Persuasion with Correlation Neglect

Gilat Levy, Inés Moreno de Barreda, Ronny Razin*

December 11, 2018

Abstract

We consider an information design problem in which a sender tries to persuade a receiver that has "correlation neglect", i.e., fails to understand that signals might be correlated. We show that the sender can change the expected posterior of the receiver in any direction. When the number of signals the sender can send is large, she can approach her first best utility. We characterize for which environments full correlation is the optimal solution; in these cases we can use a modified problem and standard concavification techniques. We show that full correlation is optimal in the familiar case of binary utilities but also more generally when utilities are super-modular and when the number of signals is large. However, in some environments full correlation is not optimal and in those cases the optimal solution involves negative correlation.

*Levy: London school of Economics; Moreno de Barreda: University of Oxford; Razin: London school of Economics. We thank Clement Minaudier for excellent research assistance. This project has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No SEC-C413.

1 Introduction

In this day and age we are constantly exposed to many different information sources. For example, as voters we are exposed to news from different media sources offline as well as online and through social media. To top that, day to day interactions with family, friends and colleagues at work expose us to even more information. We probably understand each information source separately better than we understand the complex ways in which these sources are inter-related.

The growth in the amount of information we are exposed to and its complexity, opens up opportunities for those that try to manipulate the beliefs of voters or consumers. In the last two years, more and more instances of orchestrated attempts to feed information into social media are exposed. To give one example, in the aftermath of the Brexit vote in the UK in June 2016, a key concern is a possible breach of electoral laws on behalf of the pro-Brexit side in terms of its use of individual data; specifically the UK Information Commissioner's Office (ICO) has started investigating the relationship between the Canadian data firm AggregateIQ, Vote Leave and a number of other leave campaigns in using the same data handed over by Facebook in May 2018. The ICO's initial report says: "AIQ created and, in some cases, placed advertisements on behalf of the DUP Vote to Leave campaign, Vote Leave, BeLeave, and Veterans for Britain. AIQ ran 218 ads solely on behalf of Vote Leave and directed at email addresses on Facebook. Vote Leave and BeLeave used the same data set to identify audiences and select targeting criteria for ads[...] Our regulatory concern is therefore whether, and on what basis, the two groups have shared data between themselves and others." While the ICO is only able to investigate breaches of data protection law, the Electoral Commission in the UK is separately investigating allegations of unlawful coordination between Vote Leave (the official pro-Brexit campaign during the EU referendum) and BeLeave.

In this paper we investigate how coordination across information sources -when unknown to voters and consumers- can be used strategically to affects their beliefs. While, as described above, campaigns and information sources may be correlated, voters may be unaware that this is the case. Indeed a recent empirical literature documents in different environments how people exhibit "correlation neglect" and how this can lead to extremism. Formally, correlation neglect means that individuals assume that information sources are (conditionally) independent.[1] Ortoleva and Snowberg (2015) documents how correlation neglect shapes political views, Eyster and Weizsäker (2011), Kallir and Sonsino (2009) and Enke and Zimmermann (2018) provides experimental evidence for such behaviours.[2]

Given this widespread inability to understand correlation, those that control several information sources may find ways to manipulate voters or consumers beliefs: Media owners can correlate the messages they wish to send across their different news outlets, campaign

[1] De Marzo et al. (2003), Golub and Jackson (2012) and Gagnon-Bartsch and Rabin (2015) study how correlation neglect affects the diffusion of information in social networks. Glaeser and Sunstein (2009) and Levy and Razin (2015a;b) explore the implications for group decision making in political applications.

[2] On the theory side, Ellis and Piccione (2017) provide an axiomatic characterization of individuals that cannot account for correlation. Levy and Razin (2018) discuss environments for which correlation neglect is more likely.

and thinktank managers can coordinate advertisements in a manner that is hard to detect even for savvy regulators, and editors of news aggregation sites can carefully pick the links they present. Even those who do not have such immediate access to multiple news outlets can still take advantage of individuals neglecting correlation. As Cagé et al. (2017) show, most online content is not original and is quickly repackaged and repeated.

To be sure, offline communication can yield similar situations. A lobbyist who may wish to influence a committee may coordinate a campaign in which individual members are targeted with different reports from a variety of thinktanks or scientific institutions. Committee members that exchange information among themselves may be unaware that the sources of their information were coordinated. For example, they may not be aware that the same donor gives money to all these different thinktanks. Thus, more generally strategic senders may be in a position to manipulate the correlation neglect of receivers.[3]

We analyze a model of persuasion when the receiver has correlation neglect. We use an information design model in which a sender, who cares about the action of the receiver as well as about the state of the world, can design (and commit to) a joint information structure for m experiments or signals, as a function of the state of the world. A receiver who attempts to learn the state of the world will observe the realizations of these m signals. We assume that the receiver understands the marginal distribution of each signal, so that she can interpret each of them separately. But, we also assume that she believes that the signals are all conditionally independent. Given her updated beliefs, she takes her optimal action. We then analyze the scope for manipulation of the receiver's beliefs/actions.

We show that the sender can manipulate the receiver's posteriors. Specifically, the sender can move the expected posterior of the receiver away from the prior. This is intuitive as the receiver's beliefs do not satisfy the martingale property when the sender correlates signals. What we show is that the sender can move these beliefs in any direction she wishes to, and that to do so, it is sufficient to use just two signals. Moreover, we construct this type of manipulation by fully correlating these two signals.

This ability of the sender to manipulate the receiver is, however, bounded. Whenever the receiver has only a finite number of signals there are strict bounds on how the sender can manipulate her posterior beliefs. This bound arises as the receiver is Bayesian when it comes to interpreting each signal separately. This imposes some limits on the ability of the sender to manipulate her beliefs.

When the number of signals grows large, the above constraints on manipulation are relaxed. We show that in the limit, the sender can manipulate the receiver in the most extreme way possible. In particular, the sender can induce any posterior beliefs that she wishes to, where these induced posteriors may differ across states. Thus, the sender can fully manipulate the receiver and achieve her first best utility.

Correlating signals or posteriors presents additional constraints on the problems con-

[3]One might recall the story of Solomon's Succession of King David as described in the Bible, 1 Kings 1. In the story, Nathan the Prophet and Bathsheva, King David's wife, want to make sure King David follows up on his promise to anoint Solomon as King. Nathan comes up with a plan by which Bathsheva will speak to King David about the matter, and "Then while you are still there speaking with the king, I will come in after you and confirm your words." The plan succeeds and King David anoints Solomon as king of Israel and Judah.

sidered for example in Aumann and Maschler (1995) and Kamenica and Gentzkow (2011). Specifically, standard concavification techniques will indeed imply an upper bound on what the sender can gain, but this upper bound is not necessarily achieved: Concavification may imply correlation patterns that cannot be obtained. For example, one cannot negatively correlate two fully informative signals. However, if the optimal concavification solution can be achieved with full positive correlation across signals, then it is always feasible.

Motivated by the above, we present and analyze a modified problem. In the modified problem we redefine the utility of the sender to depend on one signal whose realisation is fully correlated (repeated) m times. Specifically, if the optimal solution to the general problem is full correlation, then the problem is isomorphic to the modified problem in which a sender designs just one signal whose realisation the receiver observes m times. We can then use a simple modified concavification argument to design just such one signal.

We show that for the case of state-independent utilities, the solution to the modified problem converges to the optimal solution when m is large. Even if m is small, this procedure will characterize the optimal solution in a large family of utilities that includes the family of supermodular utilities but also the simple canonical binary example that we analyze in Section 4.1. In other words, the sender should use full correlation as her optimal information structure in these cases.

However, full correlation is not always optimal. For example, consider the case in which it is optimal to induce a posterior which is degenerate on one state. In this case it is wasteful to fully correlate signals, each inducing with some probability this degenerate posterior. Alternatively, by negatively correlating this degenerate posterior with interior ones, the receiver will still have a degenerate posterior when factoring in both signals while being able to use the interior posteriors in her benefit. More generally we show that submodularity will imply some negative correlation.

Our paper relates to the growing literature on persuasion, stemming from Aumann and Maschler (1995) and Kamenica and Gentzkow (2011). In Kamenica and Gentzkow (2011), it is sufficient to use one signal when facing a rational receiver. To construct a meaningful model of correlation neglect, we have to use more than one signal. We show that in some environments our problem can be translated to a standard persuasion problem with a modified utility function, and then the same concavification techniques can be applied. In some cases however, concavification that imposes Bayesian plausibility constraint on each signal is not sufficient to identify the solution. Alonso and Câmara (2016b) also analyze persuasion under behavioural assumptions, specifically that of a wrong prior for the receiver. Koessler et al. (2018) and Arieli and Babichenko (2016) analyze a multiple (rational) receiver persuasion problem. Our problem can be interpreted as a sender sending private or public messages to m receivers who share the information among themselves and believe that their information sources are independent. Meyer (2017) considers a general model which can be applied to an information design problem with many receivers. In her model the sender chooses a joint distribution to maximize an expected value function subject to *fixed* marginals distributions. For the case of binary states, she shows that when the value function is symmetric, the maximized value decreases as the marginals become more heterogeneous. Finally, in Levy et al. (2018), we consider a similar model that is applied to media markets where we consider a binary state space, binary signals, and competition among senders.

Our analysis relates more generally to the literature on misperceived or misspecified models. In some sense misperception arises endogenously in our model. While we assume that the receiver always interprets different signals as conditionally independent, this is a misspecified model only if the sender chooses to correlate her signals. Indeed, as we show, the optimal solution involves a joint information structure with correlation and so the receiver's model is misperceived. In other recent contributions, Eliaz and Spiegler (2018), Schumacher and Thysen (2018), and Ellis et al. (2018) characterize how individuals with misperceived model can be strategically manipulated.

The remainder of the paper is as follows. In Section 2 we construct the model and define the sender's problem. In Section 3 we present our general results about belief manipulation, and show how the sender can achieve her first best utility in the limit. Section 4 provides the results on the design of the optimal information structure; specifically, we analyze first the case of binary utilities to illustrate our main results and we present the modified problem where the optimal information structure exhibits full (positive) correlation. In Section 5 we consider extensions that restrict the ability of the sender to freely manipulate, where we allow the receiver to understand some correlation and also consider the case of limited attention. The Appendix includes all proofs.

2 The Model and Preliminary Results

There is a finite state space $\Omega = \{\omega_1, ..., \omega_n\}$ with a full support prior probability distribution $p \in \Delta(\Omega)$ which is common knowledge. The sender designs an information structure that consists of m distinct signals. The receiver observes the signal realisations $\boldsymbol{s} = (s_1, ..., s_m)$ and chooses an action a out of a compact set $A \subset \mathbb{R}$. Given action a and state ω, the receiver gets utility $u(a, \omega)$ and the sender gets $v(a, \omega)$.

Information Structures:
An information structure with m distinct signals or experiments is defined by $\{\boldsymbol{S}, \{q(\cdot \mid \omega)\}_{\omega\in\Omega}\}$ where $\boldsymbol{S} = \prod_{i=1}^{m} S_i$ is the cartesian product of the support of the individual signals, and $q(\cdot \mid \omega) \in \Delta(\boldsymbol{S})$ is the joint probability distribution conditional on $\omega \in \Omega$. For such information structure, let $\{S_i, \{q_i(\cdot \mid \omega)\}_{\omega\in\Omega}\}$ denote the marginal information structure for signal i derived from $\{\boldsymbol{S}, \{q(\cdot \mid \omega)\}_{\omega\in\Omega}\}$.[4] We assume S_i is finite for all $i \in \{1, ..., m\}$.[5]

As in the Bayesian Persuasion literature it will be more convenient to work with distributions over posteriors rather than signal distributions. Given a realization s_i of signal i, define the posterior induced by s_i to be,

$$\mu_{s_i}(\omega) = \frac{p(\omega)q_i(s_i \mid \omega)}{\sum_{\upsilon} p(\upsilon)q_i(s_i \mid \upsilon)} \qquad \text{for all } \omega \in \Omega.$$

[4]The marginal conditional distributions are given by $q_i(s_i \mid \omega) = \sum_{s_{-i}\in S_{-i}} q(s_i, s_{-i} \mid \omega)$ for all $s_i \in S_i$ and $\omega \in \Omega$.
[5]See Koessler et al. (2018) for a model with competing senders using infinite message sets.

