is exogenous and equal to $\rho^*(s^*)$. Hereafter, we assume that $W^{\#}(s)$ is strictly quasi-concave and then denote by s^{**} the level of the policy that maximizes $W^{\#}(s)$. Finally, for any (r, s), let

$$\hat{W}(r, s) \equiv \int_{-\infty}^{m^*(r,s)} (\delta_L(r, m) + s)\, dG(m; \rho^*(s^*)) - C(\rho^*(s^*)) - (1 + \lambda) s G(m^*(r, s); \rho^*(s^*))$$

denote the level of welfare that is attained when information is exogenous and equal to $\rho = \rho^*(s^*)$, the follower's reaction is r, the subsidy is s, and the leader engages if and only if $m < m^*(r, s)$.[30]

We then have the following result:

[28] Also recall that cognition always aggravates adverse selection when the distributions from which the mean m is drawn are Uniform, Pareto, or Exponential.

[29] Note that, holding cognition fixed at $\rho^*(0)$, an increase in the subsidy, starting from $s = 0$, always increases the friendliness of the follower's reaction. Likewise, holding r fixed at $r^*(0)$, an increase in the subsidy, starting from $s = 0$, always reduces the leader's cognition when $G_\rho(m^*(r^*(0), 0); \rho^*(0)) < 0$.

[30] Clearly, $W^{\#}(s) = \hat{W}(\hat{r}(s), s)$.

Proposition 7 (***effect of endogenous information on optimal policy***)**.** *The endogeneity of the leader's information calls for larger policy interventions (i.e., $s^* > s^{**}$) if*

$$\left(\frac{d\hat{r}(s^*)}{ds} - \frac{dr^*(s^*)}{ds}\right)\frac{\partial \hat{W}(r^*(s^*), s^*)}{\partial r} + (1+\lambda)s^* G_\rho(m^*(r^*(s^*), s^*); \rho^*(s^*))\frac{d\rho^*(s^*)}{ds} < 0,$$

whereas the opposite is true (i.e., $s^* < s^{**}$) if the above inequality is reversed.

The result is intuitive. The endogeneity of the leader's information calls for larger policies when (a) $\partial \hat{W}(r^*(s^*), s^*)/\partial r > 0$, meaning that the social value of increasing the follower's reaction beyond $r^*(s^*)$ is positive, accounting for the fact that a friendlier reaction induces more engagement which in turn comes with a larger cost to the government, (b) an increase in the subsidy, starting from s^*, triggers a larger response by the follower when information is endogenous than when it is exogenous, i.e., $dr^*(s^*)/ds > d\hat{r}(s^*)/ds$, and (c) the extra cost

$$(1+\lambda)s^* G_\rho(m^*(r(s^*), s^*); \rho^*(s^*))\frac{d\rho^*(s^*)}{ds} \tag{17}$$

that the planner incurs due to the endogeneity of information is small. Note that, when $s^* > 0$, $G_\rho(m^*(r(s^*), s^*); \rho^*(s^*)) < 0$ (which, under the MPS order, is the key condition for EC identified in Proposition 1), and $d\rho^*(s^*)/ds < 0$, the term in (17) is positive: the planner expects to pay s^* more often when the leader reduces her cognition in response to a larger subsidy. As a result, this last effect contributes to a lower policy when information is endogenous.

Next note that, when information is endogenous, the optimality of s^* reveals that

$$\frac{dr^*(s^*)}{ds}\frac{\partial \hat{W}(r^*(s^*),s^*)}{\partial r} = (1+\lambda)s^* g(m^*(r^*(s^*), s^*); \rho^*(s^*))\frac{\partial m^*(r^*(s^*),s^*)}{\partial s} + \lambda G(m^*(r^*(s^*), s); \rho^*(s^*))$$
$$+(1+\lambda)s^* \frac{d\rho^*(s^*)}{ds} G_\rho(m^*(r^*(s^*), s^*); \rho^*(s^*)).$$

Hence, when $s^* > 0$, $dr^*(s^*)/ds > 0$, and $d\rho^*(s^*)/ds G_\rho(m^*(r(s^*), s^*); \rho^*(s^*)) > 0$, necessarily $\partial \hat{W}(r^*(s^*), s^*)/\partial r > 0$. That is, under the welfare-maximizing policy s^*, welfare always increases with the friendliness of the follower's response when cognition reduces the probability of engagement (i.e., when $G_\rho(m^*(r(s^*), s^*); \rho^*(s^*)) < 0$) and the comparative statics of the equilibrium r and ρ have the same monotonicity as those of the best responses (i.e., r^* increases and ρ^* decrease with the subsidy).

To gauge some intuition about whether a higher subsidy triggers a larger response by the follower when information is endogenous than when it is exogenous, consider the interesting case in which, starting from $r = r^*(s^*)$, and holding cognition fixed at $\rho = \rho^*(s^*)$, an increase in r

reduces the follower's payoff, that is,[31]

$$\frac{d}{dr}\delta_F(r, M^-(m^*(r,s^*);\rho^*(s^*)))\Big|_{r=r^*(s^*)} < 0.$$

Then
$$\frac{dr^*(s^*)}{ds} - \frac{d\hat{r}(s^*)}{ds} \overset{sgn}{=} \frac{\partial\delta_F(r^*(s^*),m)}{\partial m}\frac{\partial}{\partial\rho}M^-(m^*(r^*(s^*),s^*);\rho^*(s^*))\frac{d\rho^*(s^*)}{ds},$$

where $\partial\delta_F(r^*(s^*),m)/\partial m$ is the sensitivity of the follower's payoff to the state (which is invariant in m under the maintained assumption that δ_F is affine in m). Hence, in the lemons case (i.e., when Assumption 2 holds, in which case $\partial\delta_F(r^*(s^*),m)/\partial m > 0$), an increase in the subsidy leads to a larger response by the follower under endogenous information when

$$\frac{\partial}{\partial\rho}M^-(m^*(r^*(s^*),s^*);\rho^*(s^*))\frac{d\rho^*(s^*)}{ds} > 0 \tag{18}$$

and a smaller response when the inequality is reversed. The opposite conclusions holds in the anti-lemon case (i.e., under Assumption 2', in which case $\partial\delta_F(r^*(s^*),m)/\partial m < 0$). This is also intuitive. Consider the case in which Assumption 2 holds (i.e., the lemons case). The follower responds more to an increase in the subsidy when information is endogenous than when it is exogenous if the increase in the subsidy leads to a reduction in cognition and, as a result of it, an alleviation of the adverse selection problem.

When applied to the Akerlof model, the above insights lead to the following:

Corollary 2 (***double dividend of the subsidy in Akerlof model***)**.** *Consider the Akerlof model of Example (a) in Subsection 2.2 and let s^* denote the optimal subsidy when information is endogenous. Assume that $G(m;\rho^*(s^*))/g(m;\rho^*(s^*))$ is increasing in m, information structures are consistent with the MPS order, and $G_\rho(m^*(r^*(s^*),s^*);\rho^*(s^*)) < 0$ (meaning that cognition aggravates adverse selection). Then, when information is exogenous and equal to $\rho^*(s^*)$, the optimal subsidy, s^{**}, satisfies $s^{**} < s^*$.*

Recall that the property that $G(m;\rho^*(s^*))/g(m;\rho^*(s^*))$ is increasing in m guarantees that Assumption 2 holds in the Akerlof model (that is, the friendliness of the follower's reaction increases with the leader's cognition if and only if a higher cognition alleviates adverse selection, i.e., it increases $M^-(m^*(r^*(s^*),s^*;\rho^*(s^*)))$. That information structures are consistent with the MPS order and $G_\rho(m^*(r^*(s^*),s^*);\rho^*(s^*)) < 0$ in turn implies that, under the optimal subsidy s^*, cognition aggravates adverse selection (see Proposition 1). Therefore, under the assumptions in the corollary, starting from s^*, if the planner were to cut the subsidy, it would trigger a larger reduction in the price offered by the buyers when information is endogenous than when it is exogenous. The optimal subsidy is thus larger under endogenous information.

