We develop a model of persuasion where the persuadee has to exert effort – the elaboration cost – in order to fully and precisely elaborate information. The persuader makes an offer to the persuadee and, aware that she is a dual process reasoner, also sends her a costly signal – the reference cue – which refers the offer to a category of offers whose average quality is known by the persuadee. Initially, the actual quality of the offer by the persuader is hidden to the persuadee, while the signal is visible. Then, the persuadee can either rely on cheap low elaboration and form expectations on the basis of the signal – thinking coarsely, i.e., by category – or engage in costly high elaboration to attain knowledge of the actual quality of the offer. This signaling setup allows to maintain the assumption that agents are both rational and Bayesian and, at the same time, to match many of the findings emphasized by well established psychological models of persuasion. In addition, the model provides novel theoretical results such as the possibility of separating equilibria that do not rely on the single-crossing property and, in particular, the emergence of a new phenomenon that we name reverse-signaling, where high types send low signals and low types send high signals.

JEL classification code: D01, D03, D82, D83.

Keywords: persuasion, coarse reasoning, peripheral and central route, heuristic and systematic reasoning, reverse-signaling, counter-signaling.
1 Introduction

Consider the following situation. A milk producer can bottle milk in either glass or plastic. Glass is more expensive than plastic, and has no impact on milk quality per se. A consumer has to decide whether to buy a bottle of milk, and can choose whether to exert high or low effort in evaluating the quality of the milk offered. If the consumer exerts low effort, then she only relies on the fact that milk is contained in a glass bottle, and she generically thinks of products in glass containers, whose average quality can be higher than that of products in plastic containers. If the consumer exerts high effort, then she reads and understands labels, recovers data from memory, and elaborates all the information extracted; this is a costly activity at the end of which she is able to fully assess the quality of the product. If the milk producer anticipates the behavior of the consumer, then he can use the container – which we call *cue* (or signal) throughout the rest of the paper – to persuade the consumer to buy his product.

In this paper we study strategic situations as the one described above, in the attempt to contribute to the general understanding of persuasion activities. In particular we show that, if we nest on a standard economic model the insights from psychology about dual process reasoning (high and low elaboration effort) and categorical thinking (expected quality is average quality in a category), then we can rationalize the behavior of a principal who – in order to persuade an agent to act in his interest – resorts to costly communicative tools – such as the glass bottle – that are otherwise useless in a model where the agent has unbounded cognitive resources. More precisely, we frame persuasion activities within a sender-receiver model – the sender is the *persuader*, the receiver is the *persuadee* – assuming that agents are both rational and Bayesian (see Subsection 7.1 for the related economics literature). In addition, we introduce cognitive limitations along the lines suggested by social and cognitive psychology with regard to how the persuadee can elaborate information and how the persuader can take advantage of this (see Subsection 7.2 for the related psychological literature). In particular, we follow Dewatripont and Tirole (2005) by assuming that the persuadee has to pay a cognitive cost to process information fully and precisely.\(^1\)\(^2\) However, we depart from Dewatripont and Tirole (2005) in the modeling of cues – i.e., a simple signal sent by the persuader which does not require great cognitive effort to be processed – which

---

1 Other recent contributions considering the costly acquisition of information are Dewatripont (2006), Caillaud and Tirole (2007), Tirole (2009) and Butler et al. (2013).

2 Brocas and Carrillo (2008) and Brocas (2012) stress that the evidence provided by brain sciences on the multi-system nature of the human brain should be a fundamental source of inspiration for the economic modeling of decision-making (see Brocas and Carrillo, 2014, for a focused survey).
they assume to contain information on the trustworthiness of the persuader; instead, we propose to model cues as signals that refer an object to a category of objects, and to follow Mullainathan (2002) in assuming that individuals may suffer from “coarse thinking”, i.e., they might be unable to distinguish objects falling in the same category. More precisely, we assume that the persuadee suffers from coarse thinking whenever she chooses not to bear the cognitive cost of elaborating carefully the message sent by the persuader. Thanks to this we can investigate the strategic use of reference cues by the persuader and we are able to match many findings of psychological theories of persuasion. In particular, this assumption allows us to model persuasion activities such as the one considered in the field experiment by Bertrand et al. (2010) – where prospective borrowers are offered loans by mailed advertisement cards with different cues – which can hardly fit the model proposed by Dewatripont and Tirole (2005) since the latter only considers cues related to the sender’s expertise.

We briefly sketch the basic working of our model. The persuader makes an offer of unknown quality to the receiver, and provides a reference cue – like the packaging of a product – that gives summary information about the expected quality of the offer; the persuadee can then decide to exert a low cognitive effort in processing the offer – low elaboration – and take a decision on the sole basis of the reference cue, or she can scrutinize the offer carefully – high elaboration – and obtain a precise knowledge of the offer received but also bear a cognitive cost due to information processing. Importantly, we do not treat low elaboration as a behavioral shortcut, but as ignorance that the persuadee is aware of and able to quantify, so that she can form expectations and choose the elaboration level to maximize expected utility. We stress that this is where coarse thinking crucially comes in: instead of investigating all benefits and costs of the current offer, the persuadee uses the reference cue to assign the offer to a category of similar offers – all associated with the same reference cue — and then evaluates the offer according to the average quality of its category, taking a decision on the sole basis of this information and her priors. The trade-off at play is the following: low elaboration leads to take a decision that is good on average, but that is not necessarily the


4The marketing literature focuses on the strategic use of cues to influence how consumers perceive quality, and points out that very effective cues might have little to do with actual quality of products (Teas and Agarwal, 2000). Experimental and survey evidence indicates that these cues can exert a sizeable impact on consumers’ valuation (see, e.g., Sáenz-Navajas et al., 2013; Woodside, 2012, for wine and food products, respectively). Our model can be seen as a way to rationalize the use of such cues in the attempt to persuade consumers to buy (also, see Bagwell, 2007, for detailed references on persuasive advertisement).
best choice for the current offer (only high elaboration guarantees to always take the best choice available); however, low elaboration allows to save on cognitive costs, and hence the persuadee can well decide not to engage in high elaboration, especially when beliefs about the quality of the current offer are quite extreme (in which case the persuadee does not expect to learn much from high elaboration). Since the persuader anticipates all this, he can make a strategic use of the reference cue.

Despite its simplicity, this model and its extensions prove able to predict behavior that is in line with well established findings on persuasion in the psychological literature (Petty and Cacioppo, 1986a; Eagly and Chaiken, 1993). The persuasive message sent by the persuader is interpreted by the persuadee differently under high elaboration and low elaboration (Remark 1). A greater motivation or a better cognitive ability by the persuadee makes the recourse to high elaboration more likely (Remarks 2 and 3). The use of reference cues by the persuader affects the elaboration effort and the reaction to the offer by the persuadee (Remark 4). Persuasion obtained under high elaboration by the persuadee is stabler than if obtained under low elaboration (Remark 5). Antecedents of elaboration such as arousal or prior knowledge can influence the intensity of elaboration and induce biased elaboration (Remarks 6 and 7).

Our analysis also shows that explicitly considering elaboration costs in a signaling framework where signals are reference cues allows for the emergence of a variety of possible equilibrium behaviors, some of which cannot be easily accommodated in a standard signaling model. First, there are pooling equilibria where all types of persuaders employ the same reference cue, either high or low. Second, there are separating equilibria where different types send different reference cues. As one might expect, there is a separating equilibrium where high quality persuaders employ high reference cues while low quality persuaders employ low reference cues – this outcome resembling standard separating equilibria of signaling models. We stress, however, that separating equilibria can exist in our model even if different types incur the same costs for reference cues and obtain the same gains from having their offer accepted, i.e., the single-crossing property does not hold. Standard separation of types is still possible because if the persuadee decides to engage in high elaboration then she can condition acceptance on types, and in such a case different types face different expected gains. We also stress that separating equilibria can exist only for elaboration costs that are mild relatively to the stake associated with the offer. Indeed, if elaboration costs are sufficiently low, then the persuadee relies on high elaboration whatever cue she observes, and hence both persuader types find it optimal to save money and pool on the cheap reference cue; if, instead, elaboration costs are sufficiently high, then the persuadee never relies on high elaboration, and similarly there is no reason for either persuader type to choose the more
costly cue, with the result that both types pool on the cheap reference cue also in this case.

Perhaps more interestingly, a different form of separation between types can emerge in this setup: separating equilibria where high quality persuaders go with low reference cues and low quality persuaders go with high reference cues. We name this phenomenon reverse-signaling, which has to be distinguished from the related phenomenon of counter-signaling. Feltovich et al. (2002) define counter-signaling as a situation where a sender has a quality that can be mistaken only for close qualities, and this allows the emergence of a signaling outcome where medium-quality senders choose high signals to separate from low-quality senders, while high-quality senders choose low signals to separate from medium-quality senders, thus yielding an inverted U-shaped relationship between types and signals (for this requiring at least three types).

Reverse-signaling, instead, leads to a negative monotonic relationship between types and signals, and is triggered by a totally different mechanism. A low quality persuader relies on a high cue that refers to a category where average quality is high enough, so that the persuadee chooses to accept the offer without exerting high elaboration effort. A high quality persuader prefers to save on the cost of the signal, choosing a cheaper cue that refers to a category where average quality is intermediate, so that the persuadee exerts high elaboration effort and accepts the offer after discovering it is of high quality. The low quality persuader does not find it profitable to choose the cheaper cue, because his offer would be discovered to be of low quality and hence rejected. We note that this result cannot arise in the model by Dewatripont and Tirole (2005) as in their setup the sender is not allowed to manipulate the cue, but only to either send a truthful one or do not send it at all. We also stress that our result hinges on coarse thinking by the persuadee, which makes the manipulation of the cue potentially worthwhile (Mullainathan et al., 2008). In this sense, the reverse-signaling arising in our model can be seen as a case of profitable deception similar to that studied in Heidhues et al. (2014), where naive consumers overestimate the net value of products that they buy; however, differently from Heidhues et al. (2014), in our model agents are not necessarily naive (or sophisticated), but can decide to be so in order to save on the cognitive effort, sustaining deception in equilibrium.

The paper is organized as follows. Section 2 presents the model in three steps: the modeling of the elaboration of information (Subsection 2.1), the optimal behavior of the persuadee in terms of elaboration level and reaction to the offer (Subsection 2.2), and the optimal behavior of the persuader who anticipates the behavior of the persuadee and tries to take advantage of this (Subsection 2.3). Section 3 presents all persuasion equilibria of our model,
distinguishing between pooling equilibria (Subsection 3.1), standard separating equilibria (Subsection 3.2), and separating equilibria where reverse-signaling emerges (Subsection 3.3); uniqueness and existence of equilibria are also discussed and related to the degree of coarse thinking (Subsection 3.4). Section 4 provides a discussion of the potential interpretations of the model, with various examples. Section 5 discusses how our model relates to psychological findings. Section 6 extends our stylized model in various directions, with the aim of checking the robustness of results: many offer qualities (Subsection 6.1), many offer categories and reference cues (Subsection 6.2), a continuum of elaboration intensities (Subsection 6.3), and cues with fully endogenous quality (Subsection 6.4). Section 7 surveys the relevant literature on persuasion, distinguishing between contributions from economics (Subsection 7.1) and those from psychology (7.2), and some recent papers that in different ways model the rational allocation of scarce cognitive resources (Subsection 7.3). Finally, Section 8 concludes, summarizing the contribution and showing lines for future research.

2 The model

We introduce the model in three steps. Firstly, we describe the message received by the persuadee (to whom we refer as “she”) and how she can elaborate the information it contains. Secondly, we study her behavior with respect to both the choice of the elaboration level and the reaction to the offer. Finally, we introduce the persuader (to whom we refer as “he”) and we analyze his strategic choice regarding the reference cue.

2.1 Message processing: High and low elaboration

We follow the psychological literature in assuming that the decision-maker (DM) can process the message at two different levels of elaboration. More precisely, DM can decide to exert a high cognitive effort and scrutinize the message carefully. This can be interpreted as the use of the “central route” in the ELM or the use of the “systematic elaboration” in the HSM. Alternatively, DM can decide to exert a low cognitive effort and adopt simple heuristics, categorizations, and rules of thumb. This can be interpreted as the use of the “peripheral route” in the ELM or the use of the “heuristic elaboration” in the HSM. We stress that DM chooses the elaboration level and is fully aware of the elaboration cost.

More formally, message processing is modeled as follows. DM faces a two-part message \((q, r) \in \{G, B\} \times \{x, y\}\) associated with an offer which she has to decide upon. Part \(r \in \{x, y\}\) of the message is a reference cue, i.e., a piece of information which allows DM to refer the
offer to a specific category of offers and, through this, to infer the expected quality of the offer she is facing. Part \( q \in \{G, B\} \) of the message contains the information regarding the actual quality of the offer: if \( q = G \) then quality is good, while if \( q = B \) then quality is bad.