The joint distribution of the signals $\{q(\cdot \mid \omega)\}_{\omega \in \Omega}$ induces a collection of conditional distributions, $\{\tau(\cdot \mid \omega)\}_{\omega\in\Omega} \subset \Delta(\Delta(\Omega)^m)$, over vectors of posteriors, such that

$$\tau((\mu_{s_1}, ..., \mu_{s_m}) \mid \omega) = q((s_1, ..., s_m) \mid \omega) \quad \text{for all } \omega \in \Omega.$$

Kamenica and Gentzkow (2011) characterize the set of *unconditional* distributions over posteriors that can be induced by a signal structure, along with a Bayesian updater. Such a set corresponds to the distributions that satisfy *Bayesian plausibility*, i.e., such that the expected posterior coincides with the prior distribution. As the sender might want to correlate signals depending on the realisation of the state, it will be important for us to characterize the set of *conditional* distributions over posteriors that can be generated by a signal with a Bayesian updater. In particular, some correlation structures might not be feasible even though they generate unconditional distributions that satisfy Bayes plausibility. A simple example is provided below.

Example 1. *Consider $\Omega = \{0, 1\}$ and $p(0) = p(1) = \frac{1}{2}$ and denote by μ_1 the degenerate posterior on $\omega = 1$ and μ_0 the degenerate posterior on $\omega = 0$. Consider two fully revealing signals. Each signal generates posteriors μ_1 and μ_0 with probability $\frac{1}{2}$ which is Bayes-plausible, however, it is not possible to negatively correlate them because conditional on the realization of the state only one of the two posteriors can be induced.*

Given the joint conditional distributions $\{\tau(\cdot \mid \omega)\}_{\omega\in\Omega}$ we denote by $\{\tau_i(\cdot \mid \omega)\}_{\omega\in\Omega} \subset \Delta(\Delta(\Omega))$ the marginal conditional distributions over posteriors corresponding to the i'th signal and $\tau_i(\cdot) = \sum_{\omega\in\Omega} p(\omega)\tau_i(\cdot \mid \omega) \in \Delta(\Delta(\Omega))$ the corresponding marginal unconditional distribution. The following lemma characterises the families of joint conditional distributions that can be induced by m signals:

Lemma 1. *The joint conditional distributions $\{\tau(\cdot \mid \omega)\}_{\omega\in\Omega}$ are inducible by m signals, given a Bayesian updater, if and only if for any $\mu \in \Delta(\Omega)$ and any $i \in \{1, ..., m\}$,*

$$\tau_i(\mu \mid \omega) = \frac{\tau_i(\mu)\mu(\omega)}{p(\omega)} \quad \textit{for all } \omega \in \Omega \tag{1}$$

Note that if $\{\tau_i(\cdot \mid \omega)\}_{\omega\in\Omega}$ satisfies condition (1), then the unconditional distribution $\tau_i(\cdot)$ satisfies Bayesian plausibility, but the reverse is not true.[6]

In the rest of the paper we abstract from explicit signal structures and work with its corresponding conditional distributions over posteriors.

Correlation Neglect Assumption:

We assume that the receiver observes the marginal information structure for every signal

[6]To see this, suppose that $\{\tau_i(\cdot \mid \omega)\}_{\omega\in\Omega}$ satisfies condition (1) then,

$$\sum_{\mu} \tau_i(\mu)\mu(\omega) = \sum_{\mu} p(\omega)\tau_i(\mu \mid \omega) = p(\omega) \sum_{\mu} \tau_i(\mu \mid \omega) = p(\omega).$$

However, consider the following conditional distributions in the setting of Example 1: $\tau(\mu_1 \mid \omega = 0) = 1 = \tau(\mu_0 \mid \omega = 1)$. These conditional distributions do not satisfy condition (1) even though the unconditional distribution $\tau(\mu_1) = \tau(\mu_0) = \frac{1}{2}$ is Bayes-plausible.

i, but she believes that the signals are conditionally independent. The receiver therefore correctly interprets each signal in isolation but ignores the potential correlation between the signals. In Section 5.2 we extend the model and our results to consider a receiver who can understand some degree of correlation.

We now describe how a receiver with correlation neglect updates her beliefs in term of posterior distributions. Given m signals with conditional distributions $\{\tau_i(\cdot \mid \omega)\}_{\omega\in\Omega}$ for $i = 1, ..., m$, and a vector of posteriors $\boldsymbol{\mu} = (\mu_1, ..., \mu_i, ..., \mu_m)$ where μ_i is the posterior generated by signal i, we denote by $\mu^{CN}(\boldsymbol{\mu})$ the posterior that a receiver with correlation neglect forms. The following lemma characterises $\mu^{CN}(\boldsymbol{\mu})$.[7]

Lemma 2. *Given a prior p and a vector of posteriors $\boldsymbol{\mu} = (\mu_1, ..., \mu_m)$, the posterior belief of a receiver with correlation neglect is:*

$$\mu^{CN}(\omega \mid \boldsymbol{\mu}) = \frac{\frac{\prod_{i=1}^{m} \mu_i(\omega)}{p(\omega)^{m-1}}}{\sum_{\upsilon\in\Omega} \frac{\prod_{i=1}^{m} \mu_i(\upsilon)}{p(\upsilon)^{m-1}}}$$

Note that if the sender sends only one signal, $m = 1$, then $\mu^{CN}(\mu) = \mu$. In such case, the receiver acts as a rational agent.

Once the receiver updates her beliefs, she chooses the optimal action given those beliefs. Given a posterior $\mu \in \Delta(\Omega)$, we denote by a_μ the receiver's optimal choice:

$$a_\mu \in \arg\max_{a\in A} \sum_{\omega\in\Omega} \mu(\omega)u(a, \omega).$$

In case of multiple solutions we will assume the receiver chooses the sender-preferred solution in order to guarantee existence of the sender's solution.

It will be useful for comparisons to define the *First Best* benchmark for the sender. For any $\omega \in \Omega$, define by $\mu^{FB}_\omega = \arg\max_\mu v(a_\mu, \omega)$. This is the posterior that would generate the optimal action for the sender at state ω.

The Sender's Problem:

Given a joint information structure τ, with conditional distributions $\{\tau(\cdot \mid \omega)\}_{\omega\in\Omega}$, we denote by $\hat{v}(\tau)$ the expected utility for the sender:

$$\hat{v}(\tau) = \sum_{\omega\in\Omega} p(\omega) \sum_{\boldsymbol{\mu}\in\Delta(\Omega)^m} \tau(\boldsymbol{\mu} \mid \omega)v(a_{\mu^{CN}(\boldsymbol{\mu})}, \omega)$$

[7]This result exists in Sobel (2014), Proposition 5. The multiplicative form also arises in axiomatic approaches to combining forecasts, see Morris (1977) and Bordley (1982).

The sender's problem is to choose $\{\tau(\cdot \mid \omega)\}_{\omega\in\Omega} \subset \Delta(\Delta(\Omega)^m)$ to solve:

$$\begin{aligned} &\max_\tau \ \hat{v}(\tau) \\ s.t. \quad &\mu^{CN}(\omega \mid \boldsymbol{\mu}) = \frac{\frac{\prod_{i=1}^{m}\mu_i(\omega)}{p(\omega)^{m-1}}}{\sum_{\upsilon\in\Omega}\frac{\prod_{i=1}^{m}\mu_i(\upsilon)}{p(\upsilon)^{m-1}}} \\ and \quad &\mu_i(\omega) = \frac{p(\omega)\tau_i(\mu_i|\omega)}{\tau_i(\mu_i)} \quad \forall \mu_i \text{ s.t. } \tau_i(\mu_i) > 0, \quad \forall i \in \{1, ..., m\} \end{aligned}$$

The first constraint corresponds to the updating of the receiver with correlation neglect and the second constraint requires all the marginal conditional distributions to be inducible by a signal structure.

Existence of a solution is guaranteed by the selection of the sender-preferred action. This implies that $\hat{v}(\tau)$ is upper semicontinuous and combined with the compactness of the set of all joint distributions this implies the existence of a maximum.

Remark 1. *Note that the expected utility of the sender, $\hat{v}(\tau)$, cannot be written as a function of her belief only, or the receiver's beliefs only. Given some joint information structure that does not satisfy independence across the different signals, the sender's and the receiver's beliefs do not coincide, and in general there is no fixed mapping from the sender's beliefs to that of the receiver that does not depend on the specific signal structure used. Hence standard concavification arguments will in general not be applicable.*[8]

If the receiver was rational, the sender would have no interest in increasing the number of signals she has at her disposal: as Kamenica and Gentzkow (2011) show, any posterior distribution that can be generated with multiple signals could also be generated with only one signal. Therefore the case of $m = 1$ corresponds to the case of a rational receiver. In contrast, when the receiver has correlation neglect, adding more signals does increase the set of distributions over posteriors that can be generated.

Before turning to the next section we introduce a preliminary result that simplifies the analysis of the paper. It states that we can, without loss of generality, focus on signal structures that use homogeneous signals, i.e., signals with the same marginal distributions.

Lemma 3. *For any joint information structure τ, and its associated expected utility for the sender $\hat{v}(\tau)$, there exists a joint information structure τ' with homogenous marginals, i.e., $\tau'_i = \tau'_j$, for all $i, j \in \{1, ..m\}$, such that $\hat{v}(\tau') = \hat{v}(\tau)$.*

The result relies on the fact that the receiver's updating rule $\mu^{CN}(\cdot)$ is symmetric in $(\mu_1, ..., \mu_m)$ and hence the *induced* utility of the sender, $v(a_{\mu^{CN}(\mu)}, \omega)$, is also symmetric in the different m individual posteriors she generates. Hence, any expected utility $\hat{v}(\tau)$ for some τ can be replicated by permutating the marginal distributions of the m signals.

[8]This differs from the analysis of a behavioural receiver in Alonso and Câmara (2016a), in which even though the sender and the receiver have different priors there is a bijection between the sender's and the receiver's posteriors.

Considering a joint information structure which is an average of all possible permutations will achieve this utility and will have homogeneous marginal signals. Henceforth we will focus on homogeneous signals.

3 The Scope for Manipulation

This section presents some general results on the scope of manipulation given correlation neglect. The first main result shows that by using just two signals and fully correlating them, the sender is able to move the expected posterior of the receiver in any direction she might want.

Proposition 1. *For any distribution $q \in \Delta(\Omega)$, there exists $\epsilon > 0$ and a signal structure τ with two fully correlated signals, such that:*

$$E_\tau(\mu^{CN}(\boldsymbol{\mu})) = (1-\epsilon)p + \epsilon q$$

While the fact that the expected posterior does not equal the prior is not surprising given the receiver's wrong beliefs, what we show in the proof is that the expected posterior can be moved in *any* direction away from p.

To provide the intuition behind the proof let us consider the case of binary states. Figure 1 illustrates how a receiver with correlation neglect forms her posterior after receiving a particular signal realization twice. The horizontal axis corresponds to the posterior (probability that the state is one) when a single signal is sent. The vertical axis corresponds to the correlation neglect posterior given that two identical realizations have been received.

