Calibrated Forecasting and Persuasion

Atulya Jain

Vianney Perchet

Abstract

How should an expert send forecasts to maximize her payoff given she has to pass a calibration test? We consider a dynamic game where an expert sends probability forecasts to a decision-maker. The decision-maker, based on past outcomes, verifies the claims of the expert using the calibration test. We find the optimal forecasting strategy by reducing the dynamic game in terms of a static persuasion problem for the class of stationary ergodic processes. We characterize the value of expertise by showing that an informed expert can achieve the best outcome in the persuasion problem, while an uninformed expert can only achieve the worst. We also compare the calibration test and regret minimization as heuristics for decision-making. We show that an expert can always guarantee the calibration benchmark and in some instances, she can guarantee strictly more.
Keywords: strategic forecasting, calibration, regret, Bayesian persuasion, approachability.

JEL Codes: C72, C73.

1 Introduction

Probability forecasts are widely used by experts to provide information about stochastic events. The forecasts shape the beliefs of decision-makers and persuade them to take specific actions. For example, investors rely on forecasts by financial analysts to determine which asset will achieve the best performance. However, the decision-maker follows the expert's forecasts only if they are credible. One way to determine credibility is to perform statistical tests on data and verify the claims of the expert. We focus on an objective and reasonable test: *calibration*. It is based on the frequency interpretation of probability. The basic idea is to check if the forecast of a state matches the actual proportion of times the state occurred when the forecast was announced a large number of times. For example, the investor checks if an asset outperformed others 70% of the days on which an analyst claimed the chance of it being the best was 0.7. Calibration is central to forecasting and is used to assess the accuracy of prediction markets [Page and Clemen, 2012]. Decision-makers rely on accurate forecasters to take optimal actions. But in many settings, the expert herself has skin in the game and is impacted by the decision-maker's action. This preference misalignment leads to strategic forecasting. For example, if a financial analyst receives a high commission when a particular asset is bought, her predictions might favour this asset. We study the extent of an expert's utility gain from strategic forecasting under a dynamic decision problem. Our main focus is on an *informed expert* who knows the data-generating process and could pass any complex statistical test, including calibration. Given that she can be tested on the basis of some calibration criterion (and failure leads to a large loss), how should an expert send forecasts to maximize her utility?

To do this, we develop a dynamic sender-receiver model. The state of nature evolves over time according to a stochastic process. At every period, the sender sends a probabilistic forecast about that day's state to the receiver. The receiver performs the calibration test to determine the sender's credibility. If the sender passes the calibration test, the receiver takes the forecasts at face value and plays as if the state is drawn according to the forecast. Else he takes a punishment action that results in a loss for the sender. The sender seeks to persuade the receiver to choose actions that are aligned with her preferences.

One of the basic assumptions made in dynamic sender-receiver models is that the receiver either knows the true distribution of the process or has a prior belief over the states. Thus, given the sender's strategy, he can perfectly analyze the messages, deduce the posterior beliefs and take actions. Handling such beliefs in equilibrium, even in simple situations, is a daunting task. In contrast to the Bayesian setting, we focus on the case where the receiver does not know the distribution of the state nor has any prior belief over the distribution to begin with. He simply verifies the claims of the sender using the data so far: the forecasts and the states. The receiver uses the *calibration test* to determine the credibility of the sender. An informed sender who provides honest forecasts passes the calibration test with a very high probability.

We first analyze the case of an informed sender, who knows the data-generating process. She can always pass the calibration test by reporting honestly but there can be other strategies that passes the test. We show the requirement to pass the calibration test limits the distribution of forecasts that can be induced. For a natural class of stochastic processes, we show that any distribution of forecasts can be implemented by a calibrated forecasting strategy if and only if it is (weakly) less informative than distribution of forecasts one gets from honest reporting. Overall, the forecasts need to be accurate but can be less precise than honest forecasts.

Our main result shows that the sender's highest payoff coincides with the sender's equilibrium payoff in a static persuasion problem. The state space of the persuasion problem is given by the set of all possible forecasts of the dynamic game and the prior distribution is given by the distribution of honest forecasts. In the static problem, a sender persuades an receiver by committing to a signaling policy. Kamenica and Gentzkow [2011] and Arieli et al. [2020] characterize the optimal signaling policy that maximizes the sender's ex-ante expected payoff. The optimal forecasting strategy is given by the optimal signaling policy of the persuasion problem. The strategy is stationary as it only depends on the period's conditional distribution of the states. To summarize, we solve the dynamic forecasting game by showing it is equivalent to a static persuasion problem.