Whenever DM is aware of a message \((q, r)\), she chooses a level of cognitive effort \( e \in \{H, L\} \) to elaborate it. In particular, DM can exert a low effort of elaboration \( L \) and only acquire knowledge of part \( r \) of the message, or exert a high effort of elaboration \( H \) and acquire also part \( q \) of the message. The elaboration level \( L \) is assumed to be costless and to occur as soon as the message is received, so that \( r \) is always observed. Instead, the elaboration level \( H \) can be activated after observing part \( r \) and requires to bear a cost of elaboration, denoted with \( c_e > 0 \).

We emphasize that part \( r \) of the message is informative in the sense that it refers to a category of objects of which DM knows the average quality. This is a simple way to model the fact that, when choosing \( e = L \), DM is affected by “coarse thinking”, i.e., DM puts different offers with a common characteristic in the same mental category and treats them all in the same way (Fryer and Jackson, 2008; Mullainathan et al., 2008).

Figure 1 is a graphical representation of the different information on quality that can be drawn from the same message depending on the elaboration level chosen by DM.

### 2.2 Persuadee behavior: Elaboration level and reaction

We now study the costs and benefits for DM with regard to her choices about the elaboration level and the reaction to the offer. To help focus attention on the behavior of the persuadee, for now we abstract from the decision problem of the persuader – in a sense, we solve the model backwards.

DM initially processes the message received under \( e = L \), so she observes the reference cue \( r \) while the actual quality \( q \) of the offer remains hidden. At this stage DM has a precise belief on the offer quality, that is conditional of the reference cue observed. Given such a belief, DM evaluates the convenience of increasing effort to \( e = H \), paying \( c_e \) and acquiring the knowledge of \( q \). Finally, DM has to decide whether to accept the offer, which we denote with \( Y \), or reject it, which we denote with \( N \).

The payoffs for DM are the following. If DM accepts the offer when \( q = G \), then she obtains \( U_G > 0 \). If instead she accepts the offer when \( q = B \), then she obtains \( U_B < 0 \). If DM rejects the offer, then she obtains a null payoff independently of \( q \). In any case, if DM exerts \( e = H \), then she also has to bear the elaboration cost \( c_e \).

In most cases we will stick to the interpretation that \( Y \) means acceptance and \( N \) means rejection. However we stress that other interpretations are possible. For instance, if the offer
is a consumer good to be bought, then $Y$ could mean to buy a large quantity of the good and $N$ to buy a small quantity; also, $Y$ could mean to pay a high price per unit while $N$ to pay a low price per unit. Both cases can easily be accommodated by the model since what really matters for the results is that $Y$ is preferred when quality is $G$ and $N$ is preferred when quality is $B$. We observe that $U_G$ and $U_B$ can be interpreted as relative convenience to choose $Y$ over $N$ when, respectively, quality is $G$ and quality is $B$.

Given the structure of DM’s payoffs, we immediately recognize that, if DM exerts $e = H$, then she finds it profitable to choose $Y$ in case $q = G$ and $N$ in case $q = B$. We indicate such behavior with $HYN$. Other behaviors are possible, that make use of high elaboration but do not react optimally to the knowledge of actual quality. To simplify the analysis, we neglect those behaviors. We denote behaviors that make use of low elaboration with $LY$ and $LN$, respectively leading to acceptance and rejection. Therefore, a strategy for DM can be described by a function $\delta : \{x, y\} \rightarrow \{LY, LN, HYN\}$.
We now proceed to compare the expected payoffs that DM obtains by choosing HYN, LY, and LN. We denote with $\mu$ the belief about the offer quality, conditional on the observed reference cue. More precisely, $\mu \in [0, 1]$ is the probability that the offer is of quality $G$, after a reference cue has been observed. At this stage of the analysis, we treat $\mu$ as a parameter. We observe that:

- the choice of LY leads to an expected utility of $\mu U_G - (1 - \mu)|U_B|$;
- the choice of HYN leads to an expected utility of $\mu U_G - c_e$;
- the choice of LN leads to an expected utility of 0.

By solving, with respect to $\mu$, for the highest expected payoff or utility across the above ones, we directly obtain the following proposition, whose proof is hence omitted.

**PROPOSITION 1** (Persuadee’s optimal behavior).

Let $\mu$ be DM’s belief that $q = G$. The following holds for DM:

- HYN is optimal if and only if: $\mu \geq \frac{c_e}{U_G}$ and $\mu \leq 1 - \frac{c_e}{|U_B|}$;
- LY is optimal if and only if: $\mu \geq 1 - \frac{c_e}{|U_B|}$ and $\mu \geq \frac{|U_B|}{U_G + |U_B|}$;
- LN is optimal if and only if: $\mu \leq \frac{c_e}{U_G}$ and $\mu \leq \frac{|U_B|}{U_G + |U_B|}$.

Note that HYN turns out to be optimal for some intermediate range of values of $\mu$ only if $c_e \leq \frac{U_G|U_B|}{U_G + |U_B|}$, although when equality holds there is just one value of $\mu$ for which HYN is optimal and for that value DM is indifferent between HYN, LN, and LY. The optimal behavior by DM as a function of $\mu$ is summarized in Figure 2.
Intuitively, DM is less likely to choose $H$ if the elaboration cost $c_e$ is higher, and never uses it if $c_e$ is too high. Moreover, the expected benefits for DM to choose $H$ over $L$ decrease when beliefs on quality are more extreme – i.e., a high $\mu$ that gets closer to 1 or a low $\mu$ that gets closer to 0 – because less and less uncertainty remains while the same elaboration cost has to be borne.

2.3 Persuader behavior: Strategic use of reference cues

We now introduce the persuader (P) and focus on the strategic use of reference cues. In particular, we study how P can induce in DM favorable beliefs about his offer through the sending of a reference cue $r$. This behavior on the part of P can be interpreted as an attempt to “frame” the offer (see, e.g., Mullainathan et al., 2008): the observation of $r$ induces DM to associate P’s offer to a specific mental category containing all offers sharing characteristic $r$, and consequently to evaluate P’s offer by considering the average quality of offers in such a category.

Formally, in the first stage of the game Nature selects one among the following three possibilities: with probability $\alpha_x > 0$ the offer is not made by P and belongs to the category associated with reference cue $x$, with probability $\alpha_y > 0$ the offer is not made by P and belongs to the category associated with reference cue $y$, and with probability $\alpha_P = 1 - \alpha_x - \alpha_y > 0$ the offer is made by P. We stress that probabilities $\alpha_x$, $\alpha_y$, and $\alpha_P$ are intimately related to coarse thinking. Indeed, such probabilities can be interpreted as relative frequencies of occurrence and, hence, they are naturally thought of as dependent on the relative size of categories associated with $x$ and $y$, which in turn is affected by the degree of coarse thinking. We elaborate more on this in Subsection 3.4.

If the offer does not come from P and belongs to category $x$, then it is of quality $q = G$ with probability $\beta_x$. Similarly, if the offer does not come from P and belongs to category $y$, then it is of quality $q = G$ with probability $\beta_y$. Hence, the parameters $\beta_x$ and $\beta_y$ represent the fraction of good quality offers in, respectively, category $x$ and $y$, not taking into account the behavior of P. Without loss of generality we assume that $\beta_x > \beta_y$, i.e., on average $x$ refers to a higher quality than $y$. An attempt to endogenize $\beta_x$ and $\beta_y$ is provided in Subsection 6.4.

Instead, if the offer comes from P, then the game unfolds as follows: the quality of the offer – i.e., P’s type – is $q = G$ with probability $\alpha_G$, and $q = B$ with probability $\alpha_B = 1 - \alpha_G$, and then P chooses a reference cue $r \in \{x, y\}$ to be associated with the offer. Probabilities $\alpha_G$ and $\alpha_B$ can be interpreted as, respectively, the fraction of good quality offers and the fraction of bad quality offers when P is called into play. We remark that the quality of an
offer should be interpreted in a broad sense, as something that correlates positively with DM’s utility. The cost for P of choosing reference cue \( r \) is \( c_r \). Since \( \beta_x > \beta_y \), to rule out uninteresting cases we assume that \( c_x > c_y = 0 \), i.e., we posit that referring the offer to a category of higher average quality is more costly for P, and we normalize \( c_y \) to zero. Quality \( q \) is known to P, so a strategy for P is a function \( \rho : \{ G, B \} \rightarrow \{ x, y \} \) indicating which reference cue is chosen conditionally on the quality of the offer selected by Nature. The payoff for P is \( V > c_x \) in case DM accepts the offer while it is 0 if DM rejects the offer. In any case, the cost \( c_r \) must be borne. We stress that the null payoff obtained by P when DM chooses \( N \) should be seen as a normalization, so to be consistent with different interpretations of \( Y \) and \( N \); for instance, if \( N \) means to buy at a low price or a small quantity, then P would earn a low but possibly positive profit.

Under low elaboration DM suffers from coarse thinking, and hence she is unable to distinguish where the offer comes from and its quality, while only the reference cue is directly visible. After observing \( r \), DM updates her belief \( \mu \) (that \( q = G \)) and then decides whether to switch from \( e = L \) to \( e = H \) and acquire knowledge of \( q \) at the cost \( c_e \). Finally, DM either accepts or rejects the offer, i.e., plays \( Y \) or \( N \), potentially conditioning this choice upon the information acquired on \( q \). A summary of the decision structure of this game is provided in Figure 3.

Given DM’s behavior, the choices that maximize P’s payoff are easily established as P obtains \( V \) only if DM reacts with \( Y \), while P incurs the cost of reference \( c_r \) independently of DM’s choices. Note that both types of P want to have the offer accepted by DM, but for type \( G \) it is enough that DM plays \( HYN \), while type \( B \) needs that DM plays \( LY \); moreover, everything else being equal, both types prefer to send \( y \), since it costs less than \( x \). The following proposition summarizes P’s best replies to the optimal choices by DM, as analyzed in Proposition 1. The proof can be found in Appendix B.

**Proposition 2** (Persuader’s optimal behavior).

The following specifications of strategy \( \rho \) are optimal for P:

- \( \rho(G) = \rho(B) = y \), if DM’s strategy is such that \( \delta(x) = \delta(y) \);
- \( \rho(G) = \rho(B) = x \), if DM’s strategy is such that \( \delta(x) = LY \) and \( \delta(y) = LN \);
- \( \rho(G) = x \) and \( \rho(B) = y \), if DM’s strategy is such that \( \delta(x) = HYN \) and \( \delta(y) = LN \);
- \( \rho(G) = y \) and \( \rho(B) = x \), if DM’s strategy is such that \( \delta(x) = LY \) and \( \delta(y) = HYN \).
Figure 3: The game tree. In the left and the central branch DM faces an offer that does not come from P. In the left branch the offer is characterized by cue $x$ and has probability $\beta_x$ to be of quality $G$. In the central branch the offer is characterized by cue $y$ and has probability $\beta_y$ to be of quality $G$. Since in these parts of the game P is not involved, his payoff is set equal to 0. In the right branch DM faces an offer made by P, which has probability $\alpha_G$ to be of quality $G$. DM has two information sets each of which encompasses nodes from all three branches; one set is associated with reference cue $x$ and the other set with reference cue $y$. 
Proposition 2 relates the optimal strategy of P to the behaviors that can be optimally adopted by DM, but it is silent on whether DM finds it actually optimal to adopt them. The analysis of mutually optimal decision between P and DM is carried out in the next section. Here we just want to introduce some further notation to take into account that average quality in a category depends on P’s choice. We denote with $\hat{\beta}_r(\rho(G),\rho(B))$ the average quality of offers in the category associated with the reference cue $r$ when P’s choices, which are described by function $\rho$, are taken into account. The different values that $\hat{\beta}_r(\rho(G),\rho(B))$ can take as a function of $\rho$ are summarized in Appendix B.

3 Persuasion equilibria

The arguments provided in the previous section almost led us to the identification of equilibria. In the present section we complete the analysis characterizing three types of equilibria and discussing both existence and uniqueness. We restrict our analysis to pure strategies.

Preliminarily, we note that as long as $\alpha_x, \alpha_y, \beta_x, \beta_y$ are strictly comprised between 0 and 1, any combination of quality and reference cue occurs with positive probability. This implies that every information set of the game is reached with positive probability under any strategy profile. Therefore, every Nash equilibrium – to which we simply refer as equilibrium – is also sequential, and, hence, weak perfect Bayesian. We also note that the absence of out-of-equilibrium beliefs makes it impossible to refine the set of equilibria by applying criteria that rule out some scarcely plausible out-of-equilibrium beliefs (such as, e.g., the Intuitive Criterion or D1).

3.1 Pooling equilibria: High and low signals

A first type of equilibria is characterized by the persuader P sending a message with a reference cue that is independent of the actual quality of the offer, i.e., $\rho(G) = \rho(B) \in \{x, y\}$.