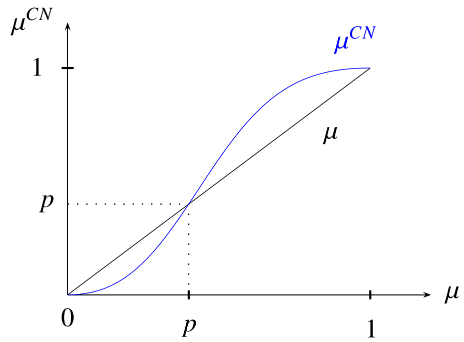
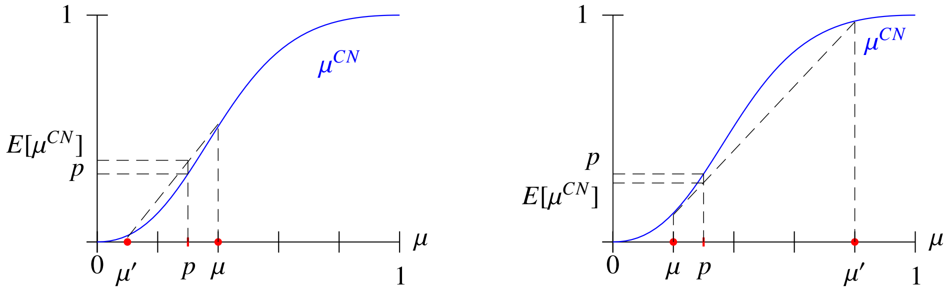


Figure 1: Correlation neglect posterior after 2 fully correlated signals.

There are two characteristics of the receiver's updating rule that are worth highlighting. First there is an amplification effect: a receiver that gets the same realization twice and thinks they are independent will have a higher posterior if the individual posterior was above the prior, and a lower posterior if the individual posterior was below the prior. Second the updating is convex below the prior and concave above it. In other words, the marginal value of information is decreasing in the sense that after the first piece of good (bad) news, the second piece of good (bad) news moves the posterior by a lower increment. The sender will be able to exploit this non-linearity of the updating rule to move the expected posterior in any desired direction. For instance suppose that the sender wants to induce an expected posterior that is above the prior. Consider a signal that with high probability induces a posterior slightly above the prior and with small probability induces a posterior below and further away from the prior (see Figure 2a). By repeating this signal twice the sender will be inducing more extreme posteriors with exactly the same probabilities as before. Given the concavity and convexity of the updating rule, this will lead to a higher expected signal. A symmetric argument could be used to generate a expected posterior below the prior (see Figure 2b).



(a) Expected posterior is above the prior

(b) Expected posterior is below the prior

Figure 2: Changing the expected posterior

The proof of Proposition 1 generalizes this intuition to finite state spaces.

Our second result shows that the scope of manipulation is bounded, i.e., the sender will not be able to induce a posterior that is far from the prior with probability approaching to one. Define by $M_\omega(\delta)$ the set of vectors of posteriors in the support of τ that lead to the receiver holding a belief that puts weight strictly higher than $1-\delta$ in state ω:

$$M_\omega(\delta) = \{\boldsymbol{\mu} \in \text{Support}(\tau) \mid \mu^{CN}(\omega \mid \boldsymbol{\mu}) > 1-\delta\}$$

Proposition 2 bounds the probability that τ allocates to $M_\omega(\delta)$.

Proposition 2. *For any $\omega, \upsilon \in \Omega$ with $\omega \neq \upsilon$, and $\delta > 0$,*

$$\tau(M_\omega(\delta) \mid \upsilon) \leq m\frac{1}{p(\upsilon)}\left(\frac{\delta}{1-\delta}\right)^{\frac{1}{m}}\left(\frac{p(\upsilon)}{p(\omega)}\right)^{\frac{m-1}{m}}$$

In particular,

$$\lim_{\delta\to 0} \tau(M_\omega(\delta)) \leq p(\omega)$$

The bound on manipulation comes from the fact that the receiver understands individual signals and hence the marginal distributions constrain the sender on her ability to persuade. For example, the sender will not be able to convince the receiver that a particular state has been realized when it has not.

Note that for any fixed $\delta > 0$, the bound on manipulation becomes more and more lenient as the number of signals increases. In fact, as the next result shows, in the limit when the number of signals increases the sender can induce the receiver to have any state-dependent posterior she might wished to.

Theorem 1. *When $v(a, \omega)$ is continuous, the sender can achieve her first best in the limit. That is, given $\{q^\omega\}_{\omega \in \Omega} \subset \Delta(\Omega)$, there exists a sequence of signal structures $\{S^m\}_{m \in \mathbb{N}}$ indexed by the number of signals, such that for any $\omega \in \Omega$, and any $\epsilon > 0$,*

$$\lim_{m \to \infty} \tau^{S^m}(\{\boldsymbol{\mu} \in Support(\tau) \mid |\mu^{CN}(\boldsymbol{\mu}) - q^\omega| < \epsilon\} \mid \omega) = 1$$

Theorem 1 states that the sender can approach her first-best, μ_ω^{FB}, at any state ω and thus manipulation is complete in the limit. The proof of Theorem 1 is constructive. The information structure uses signals with $n = |\Omega|$ realizations. Each realisation is indicative of a particular state, and the sender can correlate the realisations in a way that approaches μ_ω^{FB} in the limit.

We have shown above the unlimited scope for persuasion in the limit when $m \to \infty$; in this case the sender can achieve her first-best utility, for a general state space and for general state-dependent utilities. Away from the limit, the scope for persuasion is typically derived by using concavification techniques, which we now discuss:

Remark 2 (**Concavification and the upper bound on manipulation**)**.** *As stated in Remark 1, for state-dependent utilities, there is no a one-to-one relationship between the posteriors of the sender and the receiver that is independent of the joint distribution τ and hence the standard concavifiaction arguments are not applicable. For state-independent preferences however, the sender's expected payoff does not depend directly on the sender's posterior. Thus, given a vector of posteriors $\boldsymbol{\mu} = (\mu_1, ..., \mu_m)$, the induced utility for the sender is $\tilde{v}(\boldsymbol{\mu}) = v(a_{\mu^{CN}(\boldsymbol{\mu})})$ and we can write the expected utility of the sender as*

$$\hat{v}(\tau) = \sum_{\boldsymbol{\mu} \in \Delta(\Omega)^m} \tau(\boldsymbol{\mu}) \tilde{v}(\boldsymbol{\mu}).$$

Denote by $V(\boldsymbol{\mu})$ the concavification of $\tilde{v}(\cdot)$. It is easy to see then that for any information structure $\tau(\cdot)$,

$$\hat{v}(\tau) \leq V(\boldsymbol{p})$$

In other words, while concavifiation produces an upper bound, this upper bound $V(\boldsymbol{p})$ might not be reached. Specifically, concavification may yield correlation patterns that are not compatible with the characterisation of feasible information structures given in Lemma 1.

The remark above illustrates that concavification presents the upper bound on the sender's utility for the case of state-independent preferences. Note that when concavification prescribes full positive correlation across signals, the upper bound *will* be achieved. Full correlation across signals is always a feasible correlation structure. Specifically, it implies the design of one signal, subject to only the Bayesian plausibility constraint, and then repeating the realisation m times. We next explore the scope of manipulation when the sender uses full correlation.

4 The Modified Problem: Full Correlation

In this section we simplify the analysis by focusing on a modified problem that involves the sender designing just one signal and then fully correlating it m times: that is repeating m times each realization. We first consider the simple binary utility example to illustrate our results. We show that for this case full correlation is optimal. Building on the binary example, we construct the modified problem and show more generally under what environments the solution to this problem is also the optimal solution to the general problem. Throughout this Section we focus on state-independent utilities, given Remark 2 above.

Note that the modified problem is also interesting in its own right: There may be environments in which the sender is not able to control many information sources, but she may be aware that signals she transmits are repeated, as shown by Cagé et al. (2017). In that case, full correlation as defined in the modified problem is the only possible course of action for the sender.

4.1 Binary Utilities

Consider $\Omega = \{0, 1\}$ with prior $p = \Pr(\omega = 1) < \frac{1}{2}$. The receiver takes one of two actions, $a = 1$ if $\mu^{CN}(1 \mid \boldsymbol{\mu}) \geq \frac{1}{2}$ and $a = 0$ otherwise. In particular, given the prior, the receiver's default action is 0. Suppose that the sender has some state-independent preferences such that she prefers $a = 1$ to $a = 0$. This is the familiar binary utility model. Below, when we say the posterior of the receiver we mean her belief that the state is $\omega = 1$.

Consider first the problem with one signal, $m = 1$. In this case the receiver behaves as if she is rational. Figure 3 shows the solution; on the $x - axis$ we have the posterior of the receiver and on the $y - axis$ the utility of the sender as a function of the receiver's posterior.

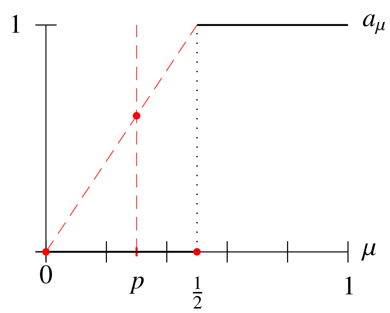


Figure 3: Binary Utilities: 1 signal

Concavification implies that the solution involves two posteriors, $\mu = \frac{1}{2}$ and $\mu = 0$. Bayesian plausibility (i.e., $E_\tau(\mu) = p$) implies that $\frac{1}{2}\tau(1/2) = p$ and hence $\tau(0) = 1 - 2p$. Using Lemma 1, we have

$$\tau(1/2 \mid 1) = 1, \quad \tau(1/2 \mid 0) = \frac{p}{1-p}, \quad \tau(0 \mid 0) = \frac{1-2p}{1-p}$$

Let us now illustrate what $\mu_m^{CN}(\mu)$ looks like when m signals are repeated (fully correlated). Figure 4 has on the $x - axis$ the posterior μ of the receiver and on the $y - axis$ we have

$\mu_m^{CN}(\mu) \equiv \frac{\frac{\mu^m}{p^{m-1}}}{\frac{\mu^m}{p^{m-1}}+\frac{(1-\mu)^m}{(1-p)^{m-1}}}$, the belief of the receiver that the state is $\omega = 1$ when she receives m replications of the posterior μ. As already mentioned in the discussion of Proposition 1, for $m > 1$, $\mu_m^{CN}(\mu)$ is higher than μ for $\mu > p$ and and lower that μ when $\mu < p$. Moreover, $\mu_m^{CN}(\cdot)$ is concave for $\mu > p$ and convex when $\mu < p$. In other words there is a "stretching" effect.

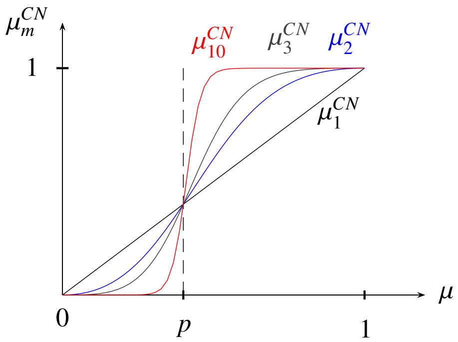


Figure 4: Correlation neglect posterior after m fully correlated signals.

It is now easy to see that already with two, fully correlated, signals, the sender can increase her utility. Specifically, the posterior $\mu_2^{CN}(\mu) = \frac{1}{2}$ can be induced by a twice-repeated $\mu^* = \frac{p^{\frac{1}{2}}}{p^{\frac{1}{2}}+(1-p)^{\frac{1}{2}}} < \frac{1}{2}$. As a result, given the Bayesian plausibility restriction, this posterior can be induced with a higher probability in state 0, specifically:

$$\tau(\mu^*|1) = 1,\ \tau(\mu^*|0) = \left(\frac{p}{1-p}\right)^{\frac{1}{2}},\ \tau(0|0) = 1 - \left(\frac{p}{1-p}\right)^{\frac{1}{2}}$$

This then leads to a higher expected posterior for the receiver, as now $E_\tau(\mu_2^{CN}(\mu)) > p$. In the left panel of Figure 5 we illustrate this result by drawing the modified utility function for the sender. Specifically, the sender gains whenever $\frac{\frac{\mu^2}{p}}{\frac{\mu^2}{p}+\frac{(1-\mu)^2}{(1-p)}} \geq \frac{1}{2}$.