[31]Note that this condition is the analog of Condition (16) above, but applied to $s = s^*$ instead of $s = 0$.

The same conclusions apply to the application of Example (b) in Subsection 2.2 where the government directly controls the price at which the sellers can trade in their assets. To see this, note that this application is formally equivalent to the Akerlof model considered above (in this application, given the price r set by the government, the private sector pays $M^{-}(r;\rho)+\Delta$ to the sellers and hence it is as if the government pays a subsidy $s = r - M^{-}(r;\rho) - \Delta$ to the sellers). The result in the previous corollary thus implies that, when information is endogenous, it is optimal for the government to design an asset-purchase program with a higher price.

The results above point to a general insight. When increasing trade is socially beneficial, cognition aggravates adverse selection, and a friendlier reaction by player F reduces the marginal value of information for player L (which is the case under the conditions for EC in Proposition 1 and for cognitive traps in Proposition 3), the social value of subsidizing trade is higher when information is endogenous than when it is exogenous. This is because subsidizing trade comes with a *double dividend*: in addition to inducing player L to engage more often, it induces L to acquire less information which in turn alleviates adverse selection and further boost welfare.

Table 1 in the Appendix summarizes some of the key results from this section and the previous ones.

7 Discussion

In this section, we discuss the connection to other investment games, and the robustness of the key insights to more flexible information-acquisition technologies.

7.1 Relation to other covert investment games

The paper's emphasis is on information acquisition, a choice motivated both by the applications and by the fact that cognitive investments are the ultimate covert investments. But the results may also be useful for other covert investments: capacity acquisition, learning by doing, arms buildup, and so on.

Suppose that there are two players, playing a "second-stage" normal-form game with actions a_L, $a_F \in \mathbb{R}$. One of the players, here player L, makes a "first-stage" investment $\rho \in \mathbb{R}$ at an increasing investment cost $C(\rho)$.[32] Payoffs are $\phi_L(a_L, a_F) - \psi(a_L, \rho) - C(\rho)$ for player L and $\phi_F(a_L, a_F)$ for player F, where all functions are C^2 and satisfy $\partial^2\psi/\partial a_L \partial\rho < 0$ and

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

[32]The analysis can be extended to the case where both players make period-1 investments. The insights are not fundamentally different from those discussed here.

for $i, j = L, F$, $j \neq i$.

That is, the investment ρ lowers player L's marginal cost of action a_L, and the strategic interaction between the two players involves either strategic complementarity (SC) or strategic substitutability (SS). For example, a_i may stand for firm i's output, ρ an investment that lowers the marginal cost of production, and the two firms' output choices may be either strategic complements or substitutes.

Assume, for simplicity, that, if player L's investment was common knowledge, the normal-form game in (a_1, a_2) would have a unique and stable equilibrium. In such a game, player F's equilibrium action $a_F(\rho^\dagger)$ as a function of player L's anticipated investment $\rho^\dagger$, is increasing in $\rho^\dagger$ under SC and decreasing in $\rho^\dagger$ under SS.

Consistently with the analysis above, suppose that player L's actual investment ρ is not observed by player F (so de facto the game is a simultaneous-move game in actions (ρ, a_L), for player L, and a_F, for player F). One can then define player L's optimal action when she deviates from her equilibrium investment. The above assumptions imply that player L's optimal action $a_L(\rho, \rho^\dagger)$ when player F expects $\rho^\dagger$ and player L's actual investment is ρ is non-decreasing in $\rho^\dagger$ under either SC or SS.

This environment is similar to that considered in the industrial organization literature on the taxonomy of business strategies[33], except for one important twist. The investment choice ρ is not observed by player F and so has no commitment effect; rather, what matters for the outcome of the normal-form game is the anticipation $\rho^\dagger$ by firm F of firm L's choice as well as the actual choice ρ by firm L (of course, in a pure-strategy equilibrium, $\rho^\dagger = \rho$).

Let $T_L(\rho, \rho^\dagger) \equiv \max_{a_L} \left\{ \phi_L(a_L, a_F(\rho^\dagger)) - \psi(a_L, \rho) - C(\rho) \right\}$ denote player L's payoff when her actual investment is ρ and player F anticipates investment $\rho^\dagger$. The above assumptions imply that, whether SC or SS prevails, for all $(\rho, \rho^\dagger)$ and $(\hat{\rho}, \hat{\rho}^\dagger)$ with $\hat{\rho} \geq \rho$ and $\hat{\rho}^\dagger \geq \rho^\dagger$, the following "expectation conformity" condition is satisfied:

$$T_L(\hat{\rho}, \hat{\rho}^\dagger) - T_L(\rho, \hat{\rho}^\dagger) \geq T_L(\hat{\rho}, \rho^\dagger) - T_L(\rho, \rho^\dagger).$$

Consequently, let ρ (alternatively, $\hat{\rho}$) denote player L's optimal investment when player F expects investment $\rho^\dagger$ (alternatively, $\hat{\rho}^\dagger$).[34] Expectation conformity implies that there is complementarity between investment and anticipation of investment: $(\hat{\rho}_1 - \rho_1)(\hat{\rho}_1^\dagger - \rho_1^\dagger) \geq 0$. This is so both when the stage-2 game involves strategic substitutes or strategic complements.

[33] See, e.g., Bulow et al (1985), and Fudenberg and Tirole (1984).

[34] That is, given $\rho^\dagger$, (a_L, ρ) is such that

$$(a_L, \rho) \in \arg\max_{\tilde{\rho}, \tilde{a}_L} \left\{ \phi_L(\tilde{a}_L, a_F(\rho^\dagger)) - \psi_L(\tilde{a}_L, \tilde{\rho}) - C(\tilde{\rho}) \right\}$$

with $\{a_L(\rho^\dagger), a_F(\rho^\dagger)\}$ denoting the Nash equilibrium of the normal-form game under common knowledge that L invested $\rho^\dagger$ (i.e., under symmetric information).

The intuition goes as follows: Suppose that firm F expects L to invest more and therefore to produce more output. It then raises its output under SC and decreases it under SS. In either case, firm L is induced to raise its output, vindicating a higher investment in the first place. It can also be checked that when there are two equilibria ($\rho = \rho^\dagger$ and $\hat{\rho} = \hat{\rho}^\dagger$), player L is better off in the high-investment one, again regardless of the type of strategic interaction (SC or SS).

Let us draw a formal analogy between the generalized lemons game of this paper and the covert investment game described above. The investment ρ in the present paper is player L's cognition, that is, her choice of information structure. To interpret the generalized lemons game as a covert investment game, it suffices to assume that, in the "stage-2" game, player L wants to take an action equal to her investment. For example, one can think of a_L as the information used by player L in the stage-2 game. In this spirit, the assumption that L maximizes her payoff by choosing $a_L = \rho$ simply reflects the idea that player L makes full use of the acquired information (see also the discussion in the next section about how to interpret ρ as capacity in entropy reduction, or the maximal slope of the induced stochastic choice rule). In this case, $\psi(a_L, \rho) = 0$ if $a_L = \rho$, whereas $\psi(a_L, \rho) = -\infty$ otherwise, which is a discontinuous version of the complementarity relationship $\partial^2 \psi_L / \partial a_L \partial \rho < 0$ in the investment game. Letting $a_F = r$, we then have that

$$\phi_L(a_L, r) \equiv \int_{-\infty}^{m^*(r)} \delta_L(m, r) dG(m; a_L)$$

and so $\partial^2 \phi_L(r(\rho^\dagger), a_L)/\partial a_L \partial r < 0$ whenever condition $\mathrm{B}(\rho, \rho^\dagger) > 0$ in Proposition 1 is satisfied, for $a_L = \rho$. Furthermore, Condition $\mathrm{A}(\rho^\dagger) < 0$ in Proposition 1 implies that $dr/da_L < 0$. Summarizing, when the two conditions of Proposition 1 are met, the lemons game can be seen as an investment game with strategic substitutes (SS). In contrast, many anti-lemon games such as those considered in Section 5 under the conditions of Proposition 5 are investment games with strategic complements.