We say that an equilibrium is pooling with high signal if $\rho(G) = \rho(B) = x$. We note that the rejection of any offer associated with cue $y$, i.e., $\delta(y) = LN$, and the acceptance of any offer associated with cue $x$, i.e., $\delta(x) = LY$, is the only behavior by DM that can sustain a pooling equilibrium with high signal. Indeed, in such a case P can have his offer accepted only by using the reference cue $x$, and this independently of whether he is of type $G$ or of type $B$. We also note that a pooling equilibrium with high signal can exist for low elaboration costs – so that $HYN$ is optimal for intermediate values of expected quality – and for high elaboration costs – so high that $HYN$ is never optimal. Figure 4 illustrates one
such equilibrium when costs of elaboration are low. The following proposition provides the conditions for the existence of a pooling equilibrium with high signal.

\[
\begin{array}{ccc}
\text{LN} & \text{HYN} & \text{LY} \\
0 & \hat{\beta}_y(\rho(G)=x, \rho(B)=x) & \mu & \hat{\beta}_x(\rho(G)=x, \rho(B)=x) & 1
\end{array}
\]

Figure 4: A pooling equilibrium where both \(G\) and \(B\) send cue \(x\), and DM finds it optimal to reply with \(LY\) to cue \(x\) and with \(LN\) to cue \(y\). Elaborations costs are low enough (i.e., \(c_e < \frac{|U_G|}{U_G + |U_B|}\)) to make \(HYN\) a best reply for intermediate levels of expected quality.

**Proposition 3** (Pooling equilibrium with high signal). The profile \((\rho, \delta)\) such that \(\rho(G) = \rho(B) = x, \delta(x) = LY, \delta(y) = LN\) is an equilibrium if and only if:

1. \(\hat{\beta}_x(\rho(G), \rho(B)) \geq 1 - \frac{c_e}{|U_B|}\) and \(\hat{\beta}_y(\rho(G), \rho(B)) \geq \frac{|U_B|}{U_G + |U_B|}\);
2. \(\hat{\beta}_y(\rho(G), \rho(B)) \leq \frac{c_e}{U_G}\) and \(\hat{\beta}_y(\rho(G), \rho(B)) \leq \frac{|U_B|}{U_G + |U_B|}\).

There is no other equilibrium profile \((\rho, \delta)\) such that \(\rho(G) = \rho(B) = x\), whatever the values of \(\hat{\beta}_x(\rho(G), \rho(B))\) and \(\hat{\beta}_y(\rho(G), \rho(B))\).

**Proof.** The last claim of the proposition follows directly from Proposition 2: for \(r = x\) to be P’s optimal choice independently of his type, DM’s choice must be such that \(\delta(x) = LY\) and \(\delta(y) = LN\).

So, let \((\rho, \delta)\) be an equilibrium. Note that, along the equilibrium path, \(\mu = \hat{\beta}_x(\rho(G), \rho(B))\) if DM sees \(r = x\) and \(\mu = \hat{\beta}_y(\rho(G), \rho(B))\) if DM sees \(r = y\). Hence, from Proposition 1 follows that 3.1 must hold for DM to find \(\delta(x) = LY\) optimal and that 3.2 must hold for DM to find \(\delta(y) = LN\) optimal.

Suppose now that 3.1 and 3.2 hold. Then, from Proposition 1 follows that \(\delta(x) = LY\) and \(\delta(y) = LN\) is optimal for DM. Hence, by Proposition 2 we can conclude that the profile \((\rho, \delta)\) is an equilibrium.
We say that an equilibrium is pooling with low signal if $\rho(G) = \rho(B) = y$. The existence of such an equilibrium depends on the convenience for DM to behave in the same way when observing $x$ or $y$, i.e., $\delta(x) = \delta(y)$. Indeed, when this occurs, P clearly finds it optimal to choose $y$ irrespectively of whether he is of type $G$ or of type $B$, since DM’s behavior is not affected by the choice of the reference cue, and $y$ costs less than $x$. We note that there are variants of this type of equilibrium, depending on the behavior held by DM. If elaboration costs are so large that $HYN$ is never optimal then there are two cases: either $\delta(x) = \delta(y) = LN$ or $\delta(x) = \delta(y) = LY$. Otherwise, if elaboration costs are not so large, then there is also a third possibility, namely that $\delta(x) = \delta(y) = HYN$. Figure 5 illustrates the case where DM always opts for $HYN$. The following proposition provides the conditions for the existence of a pooling equilibrium with low signal for each of the three variants.

**Proposition 4** (Pooling equilibrium with low signal).

4.1 The profile $(\rho, \delta)$ such that $\rho(G) = \rho(B) = y$, $\delta(x) = \delta(y) = LN$ is an equilibrium if and only if:

\[ \hat{\beta}_x(\rho(G), \rho(B)) \leq \frac{c_\rho}{U_y} \quad \text{and} \quad \hat{\beta}_x(\rho(G), \rho(B)) \leq \frac{|U_p|}{U_y + |U_B|}; \]

4.2 The profile $(\rho, \delta)$ such that $\rho(G) = \rho(B) = y$, $\delta(x) = \delta(y) = HYN$ is an equilibrium if and only if:

\[ \hat{\beta}_x(\rho(G), \rho(B)) \geq \frac{c_\rho}{U_y} \quad \text{and} \quad \hat{\beta}_x(\rho(G), \rho(B)) \leq 1 - \frac{c_\rho}{|U_B|}; \]

\[ \hat{\beta}_y(\rho(G), \rho(B)) \geq \frac{c_\rho}{U_y} \quad \text{and} \quad \hat{\beta}_y(\rho(G), \rho(B)) \leq 1 - \frac{c_\rho}{|U_B|}. \]

4.3 The profile $(\rho, \delta)$ such that $\rho(G) = \rho(B) = y$, $\delta(x) = \delta(y) = LY$ is an equilibrium if and only if:

\[ \hat{\beta}_x(\rho(G), \rho(B)) \geq 1 - \frac{c_\rho}{|U_B|} \quad \text{and} \quad \hat{\beta}_x(\rho(G), \rho(B)) \geq \frac{|U_p|}{U_y + |U_B|}; \]

\[ \hat{\beta}_y(\rho(G), \rho(B)) \geq 1 - \frac{c_\rho}{|U_B|} \quad \text{and} \quad \hat{\beta}_y(\rho(G), \rho(B)) \geq \frac{|U_p|}{U_y + |U_B|}. \]

There is no other equilibrium profile $(\rho, \delta)$ such that $\rho(G) = \rho(B) = y$, whatever the values of $\hat{\beta}_x(\rho(G), \rho(B))$ and $\hat{\beta}_y(\rho(G), \rho(B))$.

**Proof.** The last claim of the proposition follows directly from Proposition 2: for $r = y$ to be P’s optimal choice independently of his type, DM’s choice must be such that $\delta(x) = \delta(y)$. 

15
Suppose \((\rho, \delta)\) of 4.1 be an equilibrium. Note that, along the equilibrium path, \(\mu = \hat{\beta}_x(\rho(G), \rho(B))\) if DM sees \(r = x\) and \(\mu = \hat{\beta}_y(\rho(G), \rho(B))\) if DM sees \(r = y\). Hence, from Proposition 1 follows that 4.1.1 and 4.1.2 must hold for to find \(\delta(x) = \delta(y) = LN\) optimal.

Suppose now that 4.1.1 and 4.1.2. Then, from Proposition 1 follows that \(\delta(x) = \delta(y) = LN\) is optimal for DM. Hence, by Proposition 2 we can conclude that the profile \((\rho, \delta)\) is an equilibrium.

Claims 4.2 and 4.3 can be proved with analogous arguments.

Figure 5: A pooling equilibrium where both \(G\) and \(B\) send cue \(y\), and DM finds it optimal to reply with \(HYN\) to both cue \(x\) and \(y\). Elaborations costs are low enough (i.e., \(c_e < \frac{U_G}{U_G + U_B}\)) to make \(HYN\) a best reply for intermediate levels of expected quality.

### 3.2 Separating equilibrium: High quality going with high signal

A second group of equilibria is characterized by the type of separation that is typical of signaling models where there are both high quality and low quality senders: persuaders whose offer is of quality \(G\) send the reference cue \(x\) associated with high quality, while persuaders whose offer is of quality \(B\) send the reference cue \(y\) associated with low quality. More precisely, we say that an equilibrium is separating with high quality going with high signal – or that it is a separating equilibrium with signaling – if \(\rho(G) = x\) and \(\rho(B) = y\). We note that this behavior by \(P\) can form an equilibrium only if DM chooses \(\delta(x) = HYN\) and \(\delta(y) = LN\). In such a case, the persuader of type \(G\) finds it optimal to incur the cost of sending \(x\), since this leads his offer to be accepted, while the persuader of type \(B\) prefers to save on costs and send reference cue \(y\), since in no case his offer will be accepted. We also note that for this equilibrium to exist elaboration costs must be low enough so that DM
actually best replies with \( HYN \) for intermediate values of \( \mu \). Figure 6 illustrates an example of such a separating equilibrium, while the following proposition provides the conditions for its existence.

**Proposition 5** (Separating equilibrium with signaling).

The profile \((\rho, \delta)\) such that \( \rho(G) = x, \rho(B) = y, \delta(x) = HYN, \delta(y) = LN \) is an equilibrium if and only if:

1. \( \hat{\beta}_x(\rho(G), \rho(B)) \geq \frac{c_e}{U_G} \) and \( \hat{\beta}_x(\rho(G), \rho(B)) \leq 1 - \frac{e}{|U_B|} \);
2. \( \hat{\beta}_y(\rho(G), \rho(B)) \leq \frac{c_e}{U_G} \) and \( \hat{\beta}_y(\rho(G), \rho(B)) \leq \frac{|U_B|}{U_G + |U_B|} \).

**Proof.** Let \((\rho, \delta)\) be an equilibrium. Note that, along the equilibrium path, \( \mu = \hat{\beta}_x(\rho(G), \rho(B)) \) if DM sees \( r = x \) and \( \mu = \hat{\beta}_y(\rho(G), \rho(B)) \) if DM sees \( r = y \). Hence, from Proposition 1 follows that 5.1 must hold for DM to find \( \delta(x) = HYN \) optimal and that 5.2 must hold for DM to find \( \delta(y) = LN \) optimal.

Suppose now that 5.1 and 5.2 hold. Then, from Proposition 1 follows that \( \delta(x) = HYN \) and \( \delta(y) = LN \) is an optimal choice for DM. Hence, by Proposition 2 we can conclude that the profile \((\rho, \delta)\) is an equilibrium. \(\square\)

3.3 Separating equilibrium: High quality going with low signal

The last group of equilibria is perhaps the most interesting, since it is characterized by a sort of reverse-signaling behavior: separation between high quality and low quality senders.
is attained because the low quality persuader $B$ sends the reference cue $x$ associated with high quality, while the high quality persuader $G$ sends the reference cue $y$ associated with low quality. More precisely, we say that an equilibrium is separating with high quality going with low signal – or that it is a separating equilibrium with reverse-signaling – if $\rho(G) = y$ and $\rho(B) = x$. We note that this behavior by P can form an equilibrium only if DM chooses $\delta(y) = HYN$ and $\delta(x) = LY$. Indeed, in such a case the persuader of type $G$ finds it optimal to save on costs and send $y$, since this leads nevertheless his offer to be accepted, while the persuader of type $B$ finds it optimal to pay the cost of sending $x$, since this is the only way to have his offer accepted. We also note that, as for the other separating equilibrium, in order for this equilibrium to exist elaboration costs must be low enough so that DM actually best replies with $HYN$ for intermediate values of $\mu$. Figure 7 illustrates an example of a reverse-signaling equilibrium, while the following proposition provides the conditions for its existence.

**Proposition 6** (Separating equilibrium with reverse-signaling).

The profile $(\rho, \delta)$ such that $\rho(G) = y$, $\rho(B) = x$, $\delta(x) = LY$, $\delta(y) = HYN$ is an equilibrium if and only if:

$6.1 \; \hat{\beta}_x(\rho(G), \rho(B)) \geq 1 - \frac{\epsilon_x}{|U_B|}$ and $\hat{\beta}_x(\rho(G), \rho(B)) \geq \frac{|U_B|}{|U_G| + |U_B|};$

$6.2 \; \hat{\beta}_y(\rho(G), \rho(B)) \geq \frac{\epsilon_y}{|U_G|}$ and $\hat{\beta}_y(\rho(G), \rho(B)) \leq 1 - \frac{\epsilon_y}{|U_B|}.$

**Proof.** Let $(\rho, \delta)$ be an equilibrium. Note that, along the equilibrium path, $\mu = \hat{\beta}_x(\rho(G), \rho(B))$ if DM sees $r = x$ and $\mu = \hat{\beta}_y(\rho(G), \rho(B))$ if DM sees $r = y$. Hence, from Proposition 1 follows that $6.1$ must hold for DM to find $\delta(x) = LY$ optimal and that $6.2$ must hold for DM to find $\delta(y) = HYN$ optimal.

Suppose now that $6.1$ and $6.2$ hold. Then, from Proposition 1 follows that $\delta(x) = LY$ and $\delta(y) = HYN$ is an optimal choice for DM. Hence, by Proposition 2 we can conclude that the profile $(\rho, \delta)$ is an equilibrium.