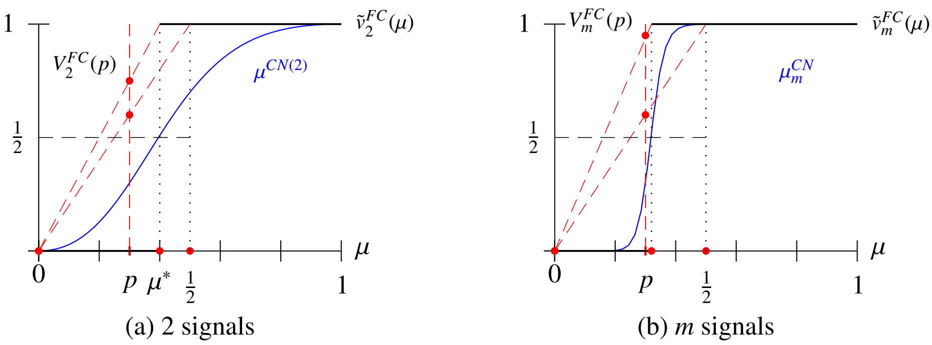


Figure 5: Binary utilities

The right panel of Figure 5 illustrates that the sender can achieve a higher utility when the number of signals increases. Specifically, we solve the following problem. The sender has to design just one signal, that satisfies Bayes-plausibility, given her modified utility

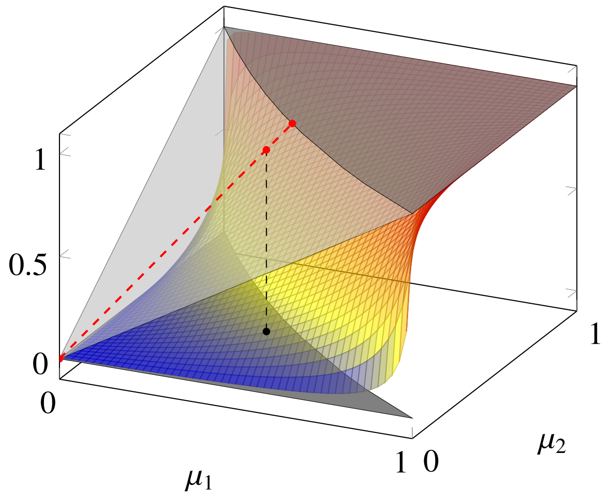


Figure 6: Concavification in the Binary States-Binary Action Model. (p=0.4)

function, which implies that she receives 1 iff

$$\frac{\frac{\mu^m}{p^{m-1}}}{\frac{\mu^m}{p^{m-1}} + \frac{(1-\mu)^m}{(1-p)^{m-1}}} \geq \frac{1}{2}$$

and 0 otherwise. Using this modified problem, we can see that in the limit, as the number of signals go to infinity, the sender's utility converges to her first best; she can induce a posterior, $\mu = \frac{p^{\frac{m-1}{m}}}{p^{\frac{m-1}{m}}+(1-p)^{\frac{m-1}{m}}}$, which is larger than but arbitrarily close to p, with probability converging to 1. Specifically, the signal she uses will have a distribution of $\tau(\mu \mid 1) = 1$, $\tau(\mu \mid 0) = \left(\frac{p}{1-p}\right)^{\frac{1}{m}}$ and $\tau(0 \mid 0) = 1 - \left(\frac{p}{1-p}\right)^{\frac{1}{m}}$. Therefore, as $m \to \infty$ will induce $\tau(\mu) \to 1$ and the sender induces action $a = 1$ with probability 1.

Finally, in order to see whether using full correlation is optimal, we compute the concavification of $\tilde{v}(\cdot)$. The case of two signals is plotted in Figure 6. By drawing that the concavification of $\tilde{v}(\cdot)$, we can compute the upper bound on the sender's expected utility specified in Remark 2. Noting that each signal has to satisfy Bayesian plausibility and hence the utility is evaluated at the prior (p, p), we see that the solution to this concavification can be achieved by mixing between two vectors of posteriors that are in the diagonal and hence correspond to fully correlated signals. Furthermore, as the solution of concavification yields full correlation, and full correlation is always feasible as long as the Bayesian plausibility constraint is satisfied, the upper bound is reached and this leads to the maximum expected utility for the sender.

The binary utility example had illustrate our key results from Section 3; the sender can gain more when $m > 1$, manipulation is bounded for any given m, but full manipulation is achieved when $m \to \infty$. In addition, we had shown that full correlation is sometimes an optimal solution, and in this case it is sufficient to consider the modified persuasion problem, where we consider the utility of the sender as depending on $\mu^{CN}(\mu, ..., \mu)$.

4.2 Optimality of Full Correlation

In the above example we saw that when the sender uses full correlation, it is equivalent to modifying her utility and using concavification to design just one signal. We will now show more generally when full (positive) correlation is optimal.

Given some state-independent preferences for the sender $v(\cdot)$, denote by $\tilde{v}_m^{FC}(\mu) \equiv v(a_{\mu^{CN}(\mu,\ldots,\mu)})$ the sender's utility when there are m replications of a posterior μ. The sender's problem can be rewritten as the following m-modified problem:

Definition 1. *In the m-modified problem, the sender has one signal, and she chooses $\tau(.) \in \Delta(\Delta(\Omega))$ to maximize $\sum_{\mu\in\Delta(\Omega)} \tau(\mu)\tilde{v}_m^{FC}(\mu)$, subject to τ being a Bayes-plaussible distribution of posteriors.*

In other words, we have taken the standard rational model with one signal and instead modified the utility of the sender to be defined over the correlation neglect posterior of an m-repeated realization of that signal. In that case, we can use concavification to solve the problem. Let $V_m^{FC}(\mu)$ denote the concavification of $\tilde{v}_m^{FC}(\mu)$.

Our first result shows that full correlation is always optimal in the limit. Recall that we focus on state-independent utilities and that $\mu^{FB} = \arg\max_\mu v(a_\mu)$.

Proposition 3. *Let $v(a)$ be a continuous function. Then the utility of the sender from the solution of the modified problem converges to her first best as the number of signals increases; that is, for any $\epsilon > 0$, $\lim_{m\to\infty} |V_m^{FC}(p) - \tilde{v}(\mu^{FB})| < \epsilon$.*

The intuition behind this proposition parallels the intuition behind the binary example (see Figure 5.b). Any desirable posterior μ^{FB} can be generated by fully correlating m posteriors and putting weight on a posterior that lies in the appropriate direction away from the prior. When m is large, then this movement away from the prior can be minimal and hence can be achieved with probability close to one. Thus, when the number of signals is large, and her preferences are state-independent, we have identified a simple solution to the problem of the sender.

Away from the limit, using the modified problem can also achieve the optimal solution for some utilities. Specifically:

Proposition 4. *Consider $\Omega = \{0, 1\}$. Suppose that the utility of the sender is continuous and supermodular in $\boldsymbol{\mu} = (\mu_1, ..., \mu_m)$. Then the optimal solution to the sender's problem is achieved by the solution to the modified problem for any m.*

The result follows directly from Lorentz (1953) and from Lemma 3. Lorentz (1953) shows that for any supermodular utility, the Fréchet bound is the joint distribution that maximizes the utility given some fixed marginals. Since by Lemma 3 we can restrict ourselves to homogeneous signals, using the upper Fréchet bound corresponds to full correlation and hence the solution of the modified problem leads to the highest value for any supermodular function.[9] Below we provide an example in which the induced utility of the sender is indeed supermodular.

[9]Meyer (2017) shows that the less heterogeneous the marginals are according to the *cumulative column majorization* order, the higher the upper Fréchet bound is with respect to the supermodular order.

Example 2 (Supermodular utility)**.** *Consider* $\Omega = \{0, 1\}$, $\Pr(\omega = 1) = \frac{1}{2}$, *a receiver whose optimal action is the expected state and a sender whose utility is single-peaked and convex, so that* $v(a_{\mu^{CN}(\boldsymbol{\mu})}) = -|\mu^{CN}(\boldsymbol{\mu}) - z|^{\alpha}$, *for some peak* $z > \frac{1}{2}$*. Suppose that* $m = 2$*. Let* α *be sufficiently small so that* $-|\mu^{CN}(\mu_1, \mu_2) - z|^{\alpha}$ *is supermodular in* (μ_1, μ_2)*. Then we know that the optimal solution involves full correlation. As a result we can look at the modified problem, where we consider the utility from* $-|\frac{\mu^2}{\mu^2+(1-\mu)^2} - z|^{\alpha}$*, so now the peak is at* $\mu = \frac{z-\sqrt{z(1-z)}}{2z-1}$ *rather than at z. This implies, by concavification, that the optimal signal (which will be repeated twice) has two posteriors, at 0 and* $\frac{z-\sqrt{z(1-z)}}{2z-1}$*, with* $\tau(\frac{z-\sqrt{z(1-z)}}{2z-1}) = \frac{2z-1}{2(z-\sqrt{z(1-z)})}$.

We conclude this Section by showing that the modified problem using full correlation is not always the optimal solution to the sender's problem. Consider the following example.

Example 3 (Suboptimality of Full Correlation)**.** *Consider* $\Omega = \{0, 1\}$ *and a receiver whose optimal action corresponds to her expected state,* μ*. Suppose that the sender has state independent preferences that are increasing and continuous for* $a < 1$*, with a discontinuity at* $a = 1$*. In other words, the sender receives a bonus when the receiver chooses action 1. Given a posterior* μ*, the sender's induced utility* $v(\mu)$ *is depicted in Figure 7.a.*

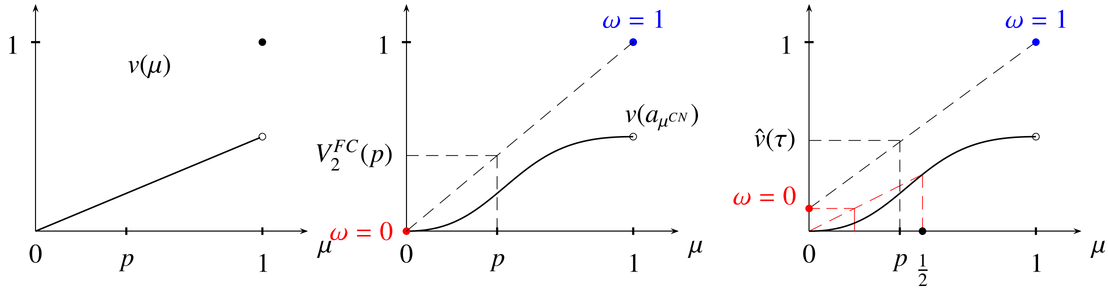
Using the modified problem, one can compute the optimal signal structure when restricted to two fully correlated signals, this corresponds to Figure 7.b. The solid line corresponds to $v(a_{\mu_2^{CN}(\mu)})$*. The dashed line corresponds to the concavification* $V_2^{FC}(\mu)$*. This procedure prescribes full informative signals, i.e., the sender perfectly reveals the state for every state and reaches an expected payoff of* $V^{FC}(p)$*.*

However, the fact that the receiver perfectly understands each marginal implies that it is enough for one of the signals to reveal the state ($\mu_i = 1$)*, in order for the receiver to learn the state* ($\mu^{CN} = 1$)*. It is therefore possible to improve upon the full correlation structure by ensuring that when* $\omega = 1$ *is realised, only one signal reveals the state, while the other pools with realisations that will occur in state* $\omega = 0$*. For example, the following information structure is an improvement over full correlation. Each signal generates three posteriors,* $\mu \in \{0, \frac{1}{2}, 1\}$*. In state 1, the signals are negatively correlated, generating vector of posteriors* $(\frac{1}{2}, 1)$ *and* $(1, \frac{1}{2})$*. In state 0, the signals are fully correlated, generating the vector of posteriors* $(0, 0)$ *and* $(\frac{1}{2}, \frac{1}{2})$*. Specifically, the joint conditional distribution is:*

$$\tau((1, \tfrac{1}{2})|\omega = 1) = \tau((\tfrac{1}{2}, 1)|\omega = 1) = \tfrac{1}{2}$$
$$\tau((0, 0)|\omega = 0) = \tfrac{1}{2}\tfrac{2-3p}{1-p} \quad \tau((\tfrac{1}{2}, \tfrac{1}{2})|\omega = 0) = \tfrac{1}{2}\tfrac{p}{1-p}$$

The new information structure is illustrated in Figure 7.c. While in state $\omega = 1$ *the sender always induces a belief of* 1*, in state* $\omega = 0$ *the sender induces a weighted average between* 0 *and* $\mu_2^{CN}(\frac{1}{2})$ *leading to a expected utility conditional on* $\omega = 0$ *that is strictly positive. The unconditional expected utility of the sender is given by* $\hat{v}(\tau)$ *which is greater than* $V_2^{FC}(p)$*.*

For the case of two signals and binary state space we have a result analogous to Proposition 4:



(a) Sender's Preferences

(b) Full revelation is optimal under full correlation.