Our second set of results characterizes the maximum payoff an uninformed sender, who does not know the data-generating process, can guarantee. Foster and Vohra [1998] show that even an uninformed sender can pass the calibration test for any process. As in the case of an informed sender, we analyze the equilibrium outcome and payoffs of the dynamic game in terms of the static persuasion problem. However, for an uninformed sender, the prior belief is given by the (limit) empirical distribution of the states. First, we model Nature as an adversary trying to prevent the sender from passing the test. The maximal payoff that she can guarantee in an adversarial environment corresponds to the sender's worst signaling policy of the persuasion problem. This payoff is (weakly) lower than the payoff an informed sender can get by honest reporting. Next, we show that if the environment is non-adversarial, the sender can guarantee much more. For a natural class of processes, the sender can (approximately) guarantee the payoff that corresponds to the no information (or babbling) signaling policy of the persuasion problem. This implies that the uninformed sender is always able to learn the empirical distribution of the states. Thus, we are able to compare what an informed and uninformed sender can guarantee and characterize the value of information for a process.

We use our forecasting model to analyze a financial app that sends forecasts to its users about a state that evolves according to a Markov chain. Additionally, the users can sequentially observe signals by paying a price to the app. The app's utility depends on the amount of signal acquired from users and its reputation for announcing precise forecasts. On one hand, a precise forecast results in higher utility from reputation but on the other hand reduces the amount of signals a user will acquire. We characterize the optimal forecasting strategy for the app using the concavification approach. The optimal forecasting strategy, at any stage, only depends on the that day's conditional distribution of the states. The forecasts need to be accurate to pass the calibration test. But even though the app has access to precise information it prefers to send garbled (less informative) forecasts.

Finally, we also model the receiver's behaviour using regret, given its close connection with calibration (see Perchet [2014]). The receiver's regret measures the difference in payoff he could have gotten and what he actually got. Heuristics based on regret minimization ensure that, in hindsight, the receiver could not have done better by playing any fixed action repeatedly. We show that a receiver has no regret if he follows the recommendations of any calibrated forecasting strategy. This provides a justification to use the calibration test as a heuristic in non-Bayesian environments. On the other hand, we show that when facing a regret minimizing receiver, the sender can guarantee the equilibrium payoff in the persuasion problem , as in the case of the calibration test. In fact, we provide an example, for a natural class of regret minimizing algorithms, where she can guarantee strictly more.

Our paper makes contributions to different strands of the literature. We develop a general framework to analyze the strategic interaction between a forecaster and a decision-maker. We introduce the calibration test as a heuristic for decision-making, which has previously been used mostly from a statistical viewpoint. We characterize the optimal forecasting strategy when the calibration test is used as the credibility criterion. We provide a micro-foundation for the commitment assumption in static persuasion models. Our results show that the dynamic interaction between a sender and receiver performing the calibration test is equivalent to a static persuasion model where the sender has ex-ante commitment. Finally, we show a novel connection between regret minimization and the calibration test, both of which can be viewed as heuristics for decision-making in non-Bayesian environments.

1.1 Related Literature

Information disclosure: Our work contributes to the literature on communication between informed senders and uninformed receivers. In cheap talk [Crawford and Sobel, 1982], the message sent by the sender is unverifiable. While in Bayesian persuasion [Kamenica and Gentzkow, 2011], the sender has exogenous commitment and her message is verifiable. Our work contributes in investigating how well dynamic cheap talk interactions substitute for the commitment assumption [Best and Quigley, 2020]. A related paper is Guo and Shmaya [2021], who consider a static model of forecasting and introduce an exogenous cost of miscalibration. Our dynamic model can be thought of as an endogenous microfoundation for their cost of miscalibration. Under reasonable assumptions, we show that the dynamic forecasting game can be characterized in terms of the static persuasion problem (with a large state space)[Arieli et al., 2020].