7.2 Flexible Information Acquisition

The analysis in the previous sections assumes that the set of experiments player L has access to leads to a collection of mean-preserving spreads (Assumption 3). The key forces responsible for expectation conformity identified in part (i) of Proposition 1, however, extend to more general information structures.

To see this, consider an arbitrary experiment $q : \Omega \to \Delta(Z)$ mapping states into probability distributions over a rich (Polish) space of signal realizations Z. Note that any such experiment, when combined with the prior G over Ω, leads to a distribution G^q of the posterior mean, m. Furthermore, when combined with the optimal engagement strategy (that is, with the strategy that, for any reaction r by player F, specifies to engage if and only $m \leq m^*(r)$), the experiment q leads to a stochastic choice rule $\sigma^q : \Omega \to [0, 1]$ specifying the probability that player L engages in each state.

Following the rational inattention literature, one can think of player L as choosing directly the rule $\sigma : \Omega \to [0,1]$ subject to an appropriate specification of the cost functional $\mathcal{C}(\sigma)$, with the interpretation that, for any σ, $\mathcal{C}(\sigma)$ is the cost of the cheapest experiment $q : \Omega \to \Delta(Z)$ that permits L to implement the stochastic choice rule σ. A couple of cost functionals that have been considered in the literature are those linked to "mutual information" and "maximal slope". Below we discuss both specifications and explain how our analysis can accommodate for these specifications.

For any experiment q, let

$$I^q = \int_\omega \int_z \ln(q(z|\omega))q(\mathrm{d}z|\omega)dG(\omega) - \int_z \ln\left(\int_\omega q(z|\omega)dG(\omega)\right)\int_\omega q(z|\omega)dG(\omega)$$

denote the mutual information between the random variables ω and z, where z is the random variable obtained by combining the prior G with the signal q. Now suppose that there exists a function $c : \mathbb{R}_+ \to \mathbb{R}_+$ such that, for any q, the cost of experiment q is given by $\mathbf{C}(q) = c(I^q)$. To facilitate the comparison with the analysis in the previous sections in which different information structures are indexed by a uni-dimensional parameter, with higher values indexing superior information structures, let player L's cognition determine the easiness by which L can absorb information. Specifically, assume that, for any $\rho \in \mathbb{R}_+$, player L's marginal cost of entropy reduction is $1/\rho$. To be able to process information at marginal cost $1/\rho$, player L must make a cognitive investment whose cost is $C(\rho)$, with the function C satisfying the same assumptions as in the baseline model. The difference is that, once ρ is chosen, player L can now pick *any* experiment q of her choice, with each experiment costing him $\frac{1}{\rho}c(I^q)$. For simplicity, one can then assume that c is the identity function (i.e., $c(I^q) = I^q$ for any q) so that the cost of each experiment q is given by the mutual information between its realizations z and the state ω (equivalently, by the reduction in entropy brought by the experiment).

Alternatively, one can let $\rho \in \mathbb{R}_+$ denote player L's "information capacity." Under this interpretation, given ρ, player L chooses the experiment that maximizes her expected payoff among those whose mutual information between ω and the realization z of the selected experiment is no greater than ρ.

It is well known that, for any cognitive investment ρ and any anticipated reaction r by player F, the experiment $q^{\rho,r}$ that maximizes player L's expected payoff net of the above cognitive cost is binary, i.e., for any ω, it assigns positive probability only to two signal realizations. Without loss of generality, label these signal realizations by $z = 1$ and $z = 0$, and interpret $z = 1$ as a "recommendation to engage" and $z = 0$ as a "recommendation to not engage." Letting $q^{\rho,r}(1|\omega)$ denote the probability that signal $q^{\rho,r}$ recommends $z = 1$ when the state is ω and $q^{\rho,r}(1) \equiv \int_\omega q^{\rho,r}(1|\omega)dG(\omega)$ the total probability of $z = 1$ under the experiment $q^{\rho,r}$, we have that the optimal signal is given by the solution to the following functional equation (see, e.g., Woodford (2009) and Yang (2015)):[35]

$$\delta_L(r,\omega) = \frac{1}{\rho}\left[\ln\left(\frac{q^{\rho,r}(1|\omega)}{1-q^{\rho,r}(1|\omega)}\right) - \ln\left(\frac{q^{\rho,r}(1)}{1-q^{\rho,r}(1)}\right)\right].$$

Next, consider the case in which the cost of inducing a stochastic choice rule $\sigma : \Omega \to [0,1]$ is given by $\mathcal{C}(\sigma) = c\left(\sup\left\{|\sigma'(\omega)|\right\}\right)$, where the function $c : \mathbb{R}_+ \cup \{+\infty\} \to \mathbb{R}_+ \cup \{+\infty\}$ is non-decreasing and satisfies $c(0) = 0$ and $c(k) < \infty$ for all $k \in \mathbb{R}_+$. Here $\sigma'(\omega)$ is the derivative of σ at ω. At any point of discontinuity of σ, $\sigma'(\omega) = +\infty$, whereas at any point ω at which σ is continuous but non-differentiable, $\sigma'(\omega)$ is the maximum between the left and the right derivative. Examples of this cost functional can be found in Robson (2001), Rayo and Becker (2007), Netzer (2009), and more recently Morris and Yang (2021).

Again, to see the connection with the analysis in the previous sections, one can think of player L's cognition as determining the maximal slope of her stochastic choice rule, selected by the player at cost $C(\rho)$ with C satisfying the same properties as in the baseline model. Given ρ, player L then selects the experiment that maximizes her expected payoff, among those inducing a stochastic choice rule σ whose maximal slope is no greater than ρ. For any ρ and r, the optimal experiment can be taken to be binary and, for any ω, it recommends $z = 1$ (i.e., engagement) with probability $q^{\rho,r}(1|\omega)$ given by

$$q^{\rho,r}(1|\omega) = \begin{cases} 1 & \text{if} \quad \omega \le m^*(r) - \frac{1}{2\rho} \\ \frac{1}{2} - \rho(\omega - m^*(r)) & \text{if} \quad m^*(r) - \frac{1}{2\rho} < \omega \le m^*(r) + \frac{1}{2\rho} \\ 0 & \text{if} \quad \omega > m^*(r) + \frac{1}{2\rho} \end{cases}$$

where $m^*(r)$ is the same engagement cutoff as in the previous sections and is such that $\delta_L(r, m^*(r)) = 0$.

What distinguishes the two examples of flexible information acquisition above from the analysis in the previous sections is that, for any cognitive choice ρ, there are multiple experiments that share the same cost (parametrized by ρ). After choosing ρ, player L then chooses the experiment that maximizes her expected payoff among those whose cost is $C(\rho)$, with the optimal choice depending on the anticipated reaction r by player F.

It is evident that, in each of the two cases of flexible information acquisition described above, under the optimal experiment $q^{\rho,r}$, when player L receives the signal $z = 1$ (equivalently, when

[35] The formula below is for when $1/\rho$ measures the marginal cost of entropy reduction. Conclusions similar to those reported below hold for the case where ρ determines the information capacity, i.e., the maximal level of entropy reduction, as in Sims (2003)'s original work on rational inattention (see also Mackowiak and Wiederholt (2009)).

she engages), her posterior mean, which is given by

$$\mathbb{E}[\omega|z=1;q^{\rho,r}] = \int \omega \frac{q^{\rho,r}(1|\omega)}{q^{\rho,r}(1)} dG(\omega),$$

with $q^{\rho,r}(1|\omega) \equiv \int_\omega q^{\rho,r}(1|\omega)dG(\omega)$, is less than $m^*(r)$, and likewise, after receiving signal $z=0$,

$$\mathbb{E}[\omega|z=0;q^{\rho,r}] = \int \omega \frac{1-q^{\rho,r}(1|\omega)}{1-q^{\rho,r}(1)} dG(\omega)$$

is greater than $m^*(r)$.