### 3.4 Uniqueness and existence of equilibria

We observe that some of the different types of equilibria described so far can actually coexist, at least for some set of parameter values. Besides potential multiplicity there are also cases where no equilibrium actually exists. Examples and some discussion of these occurrences are provided in Appendix A, where we also give a proposition summarizing all possibile cases of coexistence of equilibria.
Equilibrium inexistence and equilibrium multiplicity share a common root: the endogeneity of expected quality $\hat{\beta}_x$ and $\hat{\beta}_y$, and in particular their dependence on the behavior of P. This feature crucially hinges on the fact that beliefs $\hat{\beta}_x$ and $\hat{\beta}_y$ are endogenous, since DM is a fully Bayesian updater and hence takes into account P’s behavior when forming her conditional expectations on quality along a path of play. However, the degree of beliefs’ endogeneity is mitigated by coarse thinking, namely by the fact that under low elaboration DM cannot distinguish an offer that comes from P from an offer that does not come from P. In particular, the dependency of DM’s expectations on P’s behavior decreases in the degree of coarse thinking that, in turn, positively affects the overall likelihood that an offer does not come from P.

To formalize this claim it is useful to introduce the parameter $\chi$ which measures the degree of coarse thinking, or, to say that in other words, how coarse is coarse thinking. We assume that $\chi$ ranges from 0 to $\infty$, where $\chi$ close to 0 means that low elaboration allows in any case a good level of elaboration, making DM uncertain between the current offer by P and only a few other offers, while $\chi$ very high means that thinking is very coarse when DM resorts to low elaboration, and hence the number of offers among which she is unable to distinguish is very large. More precisely, if we denote with $N_x$, $N_y$, and $N_P$ the absolute number of offers that, respectively, fall into category $x$ and are not made by P, fall into category $y$ and are not made by P, and are made by P, we have that $N_x$ is non-decreasing in $\chi$ and $N_x \to \infty$ when $\chi \to \infty$, and an analogous assumption is made for $N_y$, while we can reasonably assume that $N_P$ does not depend on $\chi$. We also define $\alpha_x = N_x/(N_x + N_y + N_P)$,
\[ \alpha_y = \frac{N_y}{(N_x + N_y + N_P)} \quad \text{and} \quad \alpha_P = \frac{N_P}{(N_x + N_y + N_P)} \]. We then have that \( \alpha_P \) is non-increasing in \( \chi \), \( \alpha_P \to 0 \) when \( \chi \to \infty \), and both \( \alpha_x \) and \( \alpha_y \) are non-decreasing in \( \chi \) and \( \alpha_x + \alpha_y \to 1 \) when \( \chi \to \infty \).

Because more severe coarse thinking reduces the degree of endogeneity of \( \hat{\beta}_x \) and \( \hat{\beta}_y \), one can expect that for a large enough \( \chi \) both equilibrium inexistence and equilibrium multiplicity are no longer an issue. The following proposition, which proof can be found in Appendix B, states that this is indeed the case:

**Proposition 7 (Equilibrium existence and uniqueness).**

*If coarse thinking is strong enough, then almost always an equilibrium exists and is unique.*

### 4 Interpreting the model: Some examples

The model developed in this paper is aimed at improving our understanding of the phenomenon of persuasion. As such, and despite its simplicity, the model must prove to be flexible enough to fit the variety of persuading activities that, as we have argued in the Introduction, are many and widespread in our economic and social life.\(^6\) In the attempt to convince the reader that this is actually the case, we briefly discuss a wide range of situations that reasonably fit the model. Table 1 summarizes them.

As sketched in the Introduction, milk is usually sold to consumers either in glass bottles or in plastic bottles. Glass is more expensive than plastic, but can be used by the seller to induce a higher expectation. Indeed, if consumers exert low elaboration effort and observe a glass bottle containing milk, then they generically think of products in glass containers, whose average quality can be higher than that of products in plastic containers. If consumers choose to rely on high elaboration, then they have to read and understand data on labels, recover information from memory, and put effort to fully assess the quality of the product.

A seller has to decide where to set up his shop, either in the costly location where most shops sell high quality products or in the cheap location where most shops sell low quality

---

\(^6\)The model can be seen as a natural variant of Crawford and Sobel (1982) where talk is not cheap but costly in a twofold sense: the sender has to sustain a cost to send one signal, \( x \), instead of another one, \( y \), and the receiver has to incur a cost in order to learn what action is best for her, \( Y \) or \( N \), without relying on the signal. Obviously, the cost that the receiver has to bear can have a non-psychological nature. It might be a search cost or just a time cost, or even a direct monetary cost to get verifiable information. Although the model fits well these cases too, we stress that we want to focus on situations where the cost that the receiver has to bear is psychological. We leave this alternative interpretation, and the related adjustments to the model and its extensions, for future research.
Table 1: Examples of situations with a persuader-persuadee interaction that can reasonably fit the model.

<table>
<thead>
<tr>
<th>choice/market</th>
<th>P</th>
<th>DM</th>
<th>cue y</th>
<th>cue x</th>
</tr>
</thead>
<tbody>
<tr>
<td>food selling</td>
<td>milk producer</td>
<td>consumer</td>
<td>glass bottle</td>
<td>plastic bottle</td>
</tr>
<tr>
<td>shop location</td>
<td>retailer</td>
<td>consumer</td>
<td>expensive street</td>
<td>cheap street</td>
</tr>
<tr>
<td>financing</td>
<td>loan agent</td>
<td>entrepreneur</td>
<td>expensive dress</td>
<td>cheap dress</td>
</tr>
<tr>
<td>essay evaluation</td>
<td>student</td>
<td>teacher</td>
<td>long essay</td>
<td>short essay</td>
</tr>
<tr>
<td>job market</td>
<td>candidate</td>
<td>recruiter</td>
<td>many certifications</td>
<td>few certifications</td>
</tr>
<tr>
<td>project approval</td>
<td>policy-maker</td>
<td>assembly</td>
<td>merit argument</td>
<td>standard argument</td>
</tr>
<tr>
<td>fundraising</td>
<td>fundraiser</td>
<td>donor</td>
<td>face-to-face</td>
<td>by post</td>
</tr>
</tbody>
</table>

products. Consumers know the average quality of the products sold in each location, but in order to assess the quality of any single product they have to scrutinize it carefully, bearing an elaboration cost.

A loan agent working for a financial institution has to decide whether to lend money to an entrepreneur. The agent may put effort in studying the details of the investment plan for which the entrepreneur is asking money, but this can turn out to be a task requiring a substantial cognitive effort. Alternatively, the agent can look at how the entrepreneur presents himself considering, for instance, his dressing or his club memberships, and then can base her decision on the average quality of investment plans by similar entrepreneurs. Obviously, an elegant dress is more expensive than a casual one and a more exclusive club has higher fees than an unfashionable one.

A teacher, when marking a student’s essay, may spend cognitive resources to scrutinize it carefully, or she can simply choose a mark on the basis of the length of the essay, considering that longer essays are usually better than shorter ones. Writing a longer essay is evidently more time demanding for a student.

A candidate who has applied for a job needs to convince the recruiter that she is qualified for that job. Before the interview the candidate can obtain a bunch of certified qualifications, which require time and effort to get. The recruiter can exert high effort and analyze each qualification in detail, trying to understand what is the real productivity of the candidate. Alternatively, the recruiter can rely on low elaboration and just consider that many qualifications typically go with a good candidate, while a bad candidate has only few qualifications.

A policy-maker who wants to carry out a project which is subject to approval by an
assembly has to convince the members of the assembly that the benefits of the project are higher than its costs. The policy-maker can modify the project in such a way that it appears to belong to a category of projects that the assembly tends to consider favourably, e.g., improving environmental protection, defending human rights, or enforcing equality of opportunities. By doing this, however, the policy-maker can incur in a cost due to a contrast with his own preferences or due to opportunity costs in the use of the resources available for the project. The voters in the assembly can exert low elaboration effort and evaluate the project on the basis of their preferences for the relevant category of projects, or otherwise they can exert high effort and analyze the details about costs and benefits.

A fundraiser working for a non-profit organization aims at collecting voluntary contributions from a potential donor to finance a charitable initiative. The fundraiser has a leaflet, full of detailed information about the destination of donations and the trustworthiness of the non-profit organization. He can decide to send the leaflet through the postal service for a very low fare, or to deliver it in person, which is more costly in terms of both time and money. The leaflet contains all the information needed by the donor to assess the merit of the initiative, but the scarce familiarity of the donor with the specific initiative makes it cognitively costly to extract all relevant information. Alternatively, the donor can take a decision on the basis of whether the leaflet was delivered in person or via mail. The average quality of initiatives for which fundraisers have personally talked to the potential donor may well be higher than the quality of initiatives where nobody has shown up.\textsuperscript{7}

5 Consistency with psychological findings

The way in which we have modeled reference cues and costly elaboration turns out to be consistent with a number of findings of dual process models of persuasion – and, in particular, of the ELM and the HSM. For clarity we organize them in remarks.

Remark 1. \textit{The persuadee interprets the persuasive message differently under high elaboration and under low elaboration} (Petty and Cacioppo, 1986a; Eagly and Chaiken, 1993).

This fundamental idea of dual process theories is directly embedded in the way we model how DM can extract information from the message $(q, r)$: under $L$ only $r$ is observed, while under $H$ also $q$ is.

\textsuperscript{7}Della Vigna et al. (2012) found in a field experiment that a door-to-door fundraising activity is more effective for charitable purposes.
**Remark 2.** The recourse to high elaboration is more likely if persuadee’s motivation is higher (Petty and Cacioppo, 1979, 1981; Cacioppo et al., 1983; Petty et al., 1981).

To formally capture this fact, we may let $U_G$ and $U_B$ be multiplied by a parameter $\gamma$, with $\gamma$ measuring motivation. We observe that, if $c_e > \gamma U_G U_B / (U_G + |U_B|)$, then an increase in $\gamma$ may revert the inequality so that HYN passes from being never best reply to being so for some intermediate range of conditional beliefs on quality. Further, if $c_e < \gamma U_G |U_B| / (U_G + |U_B|)$, then an increase in $\gamma$ enlarges the set of beliefs against which HYN is best choice.

**Remark 3.** A greater ability to think and focus on the content of the message makes it more likely that the persuadee recurs to high elaboration (Petty et al., 1976; Petty and Cacioppo, 1986b).

An increase in DM’s cognitive skills can be translated into our model as a decrease of the elaboration cost required to find out the quality. We simply observe that a decrease of $c_e$ has the same consequences as an increase of $\gamma$ as discussed above.

**Remark 4.** The persuader can use cues to affect the persuadee’s choice of the elaboration level and, through it, the persuadee’s attitudes (Petty and Cacioppo, 1986b; Chaiken and Trope, 1999).

To see why it is so in our model, note that if DM’s optimal behavior given the belief $\hat{\beta}_x(\rho(G), \rho(B))$ differs from DM’s optimal behavior given the belief $\hat{\beta}_y(\rho(G), \rho(B))$, then P can affect DM’s decisions by choosing between $x$ and $y$.

**Remark 5.** Persuasion under high elaboration is stabler than persuasion under low elaboration (Petty and Cacioppo, 1986b; Haugtvedt and Petty, 1992).

To model this consider the following variant of our setup. Subsequently to the choice of the elaboration level, but prior to the choice to accept or reject the offer, DM receives a private signal with probability $\alpha_s$ that reveals $q$. Hence, now $LY$ leads to accept only if no signal is received or the signal reveals $G$, $LN$ leads to reject the offer only if no signal is received or the signal reveals $B$, while $HYN$ leads to the same behaviors as before – i.e., it is independent of the signal. In particular, we can see that under $HYN$, the initial decision to accept or reject is maintained with probability 1. Instead, under $LY$, an initial decision to accept is maintained with probability $1 - \alpha_s + \alpha_s \beta < 1$, while under $LN$ an initial decision to reject is maintained with probability $1 - \alpha_s + \alpha_s (1 - \beta) < 1$. We note that, incidentally, $\alpha_s$ acts against $HYN$, similarly to an increase of $c_e$. 

23
**Remark 6.** Prejudice can affect the elaboration level and induce biased elaboration (Petty and Cacioppo, 1986b; Petty et al., 1999).

Prejudice can be translated in our model by allowing DM to have biased priors about $\alpha_x$, $\alpha_y$, $\alpha_G$, $\beta_x$ and $\beta_y$, which lead to an expected quality $\tilde{\beta}_x$ and $\tilde{\beta}_y$ which are biased with respect to $\hat{\beta}_x$ and $\hat{\beta}_y$. The consequence that the choice between $H$ and $L$ is affected by prejudice is immediate: both $\tilde{\beta}_x$ and $\tilde{\beta}_y$ directly depend on the (biased) values of $\alpha_x$, $\alpha_y$, $\alpha_G$, $\beta_x$ and $\beta_y$. For instance, if DM has a prejudice against $x$ – i.e., $\tilde{\beta}_x\left(\rho(G),\rho(B)\right) < \hat{\beta}_x\left(\rho(G),\rho(B)\right)$ – then she could choose $HYN$ in place of $LY$ or $LN$ in place of $HYN$. Instead, to create the room for biased elaboration we have to modify the model further in order to have $\hat{\beta}_x$ and $\hat{\beta}_y$ playing a role also under high elaboration. This can be obtained by assuming that $H$ does not allow to observe $q$ directly, but to receive a signal on $q$ which is truthful with probability greater than one-half and smaller than one. So $\tilde{\beta}_x$ and $\tilde{\beta}_y$, and hence prejudice, can affect the updated beliefs about $q$ even under $H$.