Calibration and Expert Testing: The initial focus of this literature has been to show that an agent who only knows the test and does not know the data-generating process can pass a class of statistical tests. Foster and Vohra [1998] show that an uninformed expert can guarantee calibration without knowing the data-generating process. Most of the literature (apart from Echenique and Shmaya [2007], Gradwohl and Salant [2011] and Olszewski and Pęski [2011]) have looked at tests from a statistical standpoint without focusing on the underlying decision problem. We show that the expert's incentive to report honestly depends on the decision problem and we analyze the extent of utility gain from strategic misreporting. In line with the literature, we compare the attainable payoffs that an informed and uninformed expert can guarantee for a process. Calibration also has important applications in machine learning (see Gupta and Ramdas [2021]). We refer curious readers to Foster and Vohra [2013] and Olszewski [2015] for a good survey on the topic.

No-Regret and Online Learning: There has been a growing interest in implementing regret minimization in dynamic interactions. We focus on optimization against regret minimizing agents. The closest paper is Deng et al. [2019] . We use a natural class of regret learning algorithms called *mean-based learning algorithms* introduced by Braverman et al. [2018]. A common theme in both papers is that they show the optimizer can obtain a higher utility than the rational benchmark. On a similar note we show, against a mean-based learner, the sender can obtain a higher utility than the calibration benchmark. We solve our constrained dynamic optimization problem using tools from online learning (see Bernstein et al. [2014], Mannor et al. [2009]).

Bounded Rationality: Both the calibration test and regret minimization can be seen as explicit procedures for decision-making. On a general level, our model is a dynamic interaction between a rational sender and a boundedly rational receiver [Rubinstein, 1998]. The sender is a *strategic agent* while the receiver is a *non-strategic* agent who uses heuristics [Spiegler, 2014].

1.2 Organization

Section 2 introduces the model and defines the calibration test. In section 3, we present the persuasion problem, prove our main results for the case of informed and uninformed sender and finally provide an application. In Section 4, we consider a receiver who minimizing regret. In Section 5, we conclude and discuss future work. All omitted proofs are in Appendix A. In Appendix B, we consider an environment where the action of the decision-maker affects how the states evolves.

2 Model

We consider a dynamic game of incomplete information between two players: sender (she) and receiver (he). At each period, the sender sends a forecast (message) about that day's state which is unobserved by the receiver. The receiver then chooses an action. The state and action of that period are observed before proceeding to the next period.

Let Ω denote the finite set of states, $F \subseteq \Delta\Omega$ denote the set of forecasts over the states, and A denote the finite set of actions. Denote a "play", i.e., an infinite sequence of states, by $\omega^\infty = (\omega_1, ...)$. The state ω_t evolves over time and is governed by a stochastic process with distribution $\mu \in \Delta\Omega^\infty$. Denote by ω_t and $\omega^t = (\omega_1, ..., \omega_{t-1})$ the state and history of states at time t respectively. We assume the sender is informed and knows the distribution of the process while the uninformed receiver does not. Given $\mu \in \Delta\Omega^\infty$, the sender can compute the conditional distribution $p_t \in \Delta\Omega$ of the state ω_t given the history ω^t, i.e., $p_t = \mu(\omega_t \mid \omega^t) \in \Delta\Omega$. We assume that the conditional distribution p_t takes values from a finite set $D \subset \Delta\Omega$ [1] (the assumption holds for instance for standard finite Markov chains).

At stage t, the receiver's action a_t, together with the state ω_t, determines the stage payoff $u_S(\omega_t, a_t)$ and $u_R(\omega_t, a_t)$ for the sender and receiver respectively. The sender's goal is to maximize her long-run average payoff. She does not take the action herself but can persuade the receiver into taking actions that are aligned with her preference. She does this by providing information in the form of probabilistic forecasts. At each period t, she chooses the forecast f_t based the history of states and forecasts. Formally, the sender's forecasting strategy is a map $\sigma : \cup_{t\geq 1}(F \times \Omega)^{t-1} \to \Delta F$. Unless specified otherwise, the set of feasible forecasts $F = \Delta\Omega$.

Calibration test: To model the receiver's behaviour, we take a frequentist approach using the calibration test. Given he has no prior information nor belief over the sender's strategy, he verifies the claims of the sender by performing the calibration test on the data so far. He checks if the predicted forecasts match the empirical distribution of the states when the forecast was made. For each round t, given an error margin ϵ_t, he checks if the forecasts are close to the realized frequency of the states.