For any anticipated cognition $\rho^\dagger$, any reaction r by player F, and any cutoff m^*, then let $M^-(m^*;\rho^\dagger,r)$ denote the expected value of m conditional on $m \leq m^*$, when player L chooses cognition $\rho^\dagger$ and then selects the optimal experiment $q^{\rho^\dagger,r}$ anticipating a reaction r by player F. Then note that, for any $\rho^\dagger$ and r, when the cutoff is equal to $m^*(r)$,

$$M^-(m^*(r);\rho^\dagger,r) = \mathbb{E}[\omega|z=1;q^{\rho^\dagger,r}]$$

and $\partial M^-(m^*(r);\, \rho^\dagger,r)/\partial\rho^\dagger \overset{\text{sgn}}{=} A(m^*(r);\rho^\dagger,r)$, with

$$A(m^*(r);\rho^\dagger,r) \equiv \left[m^*(r) - M^-(m^*(r);\, \rho^\dagger,r)\right] G_\rho(m^*(r);\rho^\dagger,r) - \int_{-\infty}^{m^*(r)} G_\rho(m;\rho^\dagger,r)dm$$

where, for any m, $G(m;\rho^\dagger,r)$ denotes the probability that L's posterior mean is less than m under the experiment $q^{\rho^\dagger,r}$ and where $G_\rho(m;\rho^\dagger,r)$ denotes the partial derivative of such a probability with respect to ρ, evaluated at $\rho=\rho^\dagger$; such a derivative is computed accounting for the fact that, when ρ changes, the optimal experiment $q^{\rho,r}$ also changes.

As in the baseline model, the sign of A determines whether an increase in cognition aggravates or alleviates the adverse selection problem. Consistently with the baseline model, we then continue to interpret $\mathrm{A}(\rho^\dagger) \equiv A(m^*(r(\rho^\dagger));\rho^\dagger,r(\rho^\dagger))$ as the "adverse selection effect". As in the baseline model, $r(\rho^\dagger)$ denotes the equilibrium reaction by player F in a fictitious setting in which player L's cognition is exogenously fixed at $\rho^\dagger$. However, differently from the baseline model, in this fictitious setting, player L chooses the distribution $G(\cdot;\rho^\dagger,q)$ over her posterior mean m by selecting an experiment $q:\Omega\to\Delta(Z)$.[36] The equilibrium reaction $r(\rho^\dagger)$ is thus computed jointly with the equilibrium choice of experiment q and the equilibrium engagement strategy $a(\cdot)$.

[36] Recall that $\rho^\dagger$ only pins down the marginal cost of entropy reduction (alternatively, the maximal level of entropy reduction) or the maximal slope of the induced stochastic choice rule, leaving player L with flexibility over her choice of experiment $q:\Omega\to\Delta(Z)$.

Next, let

$$\Pi(\rho; r) \equiv U_L(0) + \int_{-\infty}^{m^*(r)} \delta_L(r, m) dG(m; \rho, r) = U_L(0) + \int_{-\infty}^{+\infty} \delta_L(r, \omega) q^{\rho,r}(1|\omega) dG(\omega)$$

denote the payoff that player L obtains under cognition ρ when expecting a reaction r by player F (with the expectation computed under the optimal experiment $q^{\rho,r}$) and then let

$$\begin{aligned} B(\rho; \rho^\dagger) &\equiv -\frac{\partial^2 \Pi(\rho; r(\rho^\dagger))}{\partial \rho \partial r} = -\int_{-\infty}^{m^*(r(\rho^\dagger))} \frac{\partial \delta_L}{\partial r}(r(\rho^\dagger), m) dG_\rho(m; \rho, r(\rho^\dagger)) \\ &= -\frac{\partial \delta_L(r(\rho^\dagger), m^*(r(\rho^\dagger)))}{\partial r} G_\rho\Big(m^*(r(\rho^\dagger)); \rho, r(\rho^\dagger)\Big) \\ &\quad + \int_{-\infty}^{m^*(r(\rho^\dagger))} \frac{\partial^2 \delta_L(r(\rho^\dagger), m)}{\partial r \partial m} G_\rho(m; \rho, r(\rho^\dagger)) dm. \end{aligned}$$

As in the baseline model, the function $B(\rho; \rho^\dagger)$ measures how a reduction in the friendliness of F's reaction around $r(\rho^\dagger)$ affects L's marginal benefit of cognition when the latter is equal to ρ. Consistently with the baseline model, we will continue to refer to $B(\rho; \rho^\dagger)$ as the "benefit of friendlier reactions" effect.

Finally, as in the previous sections, let $V_L(\rho; \rho^\dagger) \equiv \Pi(\rho; r(\rho^\dagger))$ denote the value to player L of choosing cognition ρ when player F expects cognition $\rho^\dagger$ and then responds with reaction $r = r(\rho^\dagger)$.

The following proposition establishes results analogous to those in Proposition 1 but for flexible information acquisition, when the cost of cognition is determined by either entropy reduction or the maximum slope of the induced stochastic choice rule.

Proposition 8 (***EC under flexible information acquisition***)**.** *Suppose that assumptions 1 and 2 hold and that cognition ρ determines either the marginal cost of entropy reduction or the maximum-slope of the induced stochastic choice rule.*
(i) Expectation conformity holds at $(\rho, \rho^\dagger)$ if and only if the adverse selection and the benefit of friendlier reactions effects are of opposite sign: $\mathrm{A}(\rho^\dagger)\mathrm{B}(\rho; \rho^\dagger) < 0$.
(ii) Suppose that L's cognition is given by ρ and that F's reaction is given by $r(\rho^\dagger)$. A sufficient condition for cognition to aggravate adverse selection is that $q^{\rho,r(\rho^\dagger)}(1|\omega)/q^{\rho,r(\rho^\dagger)}(1)$ is increasing in ρ for $\omega < m^*(r(\rho^\dagger))$ and decreasing in ρ for $\omega > m^*(r(\rho^\dagger))$.
(iii) Suppose that L's cognition is given by ρ and that F's reaction is given by $r(\rho^\dagger)$. A sufficient condition for a reduction in the friendliness of player F's reaction to raise player L's value of cognition is that, in addition to the property in part (ii), $q^{\rho,r(\rho^\dagger)}(1) \equiv \int q^{\rho,r(\rho^\dagger)}(1|\omega) dG(\omega)$ is non-increasing in ρ.
(iv) Therefore, a sufficient condition for expectation conformity to hold at $(\rho, \rho^\dagger)$ is that the conditions in parts (ii) and (iii) jointly hold.

(v) Suppose that $M^{-}(m^*(r(\rho^\dagger)); \rho)$ is decreasing in ρ at $\rho = \rho^\dagger$, implying that $\mathrm{A}(\rho^\dagger) < 0$, and that $\partial^2 \delta_L(r, m)/\partial r \partial m = 0$ (as is examples (a), (b), (d), and (e) in Subsection 2.2). Then $q^{\rho, r(\rho^\dagger)}(1)$ non-increasing in ρ is necessary and sufficient for expectation conformity to hold at $(\rho, \rho^\dagger)$.

Proof. See the Appendix.