**Remark 7.** The persuader can send arousing and other mood-affecting cues to induce low elaboration (Petty and Cacioppo, 1986b; Sanbonmatsu and Kardes, 1988).

We can easily introduce in our model the possibility for $P$ to induce arousal (or other mood-affecting factors) in DM, with the aim of increasing the likelihood that DM relies on low elaboration. Let us add, beside $x$ and $y$, a further characteristic of the offer: a mood cue that we indicate with $m \in \{a, n\}$, where $m = a$ means that the cue induces arousal and $m = n$ that it does not. The categories of offers are hence four: $(x, a)$, $(y, a)$, $(x, n)$, and $(y, n)$. If the offer comes from Nature, then each category has a positive probability to contain the offer. If instead the offer comes from $P$, then he has to choose not only the reference cue between $x$ and $y$, but also the mood cue between $a$ and $n$, where choosing $a$ costs $c_a > 0$ and $n$ costs nothing. Under $L$, DM observes both $r$ and $m$, and if $m = a$ the cost of undertaking $H$ increases from $c_e$ to $c_e' > c_e$. As a result we can have the persuader of type $B$ that can use the mood cue together with a reference cue to induce in DM low elaboration and, thanks to this, acceptance of his offer. Figure 8 illustrates a case where the mood cue does not convey information on average quality – i.e., categories $(x, a)$ and $(x, n)$ have the same average quality, and similarly for categories $(y, a)$ and $(y, n)$ – but nevertheless the persuader of type $B$ would like to send $m = a$. Indeed, $B$ has his offer rejected if he pools with type $G$ by sending $(y, n)$ while, thanks to arousal, if $B$ sends $(x, a)$ then DM chooses $LY$ and, hence, accepts $B$’s offer. We note that persuader of type $G$ has two reasons not to send $m = a$: first, it would cost $c_a$ and, second, it would lead DM to choose $LN$ and to reject $G$’s offer.

24
Figure 8: A pooling profile where both G and B send reference cue y and mood cue n, and where DM finds it optimal to reply with HYN to both cue x and y. It is assumed that the average quality of Nature’s offers in category (y, a) is the same of category (y, n) – and similarly for categories (x, a) and (x, n) – and that coarse thinking is strong enough to have that P’s behavior has a negligible impact on average qualities. So, expected quality $\hat{\beta}_x$ and $\hat{\beta}_y$ depend only on the reference cue, not on the mood cue. Without the possibility to send a mood cue, this profile would be a pooling equilibrium with low signal. However, if arousal can be induced then this is not an equilibrium anymore. Indeed, under mood cue a the range of beliefs for which HYN is a best reply shrinks, leading to a different outcome: if P sends x then DM reacts with LY, while if P sends y then DM reacts with LN. Hence, B would deviate sending (x, a).

6 Extensions

In this section we explore a number of extensions of our model with the aim of checking the robustness of our findings. Basically, we study if and how the characteristics of equilibria might change in a less stylized setup. We stress that the details of the model remain unchanged where not specified otherwise.

6.1 Many offer qualities

Consider the case where the quality of offers is not limited to two levels, G and B, but can take many possible values. Let us index qualities on the interval of the real line $[\underline{q}, \bar{q}]$, where $\underline{q} > 0$ is minimum quality and $\bar{q}$ is maximum quality. Nature determines the quality of the offer according to the cumulative distribution $F^N$ in case N is chosen and according to the cumulative distribution $F^P$ in case P is chosen. The values of $\beta$s are modified accordingly.

DM’s utility is given by $U(q)$, which is strictly increasing in q and takes both positive and
negative values over $[\bar{q}, q]$. In particular, there exists $\tilde{q}$ such that $U(\tilde{q}) = 0$. DM would like to accept any offer of quality $q \geq \tilde{q}$ and reject any offer of quality $q \leq \tilde{q}$. Hence, optimal choice by DM is still described by Proposition 1, where $U_G$ and $U_B$ are replaced by, respectively, $\tilde{U}_G = \int_{\tilde{q}}^{q} U(q) dF(q)$ and $\tilde{U}_B = \int_{\tilde{q}}^{q} U(q) dF(q)$, with $F = \alpha P F^P + (1 - \alpha P) F^N$.

Further, persuaders of types $q \geq \tilde{q}$ find it optimal to behave like $G$ in the model of Section 2, while persuaders of types $q \leq \tilde{q}$ find it optimal to behave like $B$. So, Proposition 2 still describes the optimal choice by P conditionally on the potentially optimal behavior by DM, where however $\rho(G)$ and $\rho(B)$ are interpreted as referring to, respectively, types in $[\hat{q}, q]$ and types in $[q, \tilde{q}]$, and where again $U_G$ and $U_B$ are replaced by $\tilde{U}_G$ and $\tilde{U}_B$.

As a consequence, the substance of the findings reported in Proposition 3 through 7 remains true when we allow for many different qualities of the offer.

### 6.2 Many offer categories and reference cues

Consider the case where the categories of objects known by DM, and to which the offer might be referred to, are not just $x$ and $y$, but possibly a large number. Let $Z$ be the set of natural numbers $\{1, 2, \ldots, n\}$ indexing the different offer categories, with $n \geq 2$ and $z \in Z$ denoting the generic reference cue. Suppose also that both $\beta_z$, which denotes the average quality for category $z$, and $c_z$, which denotes the cost of sending a cue referring to category $z$, are strictly increasing in $z$. Finally, to rule out uninteresting cases let also $c_z < V$ for all $z \in Z$.

We note that the optimal choice by DM is still described by Proposition 1. However, to describe the optimal choice by P conditionally on the potentially optimal behavior by DM one needs to generalize Proposition 2 to the case of many offer categories. We do not enter the details of this, since the forces driving the choice by P remain the same. Indeed, P will choose a reference cue that leads his offer to be accepted, if such a cue exists. We note that, as in the basic setup, type $G$ of P has more chances to attain this objective, since for him it is enough to choose a cue that induces $HYN$ as best reply by DM, while type $B$ of P needs a cue that leads to immediate acceptance – i.e., that leads DM to choose $LY$. Moreover, among cues that lead to the same best reply by DM, P will choose the one with the minimum index, since that cue is the least costly. The difference with respect to the setup of Section 2 is the larger number of choices that are available to P, and consequently the larger number of strategies for DM, since a strategy for her is now $\delta : Z \rightarrow \{LN, HYN, LY\}$. A proposition describing the best reply behavior by P should consider all possible behaviors by DM, and hence it would consist of many cases, whose explicit listing would add little to intuition.

From the detailed description of the potentially optimal choices by DM and P – i.e.,
the counterparts of Propositions 1 and 2 in this setup – one can obtain results that are in line with Proposition 3 through 6. To see that pooling and separating equilibria with both signaling and reverse-signaling can emerge with many categories as well, it is enough to think of the equilibria described for the model of Section 2 and add some further categories whose average quality induces the same best reply by DM as done by x or y, but with the associated reference cues being more costly, so that the additional categories are never chosen by P. Moreover, it is easy to understand that as coarse thinking becomes stronger and stronger, existence and uniqueness of the equilibrium are almost always ensured, similarly to what happens in with Proposition 7.

Even if all types of equilibria remain possible in the presence of more than two categories, we remark that reverse-signaling becomes the more likely outcome as the number of categories increases and average qualities are more spread all over [0, 1]. To show this formally, let us introduce a measure of how categories are densely distributed in terms of their average qualities. More precisely, given $0 \leq \beta_1 < \beta_2 < \ldots < \beta_n \leq 1$, we define $\xi$ to be equal to the largest difference of two consecutive numbers in the above sequence. In other words, $\xi$ is the minimum length that an interval of average qualities has to have to be sure that it contains at least the average quality of one category in Z. So, the lower $\xi$ is, the more densely distributed average qualities are.

We are now ready to state the following result:

**Proposition 8 (Reverse-signaling with many offer categories).**

Suppose that $c_e < U_G |U_B|/(U_G + |U_B|)$ and $\xi < \min \{1 - c_e/|U_B| - c_e/U_G, c_e/|U_B|\}$. If coarse thinking is strong enough, then almost always there exists a profile $(\delta, \rho)$ that is the unique equilibrium and such that $\hat{\beta}_{\rho(G)} < \hat{\beta}_{\rho(B)}$, $\delta(\rho(G)) = HYN$, and $\delta(\rho(B)) = LY$.

**Proof.** Since $c_e < U_G |U_B|/(U_G + |U_B|)$, then there exists an interval of beliefs against which HYN is the best reply by DM. Moreover, since $\xi < \min \{1 - c_e/|U_B| - c_e/U_G, c_e/|U_B|\}$, we are sure that there exist at least one category whose average quality induces HYN as best reply, and at least one category whose average quality induces LY as best reply. We denote with $z_{HYN}^{LY}$ the category with the minimum index among those which are best replied with HYN, and we define $z_{LY}^{LY}$ analogously. We suppose that $z_{HYN}^{LY}$ and $z_{LY}^{LY}$ are interior points in the intervals of beliefs that are best replied, respectively, with HYN and LY. By so doing we are neglecting values that have measure zero in the parameters space.

We now build a profile that we then check to be an equilibrium. For every $z \in Z$, we set $\rho(z)$ as the best action by DM against $\beta_z$, as shown by Proposition 1. Then we set $\delta(G) = z_{min}^{HYN}$ and $\delta(B) = z_{min}^{LY}$. We note that such $\delta$ selects for both $G$ and $B$ the least costly
cue that allows them to have their offer accepted by DM. This shows the optimality of P’s strategy against \( \rho \). To understand that \( \rho \) is optimal against \( \delta \), it is enough to observe that, when coarse thinking is strong enough, \( \hat{\beta}_{\text{HYN}} \) and \( \hat{\beta}_{\text{LY}} \) are arbitrarily close to, respectively, \( \beta_{\text{HYN}}^{\text{min}} \) and \( \beta_{\text{LY}}^{\text{min}} \), and hence induce the same best action by DM, since \( \beta_{\text{HYN}}^{\text{min}} \) and \( \beta_{\text{LY}}^{\text{min}} \) are interior points in the intervals of best reply behavior by DM.

Finally, uniqueness follows exactly for the same argument shown in Proposition 7, which hence we do not repeat.

6.3 A continuum of elaboration intensities: Quality signal

Consider the case where DM has not to choose between high and low elaboration, but – as suggested by Petty and Cacioppo (1986b) themselves – has to select an elaboration intensity out of an elaboration continuum. The basic idea is that the likelihood of extracting \( q \) from the message sent by P increases in the elaboration intensity. To model this let us assume that DM never observes the part \( q \) of the message sent by P, but she can extract from it a signal \( \sigma \) which conveys information on \( q \), and that the precision of \( \sigma \) depends on the elaboration intensity \( e \geq 0 \) that DM chooses. More precisely, \( p(e) \) is the probability that the signal is correct – i.e., it says \( G \) when the quality is \( G \) and \( B \) when the quality is \( B \) – and \( 1 - p(e) \) is the probability that the signal is wrong – i.e., it says \( B \) when the quality is \( G \) and \( G \) when the quality is \( B \). Further, we model the fact that the extraction of correct information about \( q \) is more likely under greater elaboration intensity by letting \( p'(e) > 0 \), and the fact that greater elaboration has increasing marginal costs by letting \( p''(e) < 0 \).

Clearly, choosing a positive level of elaboration intensity makes sense only if the acquired information is then used, i.e., only if also HYN is chosen. The expected utility of choosing HYN with elaboration intensity \( e^* \) is \( \mu p(e^*)U_G - (1 - \mu)(1 - p(e^*))|U_B| - e^* \), while the expected utilities of choosing LY and LN with null elaboration intensity are as in the model of Section 2, i.e., \( \mu U_G - (1 - \mu)|U_B| \) and 0, respectively. Solving the inequalities among these expected utilities, we find that LN with \( e^* = 0 \) is best reply for DM when \( \mu \leq (e^* + (1 - p(e^*))|U_B|)/(p(e^*)U_G + (1 - p(e^*))|U_B|) \) and \( \mu U_G - (1 - \mu)|U_B| \leq 0 \), while LY with \( e^* = 0 \) is best reply when \( \mu \geq (e^* + (1 - p(e^*))|U_B|)/(p(e^*)U_G + (1 - p(e^*))|U_B|) \) and \( \mu U_G - (1 - \mu)|U_B| \geq 0 \).

Figure 9 summarizes DM’s optimal behavior as a function of \( \mu \), providing a counterpart of Proposition 1 for the current setup. We note that while the quality of DM’s optimal behavior remains substantially unchanged with respect to the model of Section 2, the relaxation of the hypothesis of just two elaboration levels gives an extra role to \( U_G \) and \( U_B \): under HYN the optimal elaboration intensity increases, decreases, or is constant in expected quality.
depending on whether $U_G > |U_B|$, $U_G < |U_B|$, or $U_G = |U_B|$. This is because the nature of
the stake – i.e., whether it is most important to get the good offer or to avoid the bad one –
determines whether direct information on quality and expected quality are complements or
substitutes.