Formally, fixing an error margin ϵ_T, the receiver performs the ϵ_T-calibration test at round T. Based on the history of states and forecasts $h_T = (f_1, \omega_1, ...\omega_{T-1}, f_T)$, he checks if the forecasts are ϵ_T-close to the empirical frequency of the states.

Definition 1 *A T-sequence of forecasts $(f_t)_{t=1}^T$ is ϵ_T-calibrated if*

$$\sum_{f \in F} \frac{|\mathbb{N}_T[f]|}{T} \|\overline{\omega}_T[f] - f\| \leq \epsilon_T \tag{1}$$

[1] In case P is not finite, we can construct a finite ϵ-grid $P = \{p_l; l \in L\}$ such that for any $p \in \Delta\Omega$ we have a $k \in L$ such that $\|p - p_k\| \leq \epsilon$.

where, $\mathbb{N}_T[f]$ and $\overline{\omega}_T[f]$ refers to the set of stages and the empirical distribution of states when the forecast is f up to stage T respectively and $\|\cdot\|$ is the Euclidean norm, i.e.,

$$\mathbb{N}_T[f] := \{t \in \{1, ..., T\} : f_t = f\} \qquad \overline{\omega}_T[f] := \frac{\sum_{t \in \mathbb{N}_T[f]} \delta_{\omega_t}}{|\mathbb{N}_T[f]|} \tag{2}$$

where, δ_ω denotes the Dirac distribution on state ω.

Intuitively, a T-sequence of forecasts is ϵ_T-calibrated if the empirical distribution of states $\overline{\omega}_T[f]$ is close to f for all possible f (that were used, this is why, even though the set F might be infinite, the sum in the above definition is implicitly defined, over the forecasts actually sent once).

Pass: A sender passes the calibration test at stage t if the sequence of forecasts $\{f_i\}_{i=1}^t$ is ϵ_t-calibrated. In this case, the receiver takes the forecast at face value and responds as if the state ω_t is drawn randomly according to the forecast f_t. He plays as if his beliefs over the states match with the sender's forecast f_t. He plays the action $\hat{a}(f_t)$, where $\hat{a}(f_t)$ denotes the receiver's optimal action given his belief over states is f_t, i.e.,[2]

$$\hat{a}(f_t) := \underset{a \in A}{\operatorname{argmax}} \{ \sum_{\omega \in \Omega} f_t(\omega) u_R(\omega, a) \} \tag{3}$$

Fail: If the sender fails the test, he uses an action that results in a punishment cost $c > 0$ for the sender in that round. The punishment does not last forever. The receiver does not play according to the sender's forecast until she passes the calibration test at a later round. This action can be interpreted as the the receiver's safe action, which is also the sender's least preferred action. For example, in the case of financial forecasting, it could correspond to not buying any asset, which results in zero commission for the analyst. The punishment cost could also be interpreted as a loss in credibility or reputation from inaccurate predictions. One can imagine an intermediary (or platform) performing the calibration test, where in a sender's forecast is only forwarded to the receiver if it passes the calibration test. Our motivation is to study the extent of persuasion when the receiver uses the calibration test as the credibility criterion. So, our focus will be on a high punishment cost which enforces calibration.

Note that, for some $\epsilon_T > 0$, no forecasting strategy will be ϵ_T-calibrated for *all* possible T-sequence of states and forecasts. Even if a sender provides honest forecasts ($f_t = p_t$), there is a non-negligible chance that the sequences of forecasts will not be ϵ_T-calibrated. However, as we collect more data the forecasts become close to the empirical distribution of states for almost all sequences of states. This motivates the definition of the (asymptotic) calibration test.

Definition 2 *A forecasting strategy σ is ϵ-calibrated if*

[2]Given multiple optimal actions, we arbitrarily choose to break ties in favour of the sender.

$$\limsup_{T\to\infty}\sum_{f\in F}\frac{|\mathbb{N}_T[f]|}{T}\|\overline{\omega}_T[f]-f\|\leq\epsilon.\quad \mathbb{P}_{\sigma,\mu}-a.s. \tag{4}$$

A forecasting strategy σ is calibrated if it is ϵ-calibrated, for every $\epsilon > 0$.