The results in the proposition highlight a few commonalities with the corresponding results in Proposition 1 in the main text, but also a few differences. As in the baseline model, expectation conformity obtains when the effect of cognition on the severity of the adverse selection problem is of opposite sign than the effect of a reduction in the friendliness of F's reaction on the marginal value of cognition, i.e., when $\mathrm{A}(\rho^\dagger)B(\rho; \rho^\dagger) < 0$. The intuition for this result is the same as in the baseline model.

An increase in cognition aggravates the severity of the adverse selection problem when it induces L to select an experiment that makes her engage with a higher probability at low states (namely for $\omega < m^*(r(\rho^\dagger)))$ and with a lower probability at high states (namely for $\omega > m^*(r(\rho^\dagger)))$, relative to the total probability $q^{\rho, r(\rho^\dagger)}(1)$ of engaging. This is because such changes make the engagement decision by player L a more informative signal of the state being less profitable to player F.

When, in addition to this property, cognition also reduces the overall probability $q^{\rho, r(\rho^\dagger)}(1)$ that player L engages, a reduction in the friendliness of player F's reaction increases L's marginal value of cognition. The property that $q^{\rho, r(\rho^\dagger)}(1)$ decreases with ρ is equivalent to the property that $G_\rho\Big(m^*\big(r(\rho^\dagger)\big); \rho\Big) < 0$ in the baseline model. As discussed in Section 3, this condition is both necessary and sufficient for expectation conformity when L's payoffs is separable in r and ω and cognition always aggravates the adverse selection problem.

8 Conclusions

We investigate how incentives to acquire information in generalized lemons problems depend on other players' expectations about the acquired information. We show how expectation conformity, i.e., the value to conform to other players' expectations, is affected by (a) the impact of information on the severity of the adverse selection problem, (b) the sensitivity of the marginal value of information to the friendliness of other players' reactions, and (c) the overall value of engagement, as captured by the size of the gains from trade.

We then use the characterization to shed light on the connection between expectation conformity and cognitive traps, and on the role of disclosure of hard information in such games, whereby players engage in activities that prove how well or poorly informed they are.

Finally, we show how the results change in the anti-lemons case, that is, in settings in which the players' payoffs are aligned, and how a benevolent planner can improve upon the laissez-faire equilibrium by subsidizing (alternatively, taxing) trade.

There are many venues for future research. First, it would be interesting to investigate how the type of security issued to finance a project affects the incentives for information acquisition and the resulting severity of the adverse selection problem. Second, it would be interesting to study how public disclosures impact the incentives for private information acquisition. For example, in the context of stress testing, the announcement that a bank failed to pass a test may induce a conservative response by potential asset buyers which may induce asset owners to collect even more information, which in turn aggravates the severity of the adverse selection problem. This is an angle that does not seem to have been accounted for in the design of the optimal stress tests. Lastly, it would be interesting to extend the analysis by allowing both sides of the market to acquire information and investigate how strategic complementarity/substitutability in information acquisition is affected by the adverse selection problem.

References

Akerlof, G. A., (1970), "The Market for Lemons: Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, Vol. 84: 488-500.

An, M.Y. (1998) "Logconcavity versus Logconvexity: A Complete Characterization," *Journal of Economic Theory*, 80(2): 350–369.

Bar-Isaac, H., I. Jewitt, and C. Leaver (2018), "Adverse Selection, Efficiency and the Structure of Information," mimeo Oxford University.

Bergemann, D., B. Brooks, and S. Morris (2015): "The Limits of Price Discrimination," *American Economic Review*, 105, 921–57.

Bulow, J., Geanakoplos, J. and P. Klemperer (1985) "Multimarket Oligopoly: Strategic Substitutes and Complements," *Journal of Political Economy*, 93(3): 488–511.

Colombo, L., Femminis, G., and A. Pavan (2022), "Subsidies to Technology Adoption when Firms' Information is Endogenous" mimeo Universita' Cattolica of Milan, and Northwestern University.

Cramton, P., Gibbons R., and P. Klemperer (1987), "Dissolving a Partnership Efficiently," *Econometrica*, 615-632.

Cremer, J., and F. Khalil (1992), "Gathering Information Before Signing a Contract," *American Economic Review*, 82(3): 566–578.

Cremer, J., F. Khalil, and J.C. Rochet (1998), "Contracts and Productive Information Gathering," *Games and Economic Behavior*, 25(2): 174–193.

Dang, T.V., (2008), "Bargaining with endogenous information," *Journal of Economic Theory* 140, 339 – 354.

Dang, T.V., Gorton, G., Holmström B. and G. Ordonez (2017) "Banks As Secret Keepers," *American Economic Review*, 107(4): 1005–1029.

Diamond, D. W., and P. H. Dybvig (1983) "Bank runs, Deposit Insurance, and Liquidity," *Journal of Political Economy* 91.3: 401-419.

Diamond, P. and J. Stiglitz (1974) "Increases in Risk and Risk Aversion," *Journal of Economic Theory*, 8: 337–360.

Dye, R. (1985), “Disclosure of Non-proprietary Information,” *Journal of Accounting Research*, 23(1): 123-145.

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.

Hermalin, B. (1998) “Toward an Economic Theory of Leadership: Leading by Example,” *American Economic Review*, 88: 1188–1206.

Johnston, J., and D. Myatt (2006) “On the Simple Economics of Advertising, Marketing, and Product Design,” *American Economic Review*, 96(3): 756–784.

Kartik, N., and W. Zhong (2019) “Lemonade from Lemons: Information Design and Adverse Selection,” mimeo Columbia University.

Kessler, A. S. (2001), “Revisiting the Lemons Market,” *International Economic Review*, 42, 25–41.

Levin, J. (2001), “Information and the Market for Lemons,” *RAND Journal of Economics*, 32, 657–666.

Mackowiak, B. and M. Wiederholt (2009) “Optimal Sticky Prices under Rational Inattention,” *American Economic Review*, 99(3): 769-803.

Milgrom, P. (2008) “What the Seller Won’t Tell You: Persuasion and Disclosure in Markets,” *Journal of Economic Perspectives*, 22: 115–131.

Morris, S. and M. Yang, (2022) ”Coordination and Continuous Stochastic Choice,” *Review of Economic Studies* 89(5) (2022): 2687-2722.

Netzer, N. (2009), “Evolution of Time Preferences and Attitudes toward Risk,” *American Economic Review* 99(3): 937-955.

Pavan A., S, Sundaresan, and X. Vives (2022) “(In)efficiency in Information Acquisition and Aggregation through Prices,” mimeo, IESE, Imperial College, and Northwestern University.

Pavan A. and J. Tirole (2022a) “Exposure to the Unexpected, Duty of Disclosure, and Contract Design,” mimeo Northwestern University and Toulouse School of Economics.

Pavan A. and J. Tirole (2022b) “Expectation Conformity in Strategic Cognition,” mimeo Northwestern University and Toulouse School of Economics.

Philippon, T., and V. Skreta (2012) “Optimal Interventions in Markets with Adverse Selection,” *American Economic Review*, 102(1): 1–28.

Rayo, L. and G. Becker (2007), “Evolutionary Efficiency and Happiness,” *Journal of Political Economy* 115(2): 302-337.

Robson, A. J. (2001), “The Biological Basis of Economic Behavior,” *Journal of Economic Literature* 39(1): 11-33.

Ravid, D., Roesler, A.K. and B. Szentes (2022) “Learning before Trading: On The Inefficiency of Ignoring Free Information,” *Journal of Political Economy,* 130(2), 346–387.

Roesler, A.K., and B. Szentes (2017) “Buyer-Optimal Learning and Monopoly Pricing,” *American Economic Review*, 107(7): 2072–2080.

Shishkin, D. (2020) “Evidence Acquisition and Voluntary Disclosure,” mimeo, Princeton University.

Sims, C. (2003) “Implications of Rational Inattention,” *Journal of Monetary Economics* 50 (2003): 665–690.

Spier, K (1992) “Incomplete Contracts and Signaling,” *Rand Journal of Economics*, 23: 432–443.