Figure 9: In subfigures (a), (b) and (c) there exists an interval of beliefs against which $HYN$ is
best reply, which is the case for $p(e^*) > 1/2 + \frac{e^*(U_G + |U_B|)}{2U_G|U_B|}$. Moreover, $LN$ is best reply for beliefs
lower than $\frac{e^*+(1-p(e))|U_B|}{p(e)U_G+(1-p(e))|U_B|}$, while $LY$ is best reply for beliefs higher than $\frac{p(e)|U_B|-e^*}{(1-p(e))U_G+p(e)|U_B|}$, and
$HYN$ is best reply for intermediate beliefs. In subfigure (d), instead, $p(e^*) < 1/2 + \frac{e^*(U_G + |U_B|)}{2U_G|U_B|}$
and $HYN$ is never best reply, while $LN$ and $LY$ are best reply for beliefs, respectively, lower than
and higher than $\frac{|U_B|}{U_G+|U_B|}$. 

29
We also note that the optimal choice by P is still described by Proposition 2, meaning that the actual level of $e^*$ under $HYN$ is irrelevant to P. By the same token, Proposition 10 and Proposition 7 remain true. Further, the characterization of equilibria made by Proposition 3 through 6 still holds in qualitative terms, but the statements have to be adjusted by replacing conditions on $c_e$ with the appropriate conditions from those described above.

Finally, we observe that the current extension could be slightly changed by assuming that with probability $p(e)$ the signal $\sigma$ is correct, as before, but, differently from before, with probability $1 - p(e)$ the signal does not arrive at all. This represents the case where failure to extract from signal $\sigma$ reliable information on $q$ does not result in extracting wrong information, but just no information. While in this alternative setup the optimal choice of elaboration by DM would be partly different under $HYN$ – potentially u-shaped – the substance of what discussed in this subsection remains true.

6.4 Cues with fully endogenous quality

Proposition 6 tells us that a reverse-signaling equilibrium can arise in our model, and Proposition 8 states that this type of equilibrium is the only type that is possible when there are many categories and average qualities are sufficiently densely distributed. The combination of these results suggests that reverse-signaling may be a quite relevant case. However, the assumed partial exogeneity of expected average quality that is associated with each reference cue can raise doubts about the relevance of reverse-signaling. Indeed, in the model of Section 2 (and similarly in the model of Subsection 6.2) the average quality of offers not coming from P, i.e., $\beta_x$ and $\beta_y$, is exogenous and, in particular, one category of offers is of better average quality than the other by assumption, i.e., $\beta_x > \beta_y$. We observe that in a reverse-signaling equilibrium the average quality remains higher in category $x$ than in category $y$, despite the fact that when P is called to play he chooses cue $y$ if his type is $G$ and cue $x$ if his type is $B$. In order for this to occur, the decisions influencing the quality of offers in category $x$ and $y$ that are not made by P should be sustained by sound explanations. In the following we explore one particular explanation: for a subset of offers in each category, cues $x$ and $y$ denote characteristics which provide an intrinsic utility to DM, so that for such offers the choice between $x$ and $y$ directly affects the quality of the offer.

Consider the model presented in Section 2, but modified as depicted in Figure 10. Initially, either an offerer O or a persuader P are randomly selected with probability $\alpha_O$ and $\alpha_P = 1 - \alpha_O$, respectively. If P is chosen then the game unfolds as in the original model. Instead, if O is chosen then the game unfolds as follows. Firstly, a type for O is drawn, either the type $A$ – who is endowed with advanced technology – or the type $S$ – who is endowed
with standard technology; types are selected with probability $\alpha_A$ and $\alpha_S$, respectively. Secondly, $O$ chooses between $x$ and $y$. Lastly, $DM$ observes the reference cue without knowing whether the offer comes from $O$ or from $P$, and has to decide on both elaboration level and reaction.

The two types of player $O$ are characterized by different technologies that induce different costs to choose $x$. In particular, the advanced technology allows type $A$ to incur a lower cost than type $S$ to employ $x$, i.e., $c^A_x < c^S_x$. We also assume that $V - c^A_x > 0$ and $V - c^S_x < 0$, meaning that type $A$ prefers $x$, while type $S$ prefers $y$. Moreover, the offer made by $O$ is such that the cue has an intrinsic value for $DM$, meaning that $x$ qualifies the offer as good, while $y$ qualifies the offer as bad. In other words, the offer is of good quality if $x$ is chosen, while it is of bad quality if $y$ is chosen. Finally, a strategy by $O$ – which we denote by $\omega$ – is a choice between $x$ and $y$ as a function of the type, i.e., $\omega : \{A, S\} \rightarrow \{x, y\}$.

The following proposition shows that an equilibrium can arise where $P$ plays a reverse-signaling strategy. We note that such an equilibrium is crucially sustained by the fact that average quality in category $x$ is higher than average quality in category $y$, which in turn is possible thanks to the existence of a subset of offers whose quality is directly affected by $x$ and $y$.

**Proposition 9** (Reverse-signaling with endogenous quality of cues).

There exist values for $c_e$, $\alpha_0$ and $\alpha_A$ such that the following profile $(\omega, \rho, \delta)$ is an equilibrium: $\omega(A) = x$, $\omega(S) = y$, $\rho(G) = y$, $\rho(B) = x$, $\delta(x) = LY$, $\delta(y) = HYN$.

**Proof.** By Proposition 1 we know that if $c_e \leq \frac{U_G}{U_G + U_B}$ then there exists some value for the belief on quality such that $HYN$ is best reply for $DM$, and clearly for a belief on quality that is high enough $LY$ is best reply.

We observe that $\hat{\beta}_x = \frac{\alpha_O \alpha_A}{\alpha_O \alpha_A + \alpha_P \alpha_B}$ and $\hat{\beta}_y = \frac{\alpha_P \alpha_G}{\alpha_P \alpha_G + \alpha_O \alpha_S}$. By inspection of these expressions – and remembering that $\alpha_O + \alpha_P = 1$ – we see that $\hat{\beta}_x$ approaches 1 if we choose $\alpha_O$ high enough. Given all other parameters, we pick a value of $\alpha_O$ such that $LY$ is best reply against the corresponding value of $\hat{\beta}_x$. Now we turn our attention to $\hat{\beta}_y$. If $\hat{\beta}_y$ is such that the best reply by $DM$ is $LY$, we can raise $\alpha_O$ until $\hat{\beta}_y$ decreases enough to take a value such that $DM$ finds it optimal to reply with $HYN$. If, instead, $\hat{\beta}_y$ is such that the best reply by $DM$ is $LN$ – remembering that $\alpha_A + \alpha_S = 1$ – we can raise $\alpha_A$ until $\hat{\beta}_y$ increases enough to take a value such that $DM$ finds it optimal to reply with $HYN$. In both cases, $\hat{\beta}_x$ gets higher, and hence $LY$ remains best reply when $x$ is observed.

---

8We remark that the null payoff that type $S$ obtains by choosing $Y$ if $DM$ replies with $N$ should not be interpreted as “getting nothing”, since payoff levels are normalized (see Subsection 4) and, hence, such an occurrence may represent the fact that $DM$ buys at a lower price or a smaller quantity.
Figure 10: The game tree with endogenous quality of categories. In the left branch DM faces an offer that comes from O, which has probability $\alpha_A$ to be from type $A$ (the offerer endowed with advanced technology). In the right branch DM faces an offer made by P, which has probability $\alpha_G$ to be of quality $G$. DM has two information sets each of which encompasses nodes from both branches; one set is associated with reference cue $x$ and the other set with reference cue $y$. 
Therefore, by choosing $c_e$ low enough and by an appropriate choice of $\alpha_0$ and $\alpha_A$ we can obtain that DM has no incentive to deviate from the considered strategy profile. The optimality checks for O and P are straightforward, and hence omitted.

7 Related literature

In this section we review the main contributions on persuasion, distinguishing between economic and psychological literature. We focus mainly on theory.\footnote{For a recent survey of empirical evidence regarding persuasion activities, especially related to economics and politics, see Della Vigna and Gentzkow (2010).} We also briefly review some recent papers that focus on scarce cognitive resources, stressing similarities and differences with our model.

7.1 Persuasion in economics

In recent years several models have been proposed that study how a message sent by a persuader can affect a persuadee’s behavior. We can distinguish among them on the basis of different criteria.

A first criterion is whether the act of persuasion is belief-based or non-belief-based – i.e., preference-based. Non-belief based persuasion affects behavior independently of beliefs. In such a case persuasion is obtained because the message itself impacts on utility and, hence, on behavior (Stigler and Becker, 1977; Becker and Murphy, 1993). This is also reminiscent of models of persuasive advertising (Braithwaite, 1928). Instead, belief-based persuasion affects behavior by changing persuadees’ beliefs. For instance, a persuadee can be persuaded by informative communication (Stigler, 1961; Telser, 1964).

Among the models where persuasion is understood as belief-based, a further distinction can be made between models where agents are perfect Bayesian updaters and models where they are not. When agents are Bayesian updaters, a persuader, given a communication technology, chooses signals that can be costly (Nelson, 1970) or not (Gentzkow and Shapiro, 2006; Kamenica and Gentzkow, 2011) and that are correctly elaborated. Instead, non-fully Bayesian agents have limitations in the way they elaborate the signal: e.g., they are constrained by limited memory (Mullainathan, 2002; Shapiro, 2006), they double-count repeated information (De Marzo et al., 2003), they neglect the incentives of the sender (Eyster and Rabin, 2010), or they have limitations in the ability to lie (Glazer and Rubinstein, 2012).

Finally, models can be distinguished on the basis of the nature of the information sent:
hard versus soft. Hard information is actually verifiable, while soft information is not. Cheap talk models typically rely on soft information (Crawford and Sobel, 1982), while models that exploit the strategic use of verifiable information (Milgrom and Roberts, 1986) can consider a verification cost (Caillaud and Tirole, 2007), full and costless verifiability (Glazer and Rubinstein, 2006) or only partial verifiability with the receiver deciding which part to be verified (Glazer and Rubinstein, 2004).

7.2 Persuasion in psychology

In psychology the term persuasion refers to a broader and vaguer phenomenon than in economics: activities aimed at influencing others’ behaviors. However, the most relevant difference is that psychologists assume that the persuadee can process the message sent by the persuader at different degrees of effectiveness, which are associated with different cognitive costs.

Psychologists often stress that persuasion activities exploit the fact that individuals have two distinct ways of processing information when they receive a message and have to take decisions based on it (Chaiken and Trope, 1999). Theories in cognitive and social psychology that refer to this idea are typically labelled as dual process theories (Evans, 2003), and the two ways of processing information are also called System 1 and System 2. Kahneman (2003) refers to System 1 and System 2 as, respectively, intuition and reasoning. Recent neurological research (Goel et al., 2000) suggests that different parts of our brain are actually activated when using System 1 and System 2, respectively. Dual process theories have been applied to explain human behaviors in different setups (Gawronski and Creighton, 2013): persuasion, attitude-behavior relations, prejudice and stereotyping, impression formation, dispositional attribution.

There are two workhorse models of persuasion in psychology. One is the Elaboration Likelihood Model (ELM) (Petty and Cacioppo, 1986a), where the persuadee can use the “central route” – characterized by a high cognitive effort – or the “peripheral route” – characterized by a low effort. These two routes can be understood as an approximation of a continuum of elaboration intensities which a subject can use when processing information: the higher an individual’s cognitive effort, the more likely that she processes all relevant information. At the extremum characterized by highest level of elaboration individuals use all available information and integrate it with already stored information. On the contrary, at the extremum characterized by lowest level of elaboration individuals minimally scrutinize relevant information, extensively using short-cuts to process information.
The other model is the Heuristic-Systematic Model (HSM) (Chaiken et al., 1989), where the persuadee can use “systematic elaboration” – characterized by careful scrutiny – or “heuristic elaboration” – characterized by the use of simple heuristics, rule of thumbs and categorizations. The basic idea is very similar to that of the ELM. A fully systematic processing of information requires high cognitive effort and considers all relevant information. In contrast, a purely heuristic processing requires minimal cognitive effort and considers only a small amount of information.

One important feature of these models is that the persuader may have the chance to affect the choice of elaboration intensity. This naturally gives rise to a strategic interaction between the persuader and the persuadee.

7.3 The rational allocation of scarce cognitive resources

From a broader perspective, our paper belongs to the small but growing body of literature that tries to model the scarcity of human cognitive resources in a tractable and economically meaningful way, with the central idea that such a scarcity generates an allocation problem that the decision-maker solves rationally. In particular, some recent papers have explored the possibility of incorporating the cost of reasoning into models where agents have to take decisions based on the elaboration of information.

One important paper in this regard is Dewatripont and Tirole (2005), that we have already discussed in the Introduction. Another interesting attempt is the model of decision-making proposed by Dickhaut et al. (2009), where the informativeness of a signal (about the payoff associated with different options) increases in the effort that the decision-maker puts in the observation of the signal. Our model is similar in the modeling of cognitive costs, but adds the possibility of the strategic use of the signal by the sender and the explicit modeling of what happens under low cognitive effort – i.e., coarse thinking.