Intuitively, a forecasting strategy σ is calibrated if the limit empirical distribution of states exactly matches with f for all possible f that were used sufficiently often. We've already seen that the requirement to pass the ϵ_T-calibration test for all rounds T is too demanding. But is it possible for a forecasting strategy to pass the (asymptotic) calibration test? Yes, an *honest* informed sender passes the calibration test almost surely (see Dawid [1982]). Infact, we can form a sequence of error bounds $\{\epsilon_T\}_{T=1}^{\infty}$, for the finite periods, such that $\lim_{T\to\infty}\epsilon_T = 0$ and an *honest* sender only fails the ϵ_T-calibration test in finitely many periods (see Proposition 5 in the appendix). Throughout this paper, we assume the sequence of error bounds $\{\epsilon_T\}_{T=1}^{\infty}$ satisfy this property.

Assumption 3 *The sequence of error margins $\{\epsilon_T\}_{T=1}^{\infty}$ are such that $\lim_{T\to\infty}\epsilon_T = 0$ and an honest sender only fails the ϵ_T -calibration test finitely many times (almost surely).*

One cannot provide a universal sequence of error margins because the rate of convergence depends on the underlying stochastic process μ. This assumption ensures the type I error vanishes and the calibration test does not reject an honest sender. [3] If a sequence of errors did not satisfy this assumption, even an honest sender might be punished infinitely often. To clarify, our results do not depend on the specific sequence of error bounds as long as the assumption 3 is satisfied. Note, this assumption will not be valid in the case of a uninformed sender (in Section 3.3) where nature acts as an adversary and the data generating process itself is not fixed.

The overarching objective of the sender is to find the forecasting strategy that maximizes her long run average payoff:

$$\liminf_{T\to\infty}\frac{\sum_{t=1}^{T}u_S(\omega_t,a_t)}{T}. \tag{5}$$

We call it the *optimal forecasting strategy*. If the punishment cost c is small, the optimal strategy can belong outside the class of calibrated strategies. Our motivation is to use the calibration test as the credibility criterion for the sender. So, we focus on high punishment costs where the optimal forecasting strategy needs to pass the calibration test.

[3] Type I error is the event of rejecting the null hypothesis given that it is true. In our case, it refers to rejecting an honest informed sender.

Proposition 1 *There exists a punishment cost* $\bar{c} = \max_{p \in \Delta\Omega, a \in A} \mathbb{E}_p[u_S(\omega, a) - u_S(\omega, \hat{a}(p))]$ *such that for all* $c \geq \bar{c}$ *the optimal forecasting strategy has to pass the calibration test.*

If the forecasts of the sender are not calibrated, it implies that she failed the ϵ_T-calibration test in infinite rounds along some possible play. So, the sender incurs the punishment cost c in these rounds. The proposition tells us that if the punishment cost is sufficiently high, the sender can always do better by providing honest forecasts and only getting punished in finitely many rounds. This implies that the calibration test (which does not depend on the error margins) is a necessary criterion for the sender. The calibration test imposes limitations on the distribution of forecasts a sender can implement without being punished.

Assumption 4 *The punishment cost* $c \geq \overline{c}$.

Thus, to maximize her long run average payoff the sender has to pass the calibration test, i.e., equation (4) needs to hold.

3 Main Results

Our main result is an equivalence result between the dynamic forecasting game and a static persuasion problem. First, we introduce the persuasion problem. Then, we present our main results on how the dynamic forecasting game can be reduced to a specific persuasion problem.

3.1 Persuasion problem

We consider a static model of persuasion between a sender and a receiver. The state space is the simplex $\Delta\Omega$ (where Ω is finite). The players have an (atomic) prior measure $P \in \Delta(\Delta\Omega)$. The sender commits to a signalling policy $\pi : \Delta\Omega \to \Delta F$. [4] Each signal realization $s \in F$ results in a posterior belief about the state and consequently a posterior mean $q \in \Delta\Omega$. We assume the receiver's optimal action only depends on his posterior mean as his action depends on his belief over the underlying state Ω. [5] The players' utility when the posterior mean is q is given by $\hat{u}_i(q)$. [6] Overall, the signaling policy results in a distribution over posterior mean $Q \in \Delta(\Delta\Omega)$. A measure $Q \in \Delta(\Delta\Omega)$ is implementable by a signaling policy if and only if Q is a mean-preserving contraction (garbling) of P (see Arieli et al. [2020]). The goal of the sender is to implement the garbled measure Q that maximizes her expected utility.