Thereze, J. (2022), "Adverse Selection and Information Acquisition," mimeo, Princeton University.

Tirole, J. (2012) "Overcoming Adverse Selection: How Public Intervention Can Restore Market Functioning," *American Economic Review*, 102(1): 29–59.

Yang, M. (2015) "Coordination with Flexible Information Acquisition," *Journal of Economic Theory* 158: 721-738.

Woodford, M. (2009) "Information-constrained State-Dependent Pricing," *Journal of Monetary Economics*, 56: 100-124.

9 Appendix

Proof of Proposition 6. Using the envelope theorem, we have that[37]

$$W'(s) = \int_{-\infty}^{m^*(r^*(s),s)} \left[\frac{\partial \delta_L(r^*(s),m)}{\partial r}\frac{dr^*(s)}{ds} + 1\right] dG(m;\rho^*(s))$$

$$-\frac{d}{ds}\left[(1+\lambda)sG(m^*(r^*(s),s);\rho^*(s))\right].$$

The first line is simply the effect of a change in the subsidy on the leader's expected payoff (holding $\rho^*(s)$ and $m^*(r^*(s),s)$ fixed by usual envelope-theorem arguments). The second line is the (total) effect of a change in the subsidy on the cost of the program to the government.

Note that $W'(s)$ can be expressed as

$$W'(s) = \frac{dr^*(s)}{ds}\int_{-\infty}^{m^*(r^*(s),s)} \frac{\partial \delta_L(r^*(s),m)}{\partial r} dG(m;\rho^*(s)) - s(1+\lambda)\frac{d}{ds}\left[G(m^*(r^*(s),s);\rho^*(s))\right]$$

$$-\lambda G(m^*(r^*(s),s);\rho^*(s)).$$

When $W(s)$ is quasi-concave in s, the optimal s is thus strictly positive when

$$W'(0) = \frac{dr^*(0)}{ds}\int_{-\infty}^{m^*(r^*(0),0)} \frac{\partial \delta_L(r^*(0),m)}{\partial r} dG(m;\rho^*(0)) - \lambda G(m^*(r^*(0),0);\rho^*(0)) > 0$$

and strictly negative when the above inequality is reversed.

Because δ_L is affine in m, it can be expressed as $\delta_L(r,m) = a_L(r)m + b_L(r)$, for some functions $a_L(r)$ and $b_L(r)$. Assumption 1 then implies that, for any r and m, $a_L(r) < 0$ and $a'_L(r)m + b'_L(r) > 0$. This means that $W'(0) > 0$ when $dr^*(0)/ds > r^{\#}$, whereas $W'(0) < 0$ when $dr^*(0)/ds < r^{\#}$, where

$$r^{\#} \equiv \frac{\lambda}{\frac{\partial}{\partial r}\delta_L(r^*(0), M^{-}(m^*(r^*(0),0);\rho^*(0))}$$

is a strictly positive constant that depends on the primitives of the problem. For example, in the Akerlof model, $\delta_L(r,m) = -m + r$, in which case $r^{\#} = \lambda$.

[37] Here we are using the fact that, given s and $r^*(s)$, $m^*(r^*(s),s)$ and $\rho^*(s)$ maximize the leader's payoff $\int_{-\infty}^{\hat{m}} (\delta_L(r^*(s),m) + s)\, dG(m;\rho) - C(\rho)$ over $(\hat{m},\rho)$.

Next, observe that, for any s, $\rho^*(s)$ and $r^*(s)$ jointly solve the following two conditions:

$$\int_{-\infty}^{m^*(r^*,s)} \delta_F(r^*, m) dG(m; \rho^*) = 0, \tag{19}$$

and

$$\rho^* = \arg\max_{\rho} \left\{ \int_{-\infty}^{m^*(r^*,s)} \left(\delta_L(r^*, m) + s\right) dG(m; \rho) - C(\rho) \right\}. \tag{20}$$

Because δ_F is affine in m, it can be expressed as $\delta_F(r, m) = a_F(r)m + b_F(r)$, for some functions $a_F(r)$ and $b_F(r)$, with $a_F(r) > 0$ when Assumption 2 holds (lemons), and $a_F(r) < 0$ when Assumption 2' holds (anti-lemons).[38]

Hence, for any s, $r^*(s)$ solves $\delta_F(r^*, M^-(m^*(r^*, s); \rho^*(s))) = 0$. Using the implicit-function theorem, we have that

$$\frac{dr^*(s)}{ds} = -\frac{\frac{d}{ds}\delta_F(r, M^-(m^*(r,s); \rho^*(s)))\big|_{r=r^*(s)}}{\frac{d}{dr}\delta_F(r, M^-(m^*(r,s); \rho^*(s)))\big|_{r=r^*(s)}}$$

where the denominator is negative, by assumption. We thus have that $dr^*(0)/ds > r^{\#}$ if

$$\frac{d}{ds}\delta_F(r^*(0), M^-(m^*(r^*(0), s); \rho^*(s)))\bigg|_{s=0} > \Lambda \equiv \lambda \left| \frac{\frac{d}{dr}\delta_F(r, M^-(m^*(r, 0); \rho^*(0)))\big|_{r=r^*(0)}}{\frac{d}{dr}\delta_L(r, M^-(m^*(r, 0); \rho^*(0)))\big|_{r=r^*(0)}} \right|$$

whereas $dr^*(0)/ds < r^{\#}$ if the above inequality is reversed. The condition says that, at the laissez-faire equilibrium, holding the follower's reaction fixed at $r^*(0)$, a small subsidy to engagement has a strong enough positive effect on the follower's payoff, accounting for the effect that the subsidy has on both the leader's engagement threshold and cognition. Note that, in the Akerlof model of Example 1, $\Lambda = \lambda$.

Clearly,

$$\frac{d}{ds}\delta_F(r^*(0), M^-(m^*(r^*(0), s); \rho^*(s)))\bigg|_{s=0} = a_F(r^*(0))\, \frac{d}{ds}M^-(m^*(r^*(0), s); \rho^*(s))\bigg|_{s=0}.$$

The result in the proposition then follows by letting $K \equiv \Lambda/a_F(r^*(0))$ and noting that $K > 0$ when $a_F(r^*(0)) > 0$, i.e., when Assumption 2 (lemons) holds, whereas $K < 0$ when $a_F(r^*(0)) < 0$, i.e., when Assumption 2' (anti-lemons) holds. ■

Proof of Proposition 7. The optimal value of s^* solves $dW(s^*)/ds = 0$. That is, s^* solves

[38] To see this, note that, for any m^*, ρ, and r, $\int_{-\infty}^{m^*} \delta_F(r, m) dG(m; \rho) = G(m^*; \rho)\delta_F(r, M^-(m^*; \rho))$. Now fix s and drop it. The equilibrium r thus solves $a_F(r)M^-(m^*(r); \rho) + b_F(r) = 0$. Hence,

$$\frac{dr}{d\rho} = -\frac{a_F(r)\frac{\partial}{\partial\rho}M^-(m^*(r); \rho)}{\frac{\partial}{\partial r}\delta_F(r, M^-(m^*(r); \rho))}.$$

The denominator in the above expression is negative, by assumption. It follows that $dr/d\rho \overset{sgn}{=} a_F(r)\frac{\partial}{\partial\rho}M^-(m^*(r); \rho)$. Hence, $a_F(r) > 0$ when Assumption 2 holds, whereas $a_F(r) < 0$ when Assumption 2' holds.