Brocas and Carrillo (2012) model how the human brain governs both memorization and retrieval of information from memory by assuming that more precise memorization requires more cognitive resources. Our model obtains similar insights – stronger beliefs on the state of the world lead the agent to exert less cognitive effort to acquire new information – in a choice situation where the relevant signal is not retrieved from memory but chosen strategically by a persuader.

In the literature on $k$-level reasoning, Alaoui and Penta (2013) introduce costs to access higher levels of reasoning, and perform a cost-benefit analysis that allows to endogenize the level of reasoning; pursuing this line, Alaoui and Penta (2014) test this model in the lab and find evidence that supports the idea that players weight the value of thinking deeper against
the cost of reasoning. Our model resembles theirs in that there are explicit cognitive costs but, while in our model the costly activity is about information processing, in their models the costly activity is the iteration of the calculation of best-replies.

Matějka and McKay (2012), following Sims (2003), model the cost of attention by means of constraints on measures of uncertainty – e.g., Shannon entropy. By putting constraints on the amount of cognitive resources that can be devoted to information processing, they introduce elaboration costs that depend on the overall uncertainty reduction, leaving the agent quote free to allocate such uncertainty on the possible states of the world. On the contrary, we give the agent only the choice between high and low cognitive effort, assuming that when uncertainty is large – i.e., under low elaboration – the agent relies on coarse thinking; this crucially gives to the persuader enough room to exploit reference cues.

Martin (2012) studies the case of a monopolist that independently interacts with rationally inattentive consumers, showing that the monopolist can strategically exploit the consumers’ cost of attention. This paper is close to ours in allowing the strategic exploitation by the sender of the receivers’ cognitive limitations; however, while we posit that under low elaboration the receivers uses a specific heuristic which is known to the sender – i.e., coarse thinking – in Martin (2012) the receiver can choose how to allocate attention, even when it is very limited.

Caplin and Martin (2014) characterize the testable implications of a model with imperfect perception, providing a unifying framework for all these contributions in the spirit of revealed preference approach. In particular, their framework can easily accommodate dual process reasoners.

Manzini and Mariotti (2014) characterize the testable implications of a model where there is a boundedly rational agent who suffers from limited attention. Although their focus is on stochastic choice data, their model is consistent with the idea that the agent has to incur a cognitive cost to better process the available information.

---

10 A related body of literature considers the case where the decision-maker’s attention is drawn to those payoffs or characteristics which are most different or salient relative to the average or to a reference point (Bordalo et al., 2012; Kőszegi and Szeidl, 2013; Bordalo et al., 2013b). A different but still related model is found (Gennaioli and Shleifer, 2010), where only the information relative to the most representative scenarios is retrieved from memory – i.e., a form of local thinking is analyzed.
8 Conclusions

In this paper we have proposed a model of persuasion that incorporates the insights of social and cognitive psychology on dual process reasoning. We have framed persuasion activities within a sender-receiver model where agents are both rational and Bayesian, but the persuadee has to pay a cognitive cost to extract all information from the message sent by the persuader. The main novelty of our approach is the combination of two distinct features borrowed from the psychological literature: the explicit consideration of elaboration costs to extract all information from a message – i.e., high elaboration is more informative, but also more costly in terms of cognitive resources, than low elaboration – and the low elaboration as coarse thinking – i.e., if the persuadee does not pay the cognitive cost to scrutinize carefully the offer then she is unable to distinguish among offers belonging to the same category. This setup allows us to endow the persuader with a novel strategic tool of persuasion – i.e., cues – that we model as references to categories of objects. We then proceed by studying the strategic use of reference cues by the persuader aimed at manipulating the beliefs of the persuadee.

The proposed model, despite its simplicity, has proved to be rich enough to provide predictions that are in line with well documented findings in the social psychology of persuasion. Moreover, it provides a novel reason for separating equilibria to arise even in the absence of the single-crossing property and, perhaps more importantly, the emergence of a new phenomenon that we name *reverse-signaling*, where high types send low signals and low types send high signals.

The next steps in this line of research would be, we think, to develop less stylized versions of the model to be applied to specific cases of interest as well as exploring dimensions overlooked in the present paper. Among the applications one might want to focus on there is, for instance, political campaigning (where politicians exploit slogans and ideological cues to gain the votes of electors), marketing through product packaging (where firms exploit packaging as reference cues to manipulate consumers’ belief on quality), the engineering of financial products (where sellers exploit product complexity to induce low elaboration by potential purchasers), or fundraising activities (where fundraisers exploits mood cues to induce donations). One dimension left unexplored, which however is likely to play an important role in real persuasive activities, is competition among persuaders. This would go along the lines explored by Gentzkow and Kamenica (2011) for the case of purely Bayesian persuasion and by Bordalo et al. (2013a) for the case of consumers attention. Another issue to be explored is the role of prices when persuasion is aimed at selling products, taking into
account what marketing experts often stress, i.e., that prices are important cues for product quality. This would entail considering models of quality signaling through price, such as Wolinsky (1983).

Acknowledgements

We want to thank Stefano Barbieri, Luis Corchón, Francesco Filippi, Jana Friedrichsen, Massimo Morelli, Antonio Nicolò, Eugenio Peluso, Bradley Ruffle, Anastasia Shchepepetova, and Francesco Squintani, for their useful comments, which helped us to improve the paper. We also want to thank people who have provided insightful comments during the 2012 G.R.A.S.S. workshop hosted by the Franqui Foundation and the L.U.I.S.S. University in Rome, the 2013 C.E.P.E.T. workshop hosted by Udine University, the 2013 EEA-ESEM conference in Gothenburg, the 2013 EARIE conference in Evora, the 3rd workshop on “IO: Theory, Empirics, and Experiments” in Alberobello, the 2014 OLIGO workshop in Rome, the 7th M-BEES workshop in Maastricht, and the seminars held at DEM and LEM in Pisa, and at DISEI in Florence. This paper is part of the project “Persuasion with elaboration costs” financed by Einaudi Institute for Economics and Finance (EIEF), which we gratefully acknowledge. The authors also acknowledge financial support from the Italian Ministry of Education, Universities and Research under PRIN project 2012Z53REX “The Economics of Intuition and Reasoning: a Study On the Change of Rational Attitudes under Two Elaboration Systems” (SOCRATES).

References


Appendix - Multiplicity and inexistence of equilibria: Examples and results

A.1 Multiple equilibria can coexist

An example of multiplicity of equilibria is depicted in Figure 11, where the following two equilibria coexist: a pooling equilibrium with high signal and a reverse-signaling equilibrium.

\[ LN \quad HYN \quad LY \]

\[ \hat{\beta}_y(\rho(G) = x, \rho(B) = x) \quad \mu \quad \hat{\beta}_x(\rho(G) = x, \rho(B) = x) \]

\[ \hat{\beta}_y(\rho(G) = y, \rho(B) = x) \quad \hat{\beta}_x(\rho(G) = y, \rho(B) = x) \]

pooling equilibrium (high signal)

separating equilibrium (high quality with low signal)

Figure 11: An example of the coexistence of two equilibria, for \( c_e < \frac{U_G|U_B|}{U_G + |U_B|} \).

Looking at Figure 11 we see that, when both types of \( P \) choose cue \( x \), then the expected quality \( \hat{\beta}_x \) is high and DM’s best reply is \( LY \), while the expected quality \( \hat{\beta}_y \) is so low that DM’s best reply is \( LY \); this justifies the pooling equilibrium where both type \( G \) and type \( B \) choose the high signal \( x \). At the same time, if type \( G \) switches from \( x \) to \( y \), it may happen that \( \hat{\beta}_x \) remain high enough to have that DM best replies with \( LY \), and \( \hat{\beta}_y \) raises entering the region where DM best replies with \( HYN \); this is what occurs in the case represented in the figure, and it justifies the reverse-signaling equilibrium.

We note, however, that not all types of equilibria can coexist. The following proposition lists the possible cases of coexistence of equilibria:

**Proposition 10** (Equilibrium multiplicity and coexistence).

Multiple equilibria can exist, but only in the following pairs:

- a pooling where \( \rho(G) = \rho(B) = x \) and a separating where \( \rho(G) = y \) and \( \rho(B) = x \);
- a pooling where \( \rho(G) = \rho(B) = y \) and a separating where \( \rho(G) = x \) and \( \rho(B) = y \);
• a pooling where $\rho(G) = \rho(B) = x$ and a pooling where $\rho(G) = \rho(B) = y$.

The proof of Proposition 10 can be found in Appendix B, together with some preliminary technical details. Here we provide the intuition why a reverse-signaling equilibrium – where high quality goes with low signal – cannot coexist with a signaling equilibrium – where high quality goes with high signal. In a signaling equilibrium type $G$ sends cue $x$ and type $B$ sends cue $y$, and for $B$ to send cue $y$ DM’s belief conditional on cue $x$ must be low enough not to induce DM to play $LY$ – otherwise type $B$ would find it profitable to deviate from sending cue $y$ to sending cue $x$. On the contrary, in a reverse-signaling equilibrium type $G$ sends cue $y$ and type $B$ sends cue $x$, but for $B$ to send cue $x$ DM’s belief conditional on cue $x$ must be high enough to have DM play $LY$ – so that $B$ actually has the offer accepted if he sends cue $x$. This two conditions are incompatible because having $G$ sending $y$ and $B$ sending $x$ decreases the belief conditional on $x$ with respect to having $G$ sending $x$ and $B$ sending $y$, so that if beliefs are low enough to sustain a signaling equilibrium then they cannot be high enough to sustain a reverse-signaling equilibrium.

### A.2 An equilibrium may fail to exist

Another possible occurrence is that no equilibrium exists in pure strategies. An example of equilibrium inexistence is depicted in Figure 12, where it can be easily checked that for any given behavior by $P$ the best reply by DM is such that at least one type of persuader strictly gains by deviating.

![Figure 12: An example where no equilibrium exists, for $c_e \leq \frac{U_G U_B}{U_G + U_B}$](image)

Looking at Figure 12 we see that both the separating profile where high quality goes with low signal and the pooling profile with high signal cannot be equilibria because DM best replies with
HYN to cue $x$; in such a case, indeed, the persuader of type $B$ would prefer to send cue $y$ and save on costs, since he will never see his offer accepted. Similarly, both the pooling profile with low signal and the separating profile with high quality going with high signal cannot be equilibria as well, because DM best replies with $LY$ to cue $x$, and hence the persuader of type $B$ would prefer to send cue $x$ (so to have his offer accepted) instead of $y$ (being his offer rejected with such a cue). We remark again that the example depicted in the figure – and more in general the possibility that no equilibrium exists – is due to the fact that $\hat{\beta}_x$ and $\hat{\beta}_y$ move along the segment as types $G$ and $B$ change the choice of cues, thus changing also the optimal behavior of DM.

Actually, an equilibrium exists if mixed strategies are considered. To be convinced of this, start considering the pooling profile where both $G$ and $B$ choose signal $x$; as already remarked, this is not an equilibrium because type $B$ has a profitable deviation from $x$ to $y$. Let us now suppose that type $B$ chooses a mixed strategy where the probability of playing $y$ progressively increases starting from zero. Consequently, $\hat{\beta}_y$ decreases and hence remains in the region where DM best replies with $LN$, while $\hat{\beta}_x$ progressively raises until it reaches the point where DM is indifferent between $HYN$ and $LY$, and hence she can optimally randomize between $HYN$ and $LY$. In particular, she can randomize with probabilities that make type $B$ indifferent between $x$ and $y$. We have hence found a mixed strategy equilibrium where type $G$ chooses $x$, type $B$ randomizes between $y$ and $x$, and DM chooses $LN$ if $y$ and randomizes between $HYN$ and $LY$ if $x$. Starting from the pooling equilibrium where both $G$ and $B$ choose signal $y$, and following a similar reasoning, we can find another mixed strategy equilibrium with the main difference that type $G$ chooses $y$ and DM chooses $HYN$ if $y$.

B Appendix - Proofs and related technical details

B.1 Beliefs on quality conditional on cue and P’s behavior

\[
\hat{\beta}_x(\rho(G) = y, \rho(B) = y) = \beta_x \quad (1)
\]
\[
\hat{\beta}_x(\rho(G) = y, \rho(B) = x) = \frac{\alpha_x \beta_x}{\alpha_x + \alpha_P \alpha_B} \quad (2)
\]
\[
\hat{\beta}_x(\rho(G) = x, \rho(B) = y) = \frac{\alpha_x \beta_x + \alpha_P \alpha_G}{\alpha_x + \alpha_P} \quad (3)
\]
\[
\hat{\beta}_x(\rho(G) = x, \rho(B) = x) = \frac{\alpha_x \beta_x + \alpha_P \alpha_G}{\alpha_x + \alpha_P \alpha_G} \quad (4)
\]
\[
\hat{\beta}_y(\rho(G) = y, \rho(B) = y) = \frac{\alpha_y \beta_y}{\alpha_y + \alpha_P} \quad (5)
\]
\[
\hat{\beta}_y(\rho(G) = y, \rho(B) = x) = \frac{\alpha_y \beta_y + \alpha_P \alpha_G}{\alpha_y + \alpha_P \alpha_B} \quad (6)
\]
\[
\hat{\beta}_y(\rho(G) = x, \rho(B) = x) = \beta_y \quad (7)
\]
\[
\hat{\beta}_y(\rho(G) = x, \rho(B) = y) = \frac{\alpha_y \beta_y}{\alpha_y + \alpha_P \alpha_B} \quad (8)
\]
B.2 Proof of Proposition 2

Proof. Suppose that DM chooses \( \delta(x) = \delta(y) \). If \( \delta(x) = \delta(y) = LY \) or \( \delta(x) = \delta(y) = HYN \) and \( P \) is of type \( G \), then \( P \)'s payoff is \( V - c_r \). If instead \( \delta(x) = \delta(y) = LN \) or \( \delta(x) = \delta(y) = HYN \) and \( P \) is of type \( B \), then \( P \)'s payoff is \(-c_r \). Since 0 = \( c_y < c_x \), \( r = y \) is optimal for \( P \) independently of his type.