[4]To clarify, the sender does not have commitment power in the dynamic forecasting game.

[5]This assumption is needed to provide structure to the optimal signaling policy (Arieli et al. [2020], Dworczak and Martini [2019]).

[6]This utility function takes into account the receiver's optimal action, which we assume only depends on his posterior mean $q \in \Delta\Omega$.

Let P be a probability measure on a finite support $\boldsymbol{p} = (p_1,....,p_n) \in \mathbb{R}^{|\Omega|\times n}$ with mass $\boldsymbol{\lambda} = (\lambda_1,...,\lambda_n) \in \mathbb{R}^{1\times n}$ such that $\lambda_i > 0 \quad \forall i \in \{1,...,n\}$ and $\sum_{i=1}^{n} \lambda_i = 1$. A probability measure is defined in terms of the mass and the support $P := (\boldsymbol{\lambda}, \boldsymbol{p})$. Let Q be a probability measure with support on m points $\boldsymbol{q} = (q_1,...,q_m) \in \mathbb{R}^{|\Omega|\times m}$ with respective masses $\boldsymbol{\mu} = (\mu_1,....,\mu_m) \in \mathbb{R}^{1\times m}$.

Definition 5 *A probability measure $Q = (\boldsymbol{\mu}, \boldsymbol{q})$ is a simple mean-preserving contraction (smpc) of $P = (\boldsymbol{\lambda}, \boldsymbol{p})$, if there exists a row-stochastic matrix $G \in \mathbb{R}^{n\times m}$ such that:*

$$\boldsymbol{\lambda} G = \boldsymbol{\mu} \tag{6}$$

$$(\boldsymbol{\lambda p})G = (\boldsymbol{\mu q}) \tag{7}$$

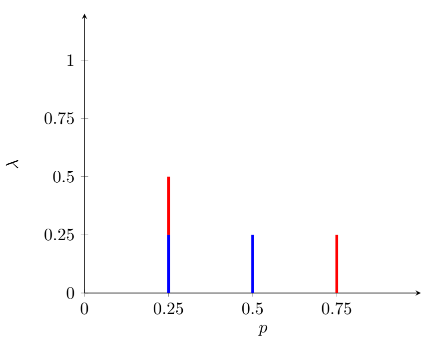
where $\boldsymbol{\lambda p} = (\lambda_1 p_1,...,\lambda_n p_n) \in \mathbb{R}^{|\Omega|\times n}$ and $\boldsymbol{\mu q} = (\mu_1 q_1,...,\mu_m q_m)\mathbb{R}^{|\Omega|\times m}$.

Definition 6 *Q is a garbling (or mean-preserving contraction) of P if there exists a sequence of spmcs $\{Q_m\}_{m=1}^{\infty} \subset G(P)$ that satisfies $Q_m \to_w Q$ (weak convergence). Denote by $G(P)$ the set of all garblings of P.*

Intuitively, a mean-preserving contractions takes the fraction G_{ij} of mass λ_i at p_i for all i and merge them together to get mass μ_j at q_j. We want to determine the *optimal garbled measure* Q^* that maximizes the sender's expected utility. Given prior measure P and utility function $\hat{u}_S$, the solution to the persuasion problem $(P, \hat{u}_S)$ is given by:

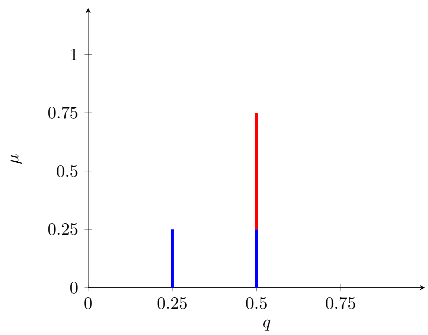
$$Per(P, \hat{u}_S) = \max_{Q\in G(P)} \mathbb{E}_Q[\hat{u}_S] = \max_{Q\in G(P)} \sum_{q\in Supp(Q)} \mu(q)\hat{u}_S(q). \tag{8}$$