$$\frac{dr^*(s^*)}{ds}\int_{-\infty}^{m^*(r^*(s^*),s^*)}\frac{\partial\delta_L(r^*(s^*),m)}{\partial r}dG(m;\rho^*(s^*))$$

$$-(1+\lambda)s^*g(m^*(r^*(s^*),s^*);\rho^*(s^*))\left[\frac{\partial m^*(r^*(s^*),s^*)}{\partial r}\frac{dr^*(s^*)}{ds}+\frac{\partial m^*(r^*(s^*),s^*)}{\partial s}\right] \tag{21}$$

$$-(1+\lambda)s^*\frac{d\rho^*(s^*)}{ds}G_\rho(m^*(r^*(s^*),s^*);\rho^*(s^*))-\lambda G(m^*(r^*(s^*),s^*);\rho^*(s^*))=0.$$

Next, observe that, when information is exogenous and equal to $\rho=\rho^*(s^*)$, using the envelope theorem and the fact that $\hat{r}(s^*)=r^*(s^*)$, we have that

$$\frac{dW^{\#}(s^*)}{ds}=\frac{\partial\hat{W}(r^*(s^*),s^*)}{\partial r}\frac{d\hat{r}(s^*)}{ds}+\frac{\partial\hat{W}(r^*(s^*),s^*)}{\partial s}$$

where

$$\frac{\partial\hat{W}(r^*(s^*),s^*)}{\partial r}=\int_{-\infty}^{m^*(r^*(s^*),s^*)}\frac{\partial\delta_L(r^*(s^*),m)}{\partial r}dG(m;\rho^*(s^*))$$

$$-(1+\lambda)s^*g(m^*(r^*(s^*),s^*);\rho^*(s^*))\frac{\partial m^*(r^*(s^*),s^*)}{\partial r}$$

and

$$\frac{\partial\hat{W}(r^*(s^*),s^*)}{\partial s}=-(1+\lambda)s^*g(m^*(r^*(s^*),s^*);\rho^*(s^*))\frac{\partial m^*(r^*(s^*),s^*)}{\partial s}-\lambda G(m^*(r^*(s^*),s^*);\rho^*(s^*)).$$

Using (21), we thus have that

$$\frac{dW^{\#}(s^*)}{ds}=\left(\frac{d\hat{r}(s^*)}{ds}-\frac{dr^*(s^*)}{ds}\right)\frac{\partial\hat{W}(r^*(s^*),s^*)}{\partial r}+(1+\lambda)s^*\frac{d\rho^*(s^*)}{ds}G_\rho(m^*(r^*(s^*),s^*);\rho^*(s^*))$$

Assuming that $W^{\#}(s)$ is quasi-concave in s, we then have that $s^{**}<s^*$ if $dW^{\#}(s^*)/ds<0$ and $s^{**}>s^*$ if the above inequality is reversed, which leads to the result in the proposition. ■

Proof of Corollary 2. The result follows from Proposition 7. In fact, in this case,

$$\frac{d\hat{r}(s^*)}{ds}-\frac{dr(s^*)}{ds}=\frac{\frac{\partial}{\partial\rho}M^-(m^*(r^*(s^*),s^*);\rho^*(s^*))\frac{d\rho^*(s^*)}{ds}}{\frac{\partial}{\partial m^*}M^-(m^*(r^*(s),s^*);\rho^*(s^*))-1}.$$

Using the fact that

$$\frac{\partial}{\partial\rho}M^-(m^*;\rho)=\frac{G_\rho(m^*;\rho)[m^*-M^-(m^*;\rho)]-\int_{-\infty}^{m^*}G_\rho(m;\rho)dm}{G(m^*;\rho)}$$

and

$$\frac{\partial}{\partial m^*}M^-(m^*;\rho)=\frac{g(m^*;\rho)[m^*-M^-(m^*;\rho)]}{G(m^*;\rho)},$$

along with the fact that $m^*(r^*(s^*),s^*)=r^*(s^*)+s^*$ and $r^*(s^*)=M^-(m^*(r^*(s^*),s^*);\rho^*(s^*))+\Delta$, we then have that

$$\frac{d\hat{r}(s^*)}{ds} - \frac{dr(s^*)}{ds} = \left((s^*+\Delta)G_\rho(m^*(r^*(s^*),s^*);\rho^*(s^*)) - \int_{-\infty}^{m^*(r^*(s^*),s^*)} G_\rho(m;\rho^*(s^*))dm\right)\frac{\frac{d\rho^*(s^*)}{ds}}{D},$$

where

$$D \equiv (s^*+\Delta)g(m^*(r^*(s^*),s^*);\rho^*(s^*)) - G(m^*(r^*(s^*),s^*);\rho^*(s^*)) < 0$$

when $G(m;\rho^*(s^*))/g(m;\rho^*(s^*))$ is increasing in m. Hence, $\frac{d\hat{r}(s^*)}{ds} - \frac{dr(s^*)}{ds} < 0$ when

$$G(m;\rho^*(s^*))/g(m;\rho^*(s^*))$$

is increasing in m, information structures are consistent with the MPS order, $G_\rho(m^*(r^*(s^*),s^*);\rho^*(s^*)) < 0$, and $d\rho^*(s^*)/ds < 0$.

Furthermore, in this case

$$\frac{\partial \hat{W}(r^*(s^*),s^*)}{\partial r} = G(m^*(r^*(s^*),s^*);\rho^*(s^*)) - (1+\lambda)s^* g(m^*(r^*(s^*),s^*);\rho^*(s^*)).$$

Using the fact that

$$\frac{dW^{\#}(s^*)}{ds} = \left(\frac{d\hat{r}(s^*)}{ds} - \frac{dr^*(s^*)}{ds}\right)\frac{\partial \hat{W}(r^*(s^*),s^*)}{\partial r} + (1+\lambda)s^*\frac{d\rho^*(s^*)}{ds}G_\rho(m^*(r^*(s^*),s^*);\rho^*(s^*))$$

as established in the proof of Proposition 7, we then have that $dW^{\#}(s^*)/ds = \frac{d\rho(s^*)}{ds}\frac{J}{D}$, where

$$J \equiv (\Delta - \lambda s^*)\, G_\rho(m^*(r^*(s^*),s^*);\rho^*(s^*))G(m^*(r^*(s^*),s^*);\rho^*(s^*))$$

$$+\left(\int_{-\infty}^{m^*(r^*(s^*),s^*)} G_\rho(m;\rho^*(s^*))dm\right)\left[(1+\lambda)s^* g(m^*(r^*(s^*),s^*);\rho^*(s^*)) - G(m^*(r^*(s^*),s^*);\rho^*(s^*))\right].$$

Note that $J < 0$ when information structures are consistent with the MPS order, $\frac{G(m;\rho^*(s^*))}{g(m;\rho^*(s^*))}$ is increasing in m, and $G_\rho(m^*(r^*(s^*),s^*);\rho^*(s^*)) < 0$.[39]

We conclude that, under the assumptions in the corollary, $dW^{\#}(s^*)/ds \overset{sgn}{=} d\rho(s^*)/ds$. To see that, under the assumption in the corollary, $d\rho(s^*)/ds < 0$, note that

$$\frac{dr^*(s)}{ds} = -\frac{\frac{\partial M^-(m^*(r^*(s),s);\rho^*(s))}{dm^*} + \frac{\partial M^-(m^*(r^*(s),s);\rho^*(s))}{d\rho}\frac{d\rho^*(s)}{ds}}{\frac{\partial M^-(m^*(r^*(s),s);\rho^*(s))}{dm^*} - 1}$$

[39]Note that, under the optimal subsidy s^*, welfare is equal to $G(m^*(r^*(s^*),s^*);\rho^*(s^*))\,(\Delta - \lambda s^*) - C(\rho^*(s^*))$. Because welfare is non-negative under the laissez-faire equilibrium (i.e., when $s=0$), it must be that $\Delta > \lambda s^*$. Also note that, when $\Delta > \lambda s^*$, $(1+\lambda)s^* g(m^*(r^*(s^*),s^*);\rho^*(s^*)) - G(m^*(r^*(s^*),s^*);\rho^*(s^*)) < D$ and hence the second line in J is negative when information structures are consistent with the MPS order and $G(m;\rho^*(s^*))/g(m;\rho^*(s^*))$ is increasing in m.