(c) Improvement over full correlation.

Figure 7: Suboptimality of Full Correlation

Proposition 5. *Suppose that $\Omega = \{0, 1\}$ and $m = 2$. When the utility of the sender is submodular in (μ_1, μ_2), the optimal information structure involves negative correlation.*

The result follows Müller and Stoyan (2002) and Lemma 3. Müller and Stoyan (2002) show that for any two marginal cumulative distributions F_1, F_2, the joint distribution that maximises the expected value of a submodular function is the lower Fréchet bound of (F_1, F_2). The lower Fréchet bound, whose cumulative distribution is given by $F^-(x_1, x_2) = \max\{0, F_1(x_1) + F_2(x_2) - 1\}$ can be implemented by the following two random vectors $F_1^{-1}(u), F_2^{-1}(1-u)$ where u denotes a random variable uniformly distributed in $[0, 1]$. Since by Lemma 3 we can use homogeneous signals, the lower Fréchet bound involves some negative correlation.

5 Discussion and extensions

We have shown how a sender can manipulate a receiver who completely ignores correlation. We now discuss how our analysis can be extended to other behavioural biases that may arise when individuals combine information sources and specifically to a receiver who can understand some degree of correlation. We then consider how the sender's ability to manipulate is curbed when the receiver has limited attention and cannot consider all m signals.

5.1 Other behavioural biases

Some of the results we presented above allow us to generalize our insights to other cognitive biases the receiver might have. In particular, assume that the receiver has some aggregation function $\mu : \Delta(\Omega)^m \to \Delta(\Omega)$ by which she aggregates the m signals realizations into one posterior belief. Correlation neglect, the aggregation rule we used in the paper, μ^{CN}, is one example of such an aggregation rule. Some of the results in the paper were about properties of the functional $\tilde{v}(\boldsymbol{\mu}) = v(a_{\mu(\boldsymbol{\mu})}, \omega)$ and did not rely on the particular structure of the function μ^{CN}.

For example, whenever $\tilde{v}(\boldsymbol{\mu})$ is symmetric in its arguments, we can focus exclusively on homogenous signal structures. Moreover, the result that concavification is a sometimes unachievable upper bound for the solution, obviously holds in this more general model. Finally, whenever $\tilde{v}(\boldsymbol{\mu})$ is supermodular we can focus our attention on a solution with full correlation and use the modified problem concavification technique to solve for it.

Some possible aggregation rules can relate to biases such as confirmation bias or projection bias. Since we focused in our model on the case of full correlation neglect, it is natural to relax this assumption to think about receivers who to some degree can think about the possibility of correlation. This is what we discuss next.

5.2 Partial correlation neglect

To model partial correlation neglect, we use the modelling technique developed in Levy and Razin (2018); they suggest a one-parameter model that captures the degree to which an individual who combines forecasts entertains correlation, which is detail-free (of the specific distribution functions). The parameter captures the set of different correlation structures that the receiver considers when interpreting or explaining multiple sources of information. The higher is this parameter, the higher degrees of correlation the receiver can contemplate. As we show, our qualitative results are robust to such an extension.

Specifically, consider a parameter $1 \leq \alpha < \infty$, such that when interpreting the signal realizations the decision maker only considers joint information structures, $(S, \Omega, p(\omega), q(s, \omega))$, so that at any state $\omega \in \Omega$ and for any vector of signals $s \in S$, the pointwise mutual information ratio is bounded:

$$\frac{1}{\alpha} \leq \frac{q(\boldsymbol{s}, \boldsymbol{\omega})}{\prod_{i=1}^{m} q^i(s^i|\omega)} \leq \alpha.$$

Note that when $\alpha = 1$, the receiver fully neglects correlation, as she can only consider independent information sources. On the other hand when $\alpha > 1$, she can consider information structures which involve correlation. One way to interpret the case of $\alpha > 1$ is to think of the receiver facing ambiguity about correlation structures; in this case α describes the extent of the ambiguity the decision maker faces.

To be more concrete, consider the binary model with $\Omega = \{0, 1\}$. Levy and Razin (2018) analyze an investment application with a safe action and a risky action which pays more when the state is 1. The receiver has ambiguity aversion la Gilboa and Schmeidler (1989) which implies that she will invest according to her minimum belief that the state is 1, implying that within our framework,

$$\mu(\boldsymbol{\mu}) = \frac{\frac{1}{\alpha} \frac{\prod_{i=1}^{m} \mu_i(1)}{p^{m-1}}}{\frac{1}{\alpha} \frac{\prod_{i=1}^{m} \mu_i(1)}{p^{m-1}} + \alpha \frac{\prod_{i=1}^{m} \mu_i(0)}{(1-p)^{m-1}}}$$

Again, with this well-defined aggregation rule over posteriors, we can use our general results. For example, if the utility function of the sender over $\mu(\boldsymbol{\mu})$ is supermodular, or if it is binary as in Section 4.1, then full correlation will be the optimal solution and we can use concavification to identify the optimal signal structure. Thus, as in our model, the sender will still use full correlation while the receiver will still consider only partial correlation.

5.3 Limited Attention

Our analysis above has assumed that the number of signals that the sender has at her disposal is exogenously given. In particular, some of our results consider what happens when the number of these signals, deemed independent by the receiver, are unboundedly large. But in many applications there would be a reason to believe that the number of such signals is restricted. To be more concrete consider for example a political strategist trying to influence voters' beliefs, by using social media to transmit messages to voters under different aliases. One constraint on the political strategist is the attention of the voter. The voter is not going to read an unbounded number of social media messages.

In practice, such strategies will lead to stochastic outcomes; for every 100 messages the strategist sends a typical voter gets a percentage of them. This implies first that there is a bound on the number of signals the voter will have and hence on the ability of the sender to persuade. But also, this implies that the sender cannot fully control which of her messages the voter will obtain. This implies that the more messages the sender has, the more uncertainty she will face with regard to the effect of her messages on the voter.

When the optimal strategy of the sender is to fully correlate the realisations of her signals, then this second effect will not be a concern. As the sender sends 'identical' messages, she will not care about which of them the reader will actually end up reading. However when the optimal strategy of the sender, with unlimited attention of the receiver, does not involve fully correlated signals, the effect of limited attention will be twofold. The limited attention will bound the influence of the sender, but will also affect her optimal correlation strategy. As a result, the sender may react to limited attention in two different ways. She may change her correlation strategy, maybe shifting towards full positive correlation, which will limit the risk of having the reader read different types of messages. She may also reconsider the number of signals she sends. If she can better control what the reader reads when she lowers the number of signals she sends, she might find it optimal to do so. Whatever the response of the sender is, limited attention will curb the power of the sender to persuade.

A Appendix

Proof of Lemma 1:

Suppose that the conditional posteriors are induced by signal i with information structure $\{S_i, \{q_i(\cdot \mid \omega)\}_{\omega\in\Omega}\}$ through Bayes' rule. Given a realisation s_i, define

$$\mu_{s_i}(\omega) = \frac{p(\omega)q_i(s_i \mid \omega)}{p(\omega)q_i(s_i \mid \omega) + \sum_{\upsilon\neq\omega} p(\upsilon)q_i(s_i \mid \upsilon)}.$$

Then such signal induces the distribution over posteriors τ_i such that $\tau_i(\mu_{s_i} \mid \omega) = q(s_i \mid \omega)$. If $\mu_{s_i}(\omega) = 0$, this implies that $q_i(s_i \mid \omega) = 0$ and so $\tau_i(\mu_{s_i} \mid \omega) = q_i(s_i \mid \omega) = 0$.
Analogously, $\mu_{s_i}(\omega \mid s_i) = 1$ if and only if $q_i(s_i \mid \omega) > 0$ and $\sum_{\upsilon\neq\omega} p(\upsilon)q_i(s_i \mid \upsilon) = 0$. In that case, $\tau_i(\mu_{s_i} \mid \omega) = q(s_i \mid \omega) > 0$.
Lastly, suppose that $q_i(s_i \mid \omega) > 0$ and $\sum_{\upsilon\neq\omega} q_i(s_i \mid \upsilon) > 0$ i.e., $0 < \mu_{s_i}(\omega) < 1$. Then,

$$q_i(s_i \mid \omega) = \frac{\mu_{s_i}(\omega)}{1 - \mu_{s_i}(\omega)} \frac{1}{p(\omega)} \sum_{\upsilon\neq\omega} p(\upsilon)q_i(s_i \mid \upsilon)$$

Therefore:

$$\begin{aligned}
\tau_i(\mu_{s_i} \mid \omega) &= q_i(s_i \mid \omega) \\
&= \tfrac{\mu_{s_i}(\omega)}{1-\mu_{s_i}(\omega)} \tfrac{1}{p(\omega)} \textstyle\sum_{\upsilon\neq\omega} p(\upsilon)q_i(s_i \mid \upsilon) \\
&= \tfrac{\mu_{s_i}(\omega)}{1-\mu_{s_i}(\omega)} \tfrac{1}{p(\omega)} \textstyle\sum_{\upsilon\neq\omega} p(\upsilon)\tau_i(\mu_{s_i} \mid \upsilon)
\end{aligned}$$

which corresponds to the condition in the Lemma. This holds for any s_i that is induced with strictly positive probability and any $\omega \in \Omega$.

Consider now a set of conditional distributions $\{\tau^i(. \mid \omega)\}_{\omega\in\Omega}$ satisfying the conditions in the Lemma. For each posterior μ in the union of the supports of $\tau^i(\cdot \mid \omega)$, we associate a distinct signal s_μ; $s_\mu \neq s_{\mu'}$ whenever $\mu \neq \mu'$. Denote by $S_i = \{s_\mu \mid \mu \in \text{Support}(\tau^i(. \mid \omega)), \omega \in \Omega\}$ and define

$$q_i(s_\mu \mid \omega) = \tau_i(\mu \mid \omega)$$

To see that $\{q_i(\cdot \mid \omega)\}_{\omega\in\Omega}$ constitutes a signal structure note that

$$\sum_{\omega} p(\omega) \sum_{s_\mu} q_i(s_\mu \mid \omega) = \sum_{\omega} p(\omega) \sum_{\mu} \tau_i(\mu \mid \omega) = 1$$

This signal obviously induces the desired posteriors. □

Proof of Lemma 2

A receiver with correlation neglect believes that all signals are conditionally independent. Therefore upon observing realisation $\boldsymbol{s} = (s_1, ..., s_m)$ which leads to posteriors $\boldsymbol{\mu} =$

$(\mu_1, ..., \mu_m)$ her posterior belief is:

$$\mu^{CN}(\omega \mid \boldsymbol{\mu}) = \mu^{CN}(\omega \mid \mathbf{s}) = \frac{q(s_1,...,s_m|\omega)p(\omega)}{\sum_v q(s_1,...,s_m|v)p(v)} = \frac{p(\omega)\prod_{i=1}^m q_i(s_i|\omega)}{\sum_v p(v)\prod_{i=1}^m q_i(s_i|v)}$$

$$= \frac{\prod_{i=1}^m \frac{q_i(s_i|\omega)p(\omega)}{\sum_v q_i(s_i|v)p(v)} \frac{1}{p(\omega)^{m-1}}}{\sum_v \prod_{i=1}^m \frac{q_i(s_i|v)p(v)}{\sum_v q_i(s_i|v)p(v)} \frac{1}{p(v)^{m-1}}} = \frac{\frac{\prod_{i=1}^m \mu_i(\omega)}{p(\omega)^{m-1}}}{\sum_v \frac{\prod_{i=1}^m \mu_i(v)}{p(v)^{m-1}}}. \square$$

Proof of Lemma 3

Consider $\tau \in \Delta(\Delta(\Omega)^m)$ and its marginals $(\tau_1, ..., \tau_m)$. Assume that not all the marginals are the same. Consider all permutations σ over the m signals and τ^σ such that $\tau^\sigma(\mu_{\sigma(1)}, ..., \mu_{\sigma(m)}) = \tau(\mu_1, ..., \mu_m)$. Given the symmetry of μ^{CN}, $\mu^{CN}(\boldsymbol{\mu}) = \mu^{CN}(\boldsymbol{\mu}^\sigma)$, and hence $\hat{v}(\tau^\sigma) = \hat{v}(\tau)$ for all permutations σ.