Suppose that DM chooses \( \delta(x) = LY \) and \( \delta(y) = LN \). \( P \)'s payoff is equal to \( V - c_x > 0 \), if \( r = x \), and to 0, if \( r = y \). Hence, \( r = x \) is optimal for \( P \) independently of his type.

Suppose that DM chooses \( \delta(x) = HYN \) and \( \delta(y) = LN \). If \( P \) is of type \( G \) then his payoff is equal to \( V - c_x > 0 \), if \( r = x \), and to 0, if \( r = y \). Hence, \( r = x \) is optimal for type \( G \). If \( P \) is of type \( B \) then his payoff is equal to \(-c_x \), if \( r = x \), and to 0, if \( r = y \). Since 0 < \( c_x \), \( r = y \) is optimal for type \( B \).

Suppose that DM chooses \( \delta(x) = LY \) and \( \delta(y) = HYN \). If \( P \) is of type \( G \) then his payoff is equal to \( V - c_x \), if \( r = x \), and to \( V \), if \( r = y \). Since 0 = \( c_y < c_x \), \( r = y \) is optimal for type \( G \). If \( P \) is of type \( B \) then his payoff is equal to \( V - c_x > 0 \), if \( r = x \), and to 0, if \( r = y \). Hence, \( r = x \) is optimal for type \( B \).

Preliminarily, we give the following straightforward results which will be used in the subsequent proof. For \( \alpha_x, \alpha_y, \alpha_G, \beta_x, \) and \( \beta_y \) strictly comprised between 0 and 1, the following inequalities necessarily hold:

\[
\begin{align*}
\hat{\beta}_x(\rho(G)=y, \rho(B)=x) &< \hat{\beta}_x(\rho(G)=y, \rho(B)=y) \quad (9) \\
\hat{\beta}_x(\rho(G)=y, \rho(B)=x) &< \hat{\beta}_x(\rho(G)=x, \rho(B)=x) \quad (10) \\
\hat{\beta}_x(\rho(G)=y, \rho(B)=y) &< \hat{\beta}_x(\rho(G)=x, \rho(B)=y) \quad (11) \\
\hat{\beta}_x(\rho(G)=x, \rho(B)=x) &< \hat{\beta}_x(\rho(G)=x, \rho(B)=y) \quad (12) \\
\hat{\beta}_y(\rho(G)=x, \rho(B)=y) &< \hat{\beta}_y(\rho(G)=x, \rho(B)=x) \quad (13) \\
\hat{\beta}_y(\rho(G)=x, \rho(B)=y) &< \hat{\beta}_y(\rho(G)=y, \rho(B)=y) \quad (14) \\
\hat{\beta}_y(\rho(G)=y, \rho(B)=y) &< \hat{\beta}_y(\rho(G)=y, \rho(B)=x) \quad (15) \\
\hat{\beta}_y(\rho(G)=y, \rho(B)=y) &< \hat{\beta}_y(\rho(G)=y, \rho(B)=x) \quad (16)
\end{align*}
\]

The following is a check that the above inequalities indeed hold:

\[
\hat{\beta}_x(\rho(G)=y, \rho(B)=x) = \frac{\alpha_x \beta_x}{\alpha_x + \alpha_p \alpha_B} = \frac{\beta_x}{1 + 1 - \frac{\alpha_p \alpha_B}{\alpha_x}} < \beta_x = \hat{\beta}_x(\rho(G)=y, \rho(B)=y) \quad (17)
\]

\[
\hat{\beta}_x(\rho(G)=y, \rho(B)=y) = \frac{\alpha_x \beta_x}{\alpha_x + \alpha_p \alpha_B} < \frac{\alpha_x \beta_x + \alpha_p \alpha_G}{\alpha_x + \alpha_p \alpha_B + \alpha_p \alpha_G} = \frac{\alpha_x \beta_x + \alpha_p \alpha_G}{\alpha_x + \alpha_p} = \hat{\beta}_x(\rho(G)=x, \rho(B)=x) \quad (18)
\]

46
\[ \beta_x(\rho(G) = y, \rho(B) = y) = \beta_x = \beta_x \left( \frac{\alpha_x + \alpha_p \alpha_G}{\alpha_x + \alpha_p} \right) = \frac{\alpha_x \beta_x + \alpha_p \alpha_G \beta_x}{\alpha_x + \alpha_p \alpha_G} < \]

\[ < \frac{\alpha_x \beta_x + \alpha_p \alpha_G}{\alpha_x + \alpha_p} = \beta_x(\rho(G) = x, \rho(B) = y) \]  

(19)

\[ \beta_x(\rho(G) = x, \rho(B) = x) = \frac{\alpha_x \beta_x + \alpha_p \alpha_G}{\alpha_x + \alpha_p} = \frac{\alpha_x \beta_x + \alpha_p \alpha_G}{\alpha_x + \alpha_p + \alpha_p \alpha_B} < \]

\[ < \frac{\alpha_x \beta_x + \alpha_p \alpha_G}{\alpha_x + \alpha_p} = \beta_x(\rho(G) = x, \rho(B) = y) \]  

(20)

\[ \beta_y(\rho(G) = x, \rho(B) = y) = \frac{\alpha_y \beta_y}{\alpha_y + \alpha_p \alpha_B} = \frac{\beta_y}{1 + \frac{\alpha_p \alpha_B}{\alpha_y}} < \beta_y = \beta_y(\rho(G) = x, \rho(B) = x) \]  

(21)

\[ \beta_y(\rho(G) = x, \rho(B) = y) = \frac{\alpha_y \beta_y + \alpha_p \alpha_G}{\alpha_y + \alpha_p} = \frac{\alpha_y \beta_y + \alpha_p \alpha_G}{\alpha_y + \alpha_p + \alpha_p \alpha_B} = \beta_y(\rho(G) = y, \rho(B) = y) \]  

(22)

\[ \beta_y(\rho(G) = y, \rho(B) = y) = \frac{\alpha_y \beta_y + \alpha_p \alpha_G}{\alpha_y + \alpha_p} = \frac{\alpha_y \beta_y + \alpha_p \alpha_G}{\alpha_y + \alpha_p + \alpha_p \alpha_B} \]

\[ < \frac{\alpha_y \beta_y + \alpha_p \alpha_G}{\alpha_y + \alpha_p} = \beta_y(\rho(G) = y, \rho(B) = x) \]  

(23)

\[ \beta_y(\rho(G) = x, \rho(B) = x) = \beta_y = \beta_y \left( \frac{\alpha_y + \alpha_p \alpha_G}{\alpha_y + \alpha_p} \right) = \frac{\alpha_y \beta_y + \alpha_p \alpha_G \beta_y}{\alpha_y + \alpha_p \alpha_G} < \]

\[ < \frac{\alpha_y \beta_y + \alpha_p \alpha_G}{\alpha_y + \alpha_p \alpha_G} = \beta_y(\rho(G) = y, \rho(B) = x) \]  

(24)

We note that inequality (17) follows from \((\alpha_p \alpha_B) / \alpha_x > 0\), inequality (18) from \(\alpha_p \alpha_G > 0\) and (2) being strictly lower than 1, inequalities (19) and (24) from \(\alpha_p \alpha_G \beta_x < \alpha_p \alpha_G\); inequalities (20) and (23) from \(\alpha_p \alpha_B > 0\), inequality (21) from \((\alpha_p \alpha_B) / \alpha_y > 0\), and inequality (22) from \(\alpha_p \alpha_G > 0\) and (8) being strictly lower than 1.

**Proof.** From (9) follows that condition 6.1 is incompatible with conditions 4.1.1 and 4.2.1. Moreover, from (15) follows that condition 6.2 is incompatible with condition 4.3.2. This proves that a separating equilibrium where \(\rho(G) = y\) and \(\rho(B) = x\) and a pooling equilibrium where \(\rho(G) = \rho(B) = y\) cannot coexist.

From (12) follows that condition 5.1 is incompatible with condition 3.1. This proves that a separating equilibrium where \(\rho(G) = x\) and \(\rho(B) = y\) and a pooling equilibrium where \(\rho(G) = \rho(B) = x\) cannot coexist.
From (9) and (11) follows that $\hat{\beta}_x(\rho(G) = y, \rho(B) = x) < \hat{\beta}_x(\rho(G) = x, \rho(B) = y)$, which in turn implies that condition 6.1 is incompatible with condition 5.1. This proves that the two types of separating equilibria cannot coexist.

To see that the two types of pooling can coexist suppose that $|U_B|/(U_G + |U_B|) < 1$. If $\beta_x$ is close enough to 1 so that 4.3.1 is satisfied and that $\beta_y$ is close enough to 0 so that condition 3.2 is satisfied. For $\alpha_G$ large enough, also 3.1 is satisfied. To have also 4.3.2 satisfied it is enough to have $\alpha_P$ and $\alpha_G$ sufficiently large.

To see that a pooling equilibrium where $\rho(G) = \rho(B) = x$ and a separating equilibrium where $\rho(G) = y$ and $\rho(B) = x$ can coexist, note that from $\hat{\beta}_x(\rho(G) = y, \rho(B) = x) < \hat{\beta}_x(\rho(G) = x, \rho(B) = y)$ follows that condition 6.1 implies condition 3.1. Moreover, by (16) we can set $c_e$, $U_G$ and $U_B$ such that $U_G\hat{\beta}_y(\rho(G) = x, \rho(B) = x) \leq c_e \leq U_G\hat{\beta}_y(\rho(G) = y, \rho(B) = x)$, and $\hat{\beta}_y(\rho(G) = x, \rho(B) = x) \leq |U_B|/(U_G + |U_B|) \leq \hat{\beta}_y(\rho(G) = y, \rho(B) = x)$, so that both 6.2 and 3.2 are satisfied.

A similar argument can be applied to show that a pooling equilibrium where $\rho(G) = \rho(B) = y$ and a separating equilibrium where $\rho(G) = x$ and $\rho(B) = y$ can coexist.

B.3 Proof of Proposition 7

Proof. We observe that $\hat{\beta}_x(\rho(G), \rho(B))$ converges to $\beta_x$ and $\hat{\beta}_y(\rho(G), \rho(B))$ converges to $\beta_y$ irrespective of $\rho$ when the degree of coarse thinking grows larger and larger, i.e., $\chi$ tends to infinity. This is evident when looking at (1), (2), (3), (4), (5), (6), (7) and (8).

We assume that $\beta_x$ and $\beta_y$ are interior points to the intervals of beliefs that determine DM’s best choices according to Proposition 1. More precisely, $\beta_x$ and $\beta_y$ are both different from $c_e/U_G$ and $1 - c_e/|U_B|$ if $c_e < U_G|U_B|/(U_G + |U_B|)$, and different from $|U_B|/(U_G + |U_B|)$ if $c_e \geq U_G|U_B|/(U_G + |U_B|)$. By so doing we are neglecting values that have measure zero in the parameters space.

We now build a profile that we then check to be an equilibrium. We set $\delta(x)$ and $\delta(y)$ equal to the best action by DM against a belief equal to $\beta_x$ and $\beta_y$, respectively, as show by Proposition 1. We set $\rho(G)$ and $\rho(B)$ equal to the best action by P conditional on $G$ and $B$, respectively, against $\delta(x)$ and $\delta(y)$, as shown by Proposition 2.

By construction, in the above profile P is best replying to $\delta$, while DM is best replying given $\beta_x$ and $\beta_y$, which are not equilibrium beliefs. However, we can choose $\chi$ high enough that, whatever $\rho$ is chosen by P, $\hat{\beta}_x(\rho(G), \rho(B))$ and $\hat{\beta}_y(\rho(G), \rho(B))$ are very close to $\beta_x$ and $\beta_y$, respectively. This means that DM is best replying even against $\hat{\beta}_x(\rho(G), \rho(B))$ and $\hat{\beta}_y(\rho(G), \rho(B))$, since $\beta_x$ and $\beta_y$ are interior points to the intervals of beliefs that determine DM’s best choices.

We have just proved equilibrium existence. To understand that such equilibrium is unique, we simply observe that, when $\chi$ is large enough, the best reply by DM is uniquely determined whatever strategy is chosen by P, and P’s best reply against such an optimal behavior by DM is uniquely determined as well.