Under the assumptions in the Corollary,

$$\frac{\partial M^-(m^*(r^*(s),s);\rho^*(s))}{dm^*}-1=DG(m^*(r^*(s^*),s^*);\rho^*(s^*))<0$$

and $\partial M^-(m^*(r^*(s),s);\rho^*(s))/d\rho<0$. Hence, $dr^*(s)/ds<0$ if $d\rho(s^*)/ds>0$. This cannot be consistent with the optimality of s^*. In fact, by cutting the subsidy, the planner would then induce a friendlier reaction by the follower, permit the leader to economize on cognition, and save on the costs of public funds. The optimality of s^* thus implies that $d\rho(s^*)/ds<0$. We conclude that $dW^{\#}(s^*)/ds<0$. The strict quasi-concavity of $W^{\#}$ then implies that $s^{**}<s^*$. ■

Proof of Proposition 8. (i) The proof follows from the same arguments that establish part (i) of Proposition 1.

(ii) Recall that

$$\frac{\partial}{\partial\rho}\left(M^-(m^*\left(r(\rho^\dagger)\right);\rho,r(\rho^\dagger))\right)=\frac{\partial}{\partial\rho}\left(\int\omega\frac{q^{\rho,r(\rho^\dagger)}(1|\omega)}{q^{\rho,r(\rho^\dagger)}(1)}dG(\omega)\right).$$

Both when the cost of information is given by entropy reduction and when it is given by maximum slope, $q^{\rho,r(\rho^\dagger)}(1|\omega)/q^{\rho,r(\rho^\dagger)}(1)$ is a decreasing function of ω. Hence, when $q^{\rho,r(\rho^\dagger)}(1|\omega)/q^{\rho,r(\rho^\dagger)}(1)$ is increasing in ρ for $\omega<m^*(r(\rho^\dagger))$ and decreasing in ρ for $\omega>m^*(r(\rho^\dagger))$, the collection of distributions $\left(F^{\rho,r(\rho^\dagger)}\right)_\rho$, indexed by ρ, with each cdf $F^{\rho,r(\rho^\dagger)}$ defined by the density

$$f^{\rho,r(\rho^\dagger)}(\omega)\equiv\frac{q^{\rho,r(\rho^\dagger)}(1|\omega)}{q^{\rho,r(\rho^\dagger)}(1)}g(\omega)$$

can be ranked according to FOSD, with $F^{\rho,r(\rho^\dagger)}\succ F^{\rho',r(\rho^\dagger)}$ for any $\rho<\rho'$. This means that $M^-(m^*(r(\rho^\dagger));\rho,r(\rho^\dagger))$ is decreasing in ρ, which implies that cognition aggravates adverse selection.

(iii) Note that

$$-\frac{\partial\Pi(\rho;r(\rho^\dagger))}{\partial r}=-q^{\rho,r(\rho^\dagger)}(1)\int\frac{\partial\delta_L(r(\rho^\dagger),\omega)}{\partial r}\frac{q^{\rho,r(\rho^\dagger)}(1|\omega)}{q^{\rho,r(\rho^\dagger)}(1)}dG(\omega).\tag{22}$$

Under Assumption 1, $\partial\delta_L(r(\rho^\dagger),\omega)/\partial r$ is increasing in ω. Hence, when $q^{\rho,r(\rho^\dagger)}(1|\omega)/q^{\rho,r(\rho^\dagger)}(1)$ is increasing in ρ for $\omega<m^*(r(\rho^\dagger))$ and decreasing in ρ for $\omega>m^*(r(\rho^\dagger))$, the integral term in (22) is decreasing in ρ (the arguments are the same as in part (ii)). Hence, given $r=r(\rho^\dagger)$, a sufficient condition for a reduction in the friendliness of player F's reaction to raise the incentive for cognition at ρ is that, in addition to the condition in part (ii) of the proposition, $q^{\rho,r(\rho^\dagger)}(1)$ is non-increasing in ρ.

(iv) The proof is an immediate implication of parts (ii) and (iii).

(v) The proof follows from fact that, in this case,

$$B(\rho;\rho^\dagger)=-\frac{\partial\delta_L}{\partial r}\Big(r(\rho^\dagger),m^*(r(\rho^\dagger))\Big)G_\rho\Big(m^*\left(r(\rho^\dagger)\right);\rho^\dagger,r(\rho^\dagger)\Big)$$

as shown in the proof of Proposition 1. Because $\mathrm{A}(\rho^{\dagger}) < 0$, the result in part (i) implies that a necessary and sufficient condition for expectation conformity to hold at $(\rho, \rho^{\dagger})$ is that $B(\rho; \rho^{\dagger}) > 0$ which is the case if and only if $G_{\rho}\Big(m^{*}\big(r(\rho^{\dagger})\big); \rho, r(\rho^{\dagger})\Big) < 0$. Because, for any ρ,

$$G\Big(m^{*}\big(r(\rho^{\dagger})\big); \rho, r(\rho^{\dagger})\Big) = \int q^{\rho, r(\rho^{\dagger})}(1|\omega) dG(\omega) \equiv q^{\rho, r(\rho^{\dagger})}(1)$$

the latter property is equivalent to $q^{\rho, r(\rho^{\dagger})}(1)$ being non-increasing in ρ. ■

	Lemons $\frac{dr}{d\rho^\dagger} \overset{sgn}{=} \frac{\partial}{\partial \rho^\dagger} M^-(m^*;\rho^\dagger)$	Anti-lemons $\frac{\mathrm{dr}}{\mathrm{d}\rho^\dagger} \overset{\mathrm{sgn}}{=} -\frac{\partial}{\partial \rho^\dagger} \mathrm{M}^-(\mathrm{m}^*;\rho^\dagger)$
Does an increase in cognition lead to unfriendlier response? ($dr/d\rho^\dagger < 0$)	Yes if • MPS and $G_\rho < 0$ • or cognition always aggravates adverse selection (e.g., Uniform, Pareto, Exp.)	No if • MPS and $G_\rho < 0$ • or cognition always aggravates adverse selection (e.g., Uniform, Pareto, Exp.)
Does an unfriendlier response increase L's demand for cognition? $\left(-\frac{\partial^2 \Pi(\rho;r)}{\partial r \partial \rho} > 0 \right)$	• Yes if MPS and $G_\rho < 0$ • No if $G\rho \geq 0$ and $\frac{\partial^2 \delta}{\partial m \partial r} = 0$ (ex. a, b, d, e)	• Yes if MPS and $G_\rho < 0$ • No if $G\rho \geq 0$ and $\frac{\partial^2 \delta}{\partial m \partial r} = 0$ (ex. g, h, j)
(Local) expectation conformity / cognitive traps	• Yes if MPS and $G_\rho < 0$ • $G_\rho < 0$ is NSC if cognition always aggravates adverse selection and $\frac{\partial^2 \delta}{\partial m \partial r} = 0$	• Yes if cognition always aggravates adverse selection, $G_\rho > 0$, and $\frac{\partial^2 \delta}{\partial m \partial r} = 0$
Engagement channel of subsidy $\frac{\partial}{\partial m^*} M^-(m^*;\rho^\dagger) \frac{\partial m^*}{\partial s}$	• Benefits player L	• Hurts player L
Cognition channel of subsidy $\frac{\partial}{\partial \rho^\dagger} M^-(m^*;\rho^\dagger) \frac{d\rho^\dagger}{ds}$	• Benefits player L if MPS and $G_\rho < 0$	• Hurts player L if MPS and $G_\rho < 0$
Total effect of subsidy on welfare	• positive if engagement + cognition channels $> K > 0$	• negative if engagement + cognition channels $> K < 0$

Table 1: summary of a few results