Denote $\tau^* = \frac{1}{m!}\sum_\sigma \tau^\sigma$. Obviously, $\hat{v}(\tau^*) = \hat{v}(\tau)$. Note that the support of τ^* is the union of the supports of the marginal distributions. Moreover, the conditional marginals derived from τ^* for each i are $\tau_i^*(\mu \mid \omega) = \frac{1}{m}\sum_{j=1}^m \tau_j(\mu \mid \omega)$ for all $\mu \in \cup_{i=1,...,m} Support(\tau_i)$, and hence are homogeneous. We now check that this distribution over posteriors τ^* indeed satisfies condition (1) and hence can be generated by homogeneous signals. For any μ in the support of τ_i^*, we want to check whether

$$\mu(\omega) = \frac{p(\omega)\tau_i^*(\mu \mid \omega)}{\sum_v p(v)\tau_i^*(\mu \mid \omega)}$$

or equivalently, that

$$\frac{\mu(\omega)}{\mu(v)} = \frac{p(\omega)\tau_i^*(\mu \mid \omega)}{p(v)\tau_i^*(\mu \mid \omega)} = \frac{p(\omega)\sum_i \tau_i(\mu \mid \omega)}{p(v)\sum_i \tau_i(\mu \mid \omega)}$$

Denote by M_μ the set of marginals $i \in \{1, ..., m\}$ such that $\mu \in Support(\tau_i)$. For all $j \notin M_\mu$, $\tau_j(\mu \mid \omega) = \tau_j(\mu \mid v) = 0$. For all $i \in M_\mu$, since τ satisfied condition (1) we have that

$$\frac{\mu(\omega)}{\mu(v)} = \frac{p(\omega)\tau_i(\mu \mid \omega)}{p(v)\tau_i(\mu \mid \omega)}$$

which implies

$$\frac{\mu(\omega)}{\mu(v)} = \frac{p(\omega)\sum_{i\in M}\tau_i(\mu \mid \omega)}{p(v)\sum_{i\in M}\tau_i(\mu \mid \omega)} = \frac{p(\omega)\sum_i \tau_i(\mu \mid \omega)}{p(v)\sum_i \tau_i(\mu \mid \omega)}. \square$$

Proof of Proposition 1

Consider two fully correlated signals, each yielding posteriors μ and μ' with probabilities α_ω and $(1 - \alpha_\omega)$ respectively in state $\omega \in \Omega$.
By Lemma 1 above, we have that $\mu(\omega) = \frac{p_\omega \alpha_\omega}{\sum_v p_v \alpha_v}$ and $\mu'(\omega) = \frac{p_\omega(1-\alpha_\omega)}{\sum_v p_v(1-\alpha_v)}$.

Given $\{\alpha_\omega\}_{\omega\in\Omega}$, the expected posterior for the receiver is given by:

$$
\begin{aligned}
E_\tau(\mu^{CN})(\varsigma) &= \sum_{\omega\in\Omega} p_\omega \left(\alpha_\omega \frac{\frac{\mu(\varsigma)^2}{p_\varsigma}}{\sum\limits_{\upsilon\in\Omega} \frac{\mu(\upsilon)^2}{p_\upsilon}} + (1-\alpha_\omega) \frac{\frac{\mu'(\varsigma)^2}{p_\varsigma}}{\sum\limits_{\upsilon\in\Omega} \frac{\mu'(\upsilon)^2}{p_\upsilon}} \right) \\
&= \sum_{\omega\in\Omega} p_\omega \left(\alpha_\omega \frac{p_\varsigma(\alpha_\varsigma)^2}{\sum\limits_{\upsilon\in\Omega} p_\upsilon(\alpha_\upsilon)^2} + (1-\alpha_\omega) \frac{p_\varsigma(1-\alpha_\varsigma)^2}{\sum\limits_{\upsilon\in\Omega} p_\upsilon(1-\alpha_\upsilon)^2} \right) \quad \forall \varsigma \in \Omega
\end{aligned}
$$

Consider $\alpha_\omega = 1 - \epsilon r_\omega$ where $0 \le r_\omega \le \frac{1}{\epsilon}$. Then we can write the expected posterior for the receiver as

$$
\sum_{\omega\in\Omega} p_\omega \left((1-\epsilon r_\omega) \frac{p_\varsigma(1-\epsilon r_\varsigma)^2}{\sum\limits_{\upsilon\in\Omega} p_\upsilon(1-\epsilon r_\upsilon)^2} + \epsilon r_\omega \frac{p_\varsigma(r_\varsigma)^2}{\sum\limits_{\upsilon\in\Omega} p_\upsilon(r_\upsilon)^2} \right) \quad \forall \varsigma \in \Omega.
$$

Taking the derivative with respect to ϵ and evaluating at $\epsilon = 0$ we get for each $\varsigma \in \Omega$:

$$
\begin{aligned}
\left.\frac{\partial E_\tau(\mu^{CN}(\varsigma))}{\partial\epsilon}\right|_{\epsilon=0} &= \frac{\partial}{\partial\epsilon} \sum_{\omega\in\Omega} p_\omega \left.\left((1-\epsilon r_\omega) \frac{p_\varsigma(1-\epsilon r_\varsigma)^2}{\sum\limits_{\upsilon\in\Omega} p_\upsilon(1-\epsilon r_\upsilon)^2} + \epsilon r_\omega \frac{p_\varsigma(r_\varsigma)^2}{\sum\limits_{\upsilon\in\Omega} p_\upsilon(r_\upsilon)^2} \right)\right|_{\epsilon=0} \\
&= p_\varsigma \left[\sum_{\omega\in\Omega} p_\omega r_\omega \Big(\frac{r_\varsigma^2}{\sum_{\upsilon\in\Omega} p_\upsilon r_\upsilon^2} - 1\Big) + 2\Big(-r_\varsigma + \sum_{\upsilon\in\Omega} p_\upsilon r_\upsilon\Big) \right] \\
&= p_\varsigma \sum_{\omega\in\Omega} p_\omega r_\omega \left[\Big(\frac{r_\varsigma^2}{\sum_{\upsilon\in\Omega} p_\upsilon r_\upsilon^2} - 1\Big) + 2\Big(1 - \frac{r_\varsigma}{\sum\limits_{\upsilon\in\Omega} p_\upsilon r_\upsilon}\Big) \right] \\
&\propto \Big(\frac{r_\varsigma^2}{\sum_{\upsilon\in\Omega} p_\upsilon r_\upsilon^2} - 1\Big) + 2\Big(1 - \frac{r_\varsigma}{\sum\limits_{\upsilon\in\Omega} p_\upsilon r_\upsilon}\Big)
\end{aligned}
$$

The first term is positive if $r_\varsigma > r_\upsilon$ and negative otherwise and the second term has opposite sign. We now define $\{r_\omega\}_{\omega\in\Omega}$ so that the expected posterior moves towards the degenerate posterior on state η. Define $r_\eta = \gamma$, $r_\omega = \gamma + \kappa$ for all $\omega \neq \eta$, where $\gamma, \kappa > 0$ are small.

Note that,

- $\sum\limits_{\upsilon\in\Omega} p_\upsilon r_\upsilon = \gamma + (1-p_\eta)\kappa$.
- $\sum\limits_{\upsilon\in\Omega} p_\upsilon r_\upsilon^2 = \gamma^2 + (1-p_\eta)(2\gamma\kappa + \kappa^2)$.

Hence we have that

$$
\begin{aligned}
\left.\frac{\partial E_\tau(\mu^{CN}(\eta))}{\partial\epsilon}\right|_{\epsilon=0} &\propto \Big(\frac{\gamma^2}{\gamma^2 + (1-p_\eta)(2\gamma\kappa+\kappa^2)} + 1 - 2\frac{\gamma}{\gamma + (1-p_\eta)\kappa}\Big) \\
&\propto (\gamma^2(\gamma + (1-p_\eta)\kappa) + (\gamma^2 + (1-p_\eta)(2\gamma\kappa+\kappa^2))(\gamma + (1-p_\eta)\kappa) \\
&\quad -2\gamma(\gamma^2 + (1-p_\eta)(2\gamma\kappa+\kappa^2)) \\
&= \kappa^2(1-p_\eta)[\gamma(1-2p_\eta) + \kappa(1-p_\eta)]
\end{aligned}
$$

and,

$$\begin{aligned}
\left.\frac{\partial E_\tau(\mu^{CN}(\omega))}{\partial \epsilon}\right|_{\epsilon=0} &\propto \left(\frac{(\gamma+\kappa)^2}{\gamma^2+(1-p_\eta)(2\gamma\kappa+\kappa^2)}+1-2\frac{\gamma+\kappa}{\gamma+(1-p_\eta)\kappa}\right)\\
&\propto ((\gamma+\kappa)^2(\gamma+(1-p_\eta)\kappa)+(\gamma^2+(1-p_\eta)(2\gamma\kappa+\kappa^2))(\gamma+(1-p_\eta)\kappa)\\
&\quad -2(\gamma+\kappa)(\gamma^2+(1-p_\eta)(2\gamma\kappa+\kappa^2))\\
&= \kappa^2 p_\eta[\gamma(2p_\eta-1)-\kappa(1-p_\eta)]
\end{aligned}$$

Hence by choosing $\frac{\gamma}{\kappa} > \frac{1-p_\eta}{2p_\eta-1}$ we have that $\left.\frac{\partial E_\tau(\mu^{CN}(\eta))}{\partial \epsilon}\right|_{\epsilon=0} > 0$ and $\left.\frac{\partial E_\tau(\mu^{CN}(\omega))}{\partial \epsilon}\right|_{\epsilon=0} < 0$ for all $\omega \neq \eta$. And we can do that for any state $\eta \in \Omega$, hence by combining these signals we can move the expectation in any direction.□

Proof of Proposition 2

The proof of Proposition 2 uses the following Lemma:

Lemma 4. *Consider a conditional distribution over posteriors $\{\tau_i(\cdot \mid \omega)\}_{\omega\in\Omega}$ induced by a signal given a Bayesian updater. Denote by $M(\omega,k) = \{\mu \in \Delta(\Omega) \mid \mu(\omega) < k\}$ the set of posteriors that gives a probability mass to ω of less than k, then*

$$\tau_i(M(\omega,k) \mid \omega) < \frac{k}{p(\omega)}$$

Proof of Lemma 4

By Lemma 1 for any $\mu \in M(\omega,k)$ that is induced with strictly positive probability, $\mu(\omega) = \frac{p(\omega)\tau_i(\mu|\omega)}{\tau_i(\mu)} < k$. But this implies that $\tau_i(\mu \mid \omega) < \frac{k}{p(\omega)}\tau_i(\mu)$ and summing over all $\mu \in M(\omega,k)$ we get $\tau_i(M(\omega,k) \mid \omega) = \sum_{\mu\in M(\omega,k)} \tau_i(\mu \mid \omega) < \frac{k}{p(\omega)} \sum_{\mu\in M(\omega,k)} \tau_i(\mu) \leq \frac{k}{p(\omega)}$. □

Proof of Proposition 2

For any $\boldsymbol{\mu} \in M_\omega(\delta)$ we have,

$$\mu^{CN}(\omega \mid \boldsymbol{\mu}) = \frac{\frac{\prod_{i=1}^m \mu_i(\omega)}{(p_\omega)^{m-1}}}{\sum_{\upsilon\in\Omega} \frac{\prod_{i=1}^m \mu_i(\upsilon)}{(p_\upsilon)^{m-1}}} > 1-\delta \quad \Leftrightarrow \quad \sum_{\upsilon\neq\omega\in\Omega} \frac{\prod_{i=1}^m \mu_i(\upsilon)}{p(\upsilon)^{m-1}} < \frac{\delta}{1-\delta}\frac{\prod_{i=1}^m \mu_i(\omega)}{p(\omega)^{m-1}}$$

This implies that for any $\boldsymbol{\mu} \in M_\omega(\delta)$ and any $\upsilon \neq \omega \in \Omega$, $\prod_{i=1}^m \mu_i(\upsilon) < \frac{\delta}{1-\delta}\left(\frac{p(\upsilon)}{p(\omega)}\right)^{m-1}$.
For any $\upsilon \neq \omega$, denote $i(\boldsymbol{\mu},\upsilon) \in \arg\min_{i=1,\ldots m} \mu_i(\upsilon)$ we have

$$(\mu_{i(\boldsymbol{\mu},\upsilon)}(\upsilon))^m < \prod_{i=1}^m \mu_i(\upsilon) < \frac{\delta}{1-\delta}\left(\frac{p_\upsilon}{p_\omega}\right)^{m-1} \quad \Leftrightarrow \quad \mu_{i(\boldsymbol{\mu},\upsilon)}(\upsilon) < \left(\frac{\delta}{1-\delta}\right)^{\frac{1}{m}}\left(\frac{p(\upsilon)}{p(\omega)}\right)^{\frac{m-1}{m}}.$$

Denote by $M(i,\upsilon,\delta) = \left\{\boldsymbol{\mu}_i''(\upsilon) < \left(\frac{\delta}{1-\delta}\right)^{\frac{1}{m}}\left(\frac{p(\upsilon)}{p(\omega)}\right)^{\frac{m-1}{m}}\right\}$. By Lemma 4,

$$\tau_i\left(M(i,\upsilon,\delta)\mid\upsilon\right) < \frac{1}{p(\upsilon)}\left(\frac{\delta}{1-\delta}\right)^{\frac{1}{m}}\left(\frac{p(\upsilon)}{p(\omega)}\right)^{\frac{m-1}{m}}.$$

As $\tau(\boldsymbol{\mu}'\mid\upsilon)\leq\tau_i(\mu_i'\mid\upsilon)$, we have that for any $\upsilon\neq\omega$:

$$\tau(M_\omega(\delta)\mid\upsilon)\leq\sum_{i=1}^{m}\tau(M(i,\upsilon,\delta)\mid\upsilon)\leq\sum_{i=1}^{m}\tau_i(M(i,\upsilon,\delta)\mid\upsilon)\leq m\frac{1}{p(\upsilon)}\left(\frac{\delta}{1-\delta}\right)^{\frac{1}{m}}\left(\frac{p(\upsilon)}{p(\omega)}\right)^{\frac{m-1}{m}}.$$

Hence,

$$\tau(M_\omega(\delta))=\sum_{\upsilon\in\Omega}p(\upsilon)\tau(M_\omega(\delta)\mid\upsilon)\leq p(\omega)+\sum_{\upsilon\neq\omega}m\left(\frac{\delta}{1-\delta}\right)^{\frac{1}{m}}\left(\frac{p(\upsilon)}{p(\omega)}\right)^{\frac{m-1}{m}}\longrightarrow_{\delta\to 0}p(\omega)$$

□

Proof of Proposition 3:

Consider μ^{FB} the optimal posterior from the point of view of the sender. Suppose first that $\mu^{FB}(\omega)\neq 0$ for all $\omega\in\Omega$. Let $\hat{\mu}$ be defined in the following way:

$$\hat{\mu}_m(\omega)=\frac{\mu^{FB}(\omega)^{\frac{1}{m}}p(\omega)^{\frac{m-1}{m}}}{\sum_{\nu\in\Omega}\mu^{FB}(\nu)^{\frac{1}{m}}p(\nu)^{\frac{m-1}{m}}}$$

which then implies that:

$$\mu_m^{CN}(\hat{\mu}_m)(\omega)=\mu^{FB}(\omega)$$

Note that since p is interior, we can always design a signal structure $\hat{\tau}_m$ with two realizations $\hat{\mu}_m$ and μ_m', where μ_m' at a distance $\delta>0$ from p, and such that the Bayesian plausibility constraint is satisfied:

$$\tau(\hat{\mu}_m)\hat{\mu}_m+(1-\tau(\hat{\mu}_m))\mu_m'=p.$$

Note that when $m\to\infty$, then $\hat{\mu}_m$ is arbitrarily close to p, as

$$\hat{\mu}_m(\omega)=\frac{p(\omega)^{1-\frac{1}{m}}}{\sum_{\nu\in\Omega}(\frac{\mu^{FB}(\nu)}{\mu^{FB}(\omega)})^{\frac{1}{m}}p(\nu)^{1-\frac{1}{m}}}\to_{m\to\infty}p(\omega).$$

As a result, given the Bayesian plausibility constraint, and maintaining μ_m' always at a fixed distance $\delta>0$ away from p, $\tau(\hat{\mu}_m)\to_{m\to\infty}1$, implying that $\lim_{m\to\infty}|\hat{v}(\tau(\hat{\mu}_m))-\tilde{v}(\mu^{FB})|<\epsilon$.

Note that we have found a signal which yields a utility arbitrarily close to μ^{FB} in the limit as m grows large. Note that as concavification provides the optimal solution for any m, then

$$\lim_{m\to\infty}|V_m^{FC}(p)-\tilde{v}(\mu^{FB})|<\lim_{m\to\infty}|\hat{v}(\tau(\hat{\mu}_m))-\tilde{v}(\mu^{FB})|<\epsilon.$$

Consider now the case in which $\mu^{FB}(\omega)=0$ for some $\omega\in\Omega$. By the continuity of the utility function, we can find μ^{SB} such that $\mu^{SB}(\omega)\neq 0$ for all $\omega\in\Omega$, with $|\tilde{v}(\mu^{SB})-\tilde{v}(\mu^{FB})|<\frac{\epsilon}{2}$ and repeat the above with μ^{SB} and $\frac{\epsilon}{2}$. □

Proof of Theorem 1

We consider m signals with the same marginal distribution. Each individual signal i generates n posteriors ($n = |\Omega|$),$\{\mu^{\omega}\}_{\omega\in\Omega}$. The conditional distribution over posteriors given i is:

$$\tau_i(\mu^{\omega}|\omega) = \alpha(m), \qquad \tau_i(\mu^{\upsilon}|\omega) = \beta(m) = \frac{1-\alpha(m)}{n-1}$$

Note that by Lemma 1, the posteriors $\{\mu^{\omega}\}_{\omega\in\Omega}$ are pinned down by $\alpha(m)$ and $\beta(m)$.[10] We choose $\alpha(m)$ and $\beta(m)$ such that, $\alpha(m) \in (\frac{1}{n}, 1)$ and

$$\lim_{m\to\infty} \alpha(m) = \frac{1}{n} \quad \text{and} \quad \lim_{m\to\infty} \frac{1}{m^2\alpha'(m)} = 0 \tag{2}$$

(For example, $\alpha(m) = \frac{1}{n} + \frac{1}{\sqrt{m}}$ would work).
We now construct a joint information structure with these marginals that in state ω will generate with probability converging to 1 (as m increases) a posterior $\mu^{CN}(\boldsymbol{\mu})$ converging to q^{ω}. In order to simplify the notation we eliminate the references to m in our variables whenever it does not lead to confusion.
Given a vector of posteriors $\boldsymbol{\mu} = (\mu_1, ..., \mu_m)$, denote by z_{υ} the proportion of posterior μ^{υ} out of the m posteriors ($z_{\upsilon} \in \{\frac{0}{m}, ..., \frac{m}{m}\}$, and $\sum_{\upsilon\in\Omega} z_{\upsilon} = 1$).
Given the vector $\boldsymbol{z} = \{z_{\upsilon}\}_{\upsilon\in\Omega}$, the receiver's posterior is given by:

$$\mu^{CN}(\omega \mid z) = \frac{\frac{\prod_{\upsilon\in\Omega}\mu^{\upsilon}(\omega)^{mz_{\upsilon}}}{p(\omega)^{m-1}}}{\sum_{\eta}\frac{\prod_{\upsilon\in\Omega}\mu^{\upsilon}(\eta)^{mz_{\upsilon}}}{p(\eta)^{m-1}}} = \frac{p(\omega)\left[\alpha^{z_{\omega}}\beta^{1-z_{\omega}}\right]^m}{\sum_{\upsilon} p(\upsilon)\left[\alpha^{z_{\upsilon}}\beta^{1-z_{\upsilon}}\right]^m} \tag{3}$$

Suppose that in state $\omega \in \Omega$ the sender wants to generate a posterior $q^{\omega} \gg 0$. Define the proportions $\boldsymbol{\zeta}^{\omega} = \{\zeta^{\omega}_{\upsilon}\}_{\upsilon\in\Omega}$ as follows:

$$\zeta^{\omega}_{\upsilon} = \frac{1}{n} + \frac{1}{m}\frac{\ln\left(\frac{q^{\omega}(\upsilon)/p(\upsilon)}{\left[\prod_{\eta} q^{\omega}(\eta)/p(\eta)\right]^{\frac{1}{n}}}\right)}{\ln\left(\frac{\alpha(m)}{\beta(m)}\right)}$$

Note that $\sum_{\upsilon\in\Omega}\zeta^{\omega}_{\upsilon} = 1$ and that given conditions (2), $\zeta^{\omega}_{\upsilon}$ converges to $\frac{1}{n}$, and for sufficiently large m, $\zeta^{\omega}_{\upsilon} > 0$ for all $\upsilon \in \Omega$. The vectors $\zeta^{\omega}_{\upsilon}$ correspond to the vectors of proportions that the sender would ideally like to generate,[11] since

$$\mu^{CN}(\boldsymbol{\zeta}^{\omega}) = q^{\omega}.$$

Denote by $\boldsymbol{z}^{\omega}$ the vector of proportions that is closest to $\boldsymbol{\zeta}^{\omega}$.[12] As m goes to infinity,

$$\boldsymbol{z}^{\omega}(m) \longrightarrow \boldsymbol{\zeta}^{\omega}(m) \quad \text{and} \quad \mu^{CN}(z^{\omega}(m)) \longrightarrow q^{\omega}.$$

[10]The probabilities $\alpha(m)$ and $\beta(m)$ are independent of the state, and hence all the signals have indeed the same marginal distributions.
[11]In general, it will not be possible to generate exactly those weights with the m posteriors unless $\zeta^{\omega}_{\upsilon}$ are rational and m is sufficiently large.
[12]Recall that $z^{\omega}_{\upsilon} \in \{\frac{0}{m}, \frac{1}{m}, .., \frac{m}{m}\}$, so the set of proportion vectors that can be generated with m signals is finite.

What remains to be shown is that the sender can design a joint distribution such that for each state $\omega \in \Omega$, the distribution induces the vector of proportion z^{ω} with probability approaching 1, and still is compatible with the marginal distributions of the signals described above.[13]

For any $\omega \in \Omega$ define

$$\begin{aligned} \gamma^{\omega} &= \min\left\{\tfrac{\alpha}{z^{\omega}_{\omega}}, \min\left\{\tfrac{\beta}{z^{\omega}_{\upsilon}} | \upsilon \in \Omega\right\}\right\}, \\ \lambda^{\omega}_{\omega} &= \alpha - \gamma^{\omega} z^{\omega}_{\omega}, \\ \lambda^{\omega}_{\upsilon} &= \beta - \gamma^{\omega} z^{\omega}_{\upsilon}. \end{aligned} \tag{4}$$

Note that $\gamma^{\omega} \leq 1$, because if not we would have $\sum_{\upsilon \in \Omega} z^{\omega}_{\upsilon} < 1$. Moreover $\gamma^{\omega} + \sum_{\upsilon} \lambda^{\omega}_{\upsilon} = 1$

Define as well $\{\delta^{\omega}\}_{\omega \in \Omega}$ the posteriors that results when all the posteriors are equal to μ^{ω}: $\delta^{\omega} = (\mu^{\omega}, ..., \mu^{\omega})$. This corresponds to a vector of proportions $\mathbf{e}^{\omega}$ in which $e^{\omega}_{\omega} = 1$ and $e^{\omega}_{\upsilon} = 0$ for all $\upsilon \neq \omega$.

Consider the following joint information structure:

For any state $\omega \in \Omega$, with probability γ^{ω} it generates the vector $\mathbf{z}^{\omega}$, with probability $\lambda^{\omega}_{\upsilon}$ it generates the vector e^{υ}. Then the conditional probabilities over posteriors are:

$$\begin{aligned} \tau^{S^m}((z^{\omega})|\omega) &= \gamma^{\omega} \\ \tau^{S^m}(\delta^{\omega}|\omega) &= \lambda^{\omega}_{\omega} \\ \tau^{S^m}(\delta^{\upsilon}|\omega) &= \lambda^{\omega}_{\upsilon} \end{aligned}$$

Clearly, this joint signal distribution generates the marginals we described:

$$\begin{aligned} \tau^{i}(\mu^{\omega}|\omega) &= \gamma^{\omega} z^{\omega}_{\omega} + \lambda^{\omega}_{\omega} = \alpha \\ \tau^{i}(\mu^{\upsilon}|\omega) &= \gamma^{\omega} z^{\omega}_{\upsilon} + \lambda^{\omega}_{\upsilon} = \beta, \quad \forall \upsilon \neq \quad \omega \end{aligned}$$

Moreover, $\alpha(m)$, $\beta(m)$ and $z^{\omega}_{\upsilon}(m)$ converge to $\frac{1}{n}$ for all $\omega, \upsilon \in \Omega$, and therefore $\gamma^{\omega}(m)$ converges to 1 and

$$\tau^{S^m}((z^{\omega})|\omega) \to 1 \quad with \quad \mu^{CN}(z^{\omega}) \longrightarrow q^{\omega}$$

Finally, if the posteriors the sender wants to achieve are not interior, we could always approximate them by a sequence of interior posteriors $\{q^{\omega}_{m}\}$ and we can construct the joint information as above in which the $\boldsymbol{\zeta}^{\omega}(m)$ is constructed using $q^{\omega}_{m} \gg 0$. □

Figure 8 illustrate the construction. The point p represents the prior and q^{ω} is the posterior that we want to achieve in state ω. The signals have three realisations $\{1, 2, 3\}$. Depending on the realisation the posterior moves closer to the corresponding state. The

[13] Note that if the sender were to induce the vector z^{ω} with probability 1 in state ω, then the marginal distribution of the signals would be:

$$\begin{aligned} \tau^{i}(\mu^{\omega}|\omega) &= z^{\omega}_{\omega} \\ \tau^{i}(\mu^{\upsilon}|\omega) &= z^{\omega}_{\upsilon} \quad \forall \upsilon \neq \quad \omega \end{aligned}$$

which in general are distinct from $\alpha(m)$ and $\beta(m)$ (since $\alpha(m)$ and $\beta(m)$ are independent of ω and generally all the z^{ω}_{υ} will be different. However, it is enough to show that for each state ω, the sender can generate the frequency vector $\mathbf{z}^{\omega}$ of marginal posteriors with probability approaching 1.

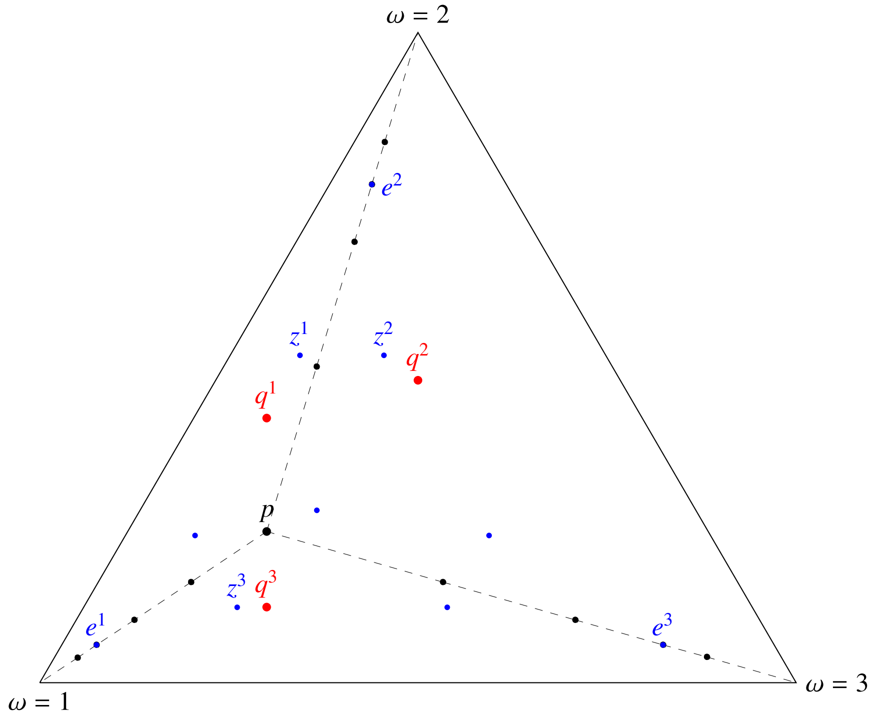


Figure 8: Approximation to Full Manipulation

small dots on the dashed lines represent the posteriors when the receiver gets s signals with the same realisation. Suppose that we have m signals. The we can find the different combinations of posteriors that we can generate given these signals. In Figure 8 we have represented in blue the set of posteriors that we can generate with $m = 3$ signals.

The steps vary with the α we consider, the closer to $\frac{1}{3}$ the smaller are the steps and the finer the grid we can achieve as we increase m.

Given a particular m and α we find among all the combination of realisations (i.e., all the possible vectors z) the one closer to q^ω for each ω, and denote it by z^ω. In Figure 1, given $m = 3$ and the α, we have that $z_1 = \frac{1}{3}(1, 2, 0)$, $z_2 = \frac{1}{3}(0, 2, 1)$ and $z^3 = \frac{1}{3}(2, 0, 1)$. (In general it will not fall exactly on q^ω unless q^ω is a vector of rational numbers, α is also rational, and we have enough signals).

Suppose that the joint signal structure only generates the vectors of frequencies z^ω. Then in state ω, z^ω_ω should be equal to α. But α is independent of the state, so that would mean that $z^\omega_\omega = z^\upsilon_\upsilon$ which is clearly not general (in particular in the example depicted $z^1_1 = \frac{1}{3} \neq \frac{2}{3} = z^2_2$). In order to be able to induce those vectors with the highest possible probability we need to complete the signal structure with the extreme vectors which have all the realisation equal to each other e^ω.

References

Alonso, R. and Câmara, O. (2016a). Bayesian Persuasion with Heterogeneous Priors, *Journal of Economic Theory* **165**: 672 – 706.

Alonso, R. and Câmara, O. (2016b). Persuading Voters, *American Economic Review* **106**(11): 3590–3605.

Arieli, I. and Babichenko, Y. (2016). Private Bayesian Persuasion, *Mimeo* .

Aumann, R. J. and Maschler, M. B. (1995). *Repeated Games with Incomplete Information*, Cambridge, MA: MIT Press.

Bordley, R. (1982). A Multiplicative Formula for Aggregating Probability Assessments, *Management Science* **28**(10): 1137–1148.

Cagé, J., Hervé, N. and Viaud, M.-L. (2017). The Production of Information in an Online World: Is Copy Right?, *CEPR Dicussion Paper 12066* .

De Marzo, P., Vayanos, D. and Zwiebel, J. (2003). Persuasion Bias, Social Influence and Unidimensional Opinions, *The Quaterly Journal of Economics* **118**(3): 909–968.

Eliaz, K. and Spiegler, R. (2018). Incentive Compatible Advertising in Non-Retail Platforms, *Working Paper* .

Ellis, A. and Piccione, M. (2017). Correlation Misperception in Choice, *American Economic Review* **107**(4): 1264–92.

Ellis, A., Piccione, M. and Zhang, X. (2018). Equilibrium Securitizarion with Diverse Beliefs, *Mimeo* .

Enke, B. and Zimmermann, F. (2018). Correlation Neglect in Belief Formation, *Review of Economic Studies* **0**: 1–20.

Eyster, E. and Weizsäker, G. (2011). Correlation Neglect in Financial Decision-Making, (Mimeo).

Gagnon-Bartsch, T. and Rabin, M. (2015). Naïve Social Learning, Mislearning, and Unlearning, (Mimeo).

Glaeser, E. L. and Sunstein, C. R. (2009). Extremism and Solcial Learning, *Journal of Legal Analysis* **1**(1).

Golub, B. and Jackson, M. O. (2012). How Homophily Affects the Speed of Learning and Best-Response Dynamics, *The Quarterly Journal of Economics* **127**(3): 35–71.

Kallir, I. and Sonsino, D. (2009). The Neglect of Correlation in Allocation Decisions, *Southern Economic Journal* **75**(4): 1045–1066.

Kamenica, E. and Gentzkow, M. (2011). Bayesian Persuasion, *American Economic Review* **101**(6): 2590–2615.

Koessler, F., Laclau, M. and Tomala, T. (2018). Interactive Information Design, *Mimeo* .

Levy, G., Moreno de Barreda, I. and Razin, R. (2018). Media Power via the Correlation of News Content, (Mimeo).

Levy, G. and Razin, R. (2015a). Correlation Neglect, Voting Behaviour and Information Aggregation, *American Economic Review* **105**(4): 1634–1645.

Levy, G. and Razin, R. (2015b). Does Polarization of Opinions Lead to Polarization of Platforms? The Case of Correlation Neglect, *Quarterly Journal of Political Science* **10**(3): 321–355.

Levy, G. and Razin, R. (2018). Combining Forecasts: Why Decision Makers Neglect Correlation, *Working Paper, LSE* .

Lorentz, G. G. (1953). An Inequality for Rearrangements, *The American Mathematical Monthly* **60**(3): 176–179.

Meyer, M. (2017). Information Design: Insights from Orderings of Dependence and Heterogeneity, *Mimeo* .

Morris, P. (1977). Combining Expert Judgments: A Bayesian Approach, *Management Science* **23**(7): 679–693.

Müller, A. and Stoyan, D. (2002). Comparison Methods for Stochastic Models and Risks, *Wiley* .

Ortoleva, P. and Snowberg, E. (2015). Overconfidence in Political Behavior, *American Economic Review* **105**(2): 504–35.

Schumacher, H. and Thysen, H. C. (2018). Equilibrium Contracts and Boundedly Rational Expectation , *Working Paper* .

Sobel, J. (2014). On the Relationship Between Individual and Group Decisions, *Theoretical Economics* (9): 163–185.