For example, consider $\Omega = \{0, 1\}$. Let the set of feasible forecasts $F = \{0.25, 0.50, 0.75\}$ and the utility be given by $\hat{u}_S(0.25) = 0, \hat{u}_S(0.50) = 1, \hat{u}_S(0.75) = 0$. The prior measure P is given below:



$$P = \begin{bmatrix} \lambda \\ p \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{bmatrix}$$

For this example, the optimal garbled measure Q^* corresponds to merging the forecasts $p = 0.25$ and $p = 0.75$ with equal weight $\lambda = 0.25$ to $p = 0.50$ (the merging is represented by red lines in the figures).



$$Q^* = \begin{bmatrix} \mu \\ q \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{bmatrix}$$

Now, we provide a sufficient condition when the persuasion problem can be solved using the concavification approach [Kamenica and Gentzkow, 2011] applied to a restricted domain.

Proposition 2 *If $Supp(P)$ is affinely independent, then the solution of the persuasion problem $(P, \hat{u}_S)$ is given by*

$$Per(P, \hat{u}_S) = Cav\ \hat{u}_S\Big|_C(\mathcal{B}(P)) \tag{9}$$

where, $Cav\ \hat{u}_S\Big|_C$ denotes the concave envelope of $\hat{u}_S$ restricted to domain $C = Co(Supp(P)))$ and $\mathcal{B}(P) = \sum_{i=1}^{n} \lambda_i p_i$ denotes the barycenter of the measure P. [7]

The proof relies on the simple characterization of $G(P)$ when $Supp(P)$ is affinely independent. The feasibility condition corresponds to Bayes plausibility in the restricted domain (see Proposition 6 in the appendix). Hence, the solution is given by the concave envelope restricted to $Co(Supp(P))$. A perfectly informed sender knows the true state $\omega \in \Omega$ and can commit to a signaling policy as a function of each state $\omega \in \Omega$. This corresponds to the standard Bayesian persuasion model [Kamenica and Gentzkow, 2011].

Corollary 7 *If the sender is perfectly informed, the solution of the persuasion problem $(P, \hat{u}_S)$ is given by*

$$Per(P, \hat{u}_S) = Cav\ \hat{u}_S(\mathcal{B}(P)). \tag{10}$$

Whitmeyer and Whitmeyer [2021] show that the persuasion problem in equation (8) attains its maximum at an extreme point of $G(P)$. Thus, if $|Supp(P)| = n$, it suffices to restrict our search to measures Q with $|Supp(Q)| \leq n$. Dworczak and Martini [2019] and Arieli et al. [2020] study the persuasion problem for the case of non-atomic measures P and when $\Omega = \{0, 1\}$. They show that it is sufficient to restrict the search of optimal signaling policy to bi-pooling policies. Finally, in Section C, we show the persuasion problem can be modeled in an equivalent way in terms of Blackwell experiments (mean-preserving spreads).

[7] $Co(A)$ refers to the convex hull of set A.

3.2 Optimal Forecasting Strategy

In this subsection, we consider an informed expert, who knows the data-generating process. We provide a necessary and sufficient condition for a forecasting strategy to be calibrated. For the class of stationary ergodic processes, the distribution of forecasts has to be a garbling of the distribution of conditionals. A garbling can be thought of as merging low-value and high-value forecasts with appropriate weights such that the overall forecast is calibrated. Finally, we characterize the optimal forecasting strategy.

Let's define the distribution of conditionals $C_\mu \in \Delta(\Delta\Omega)$:

$$C_\mu(p) = \lim_{T\to\infty} \frac{\sum_{t=1}^{T} \mathbf{1}_{\{p_t = p\}}}{T} \quad \text{(if limit exists)} \tag{11}$$

where, $p_t = \mu(\cdot \mid \omega_1, ..., \omega_{t-1}) \in \Delta\Omega$.

Note, p_t and C_μ is a random variable and depends on the realization of ω^t and ω^∞ respectively. Given a forecasting strategy σ, one can define the distribution of forecasts $F_{\mu,\sigma} \in \Delta(\Delta\Omega)$:

$$F_{\mu,\sigma}(f) = \lim_{T\to\infty} \frac{\sum_{t=1}^{T} \mathbf{1}_{\{f_t = f\}}}{T} \quad \text{(if limit exists)} \tag{12}$$

where, f_t is chosen randomly according to $\sigma(f_1, \omega_1 ..., \omega_{t-1})$. The distribution of conditionals is also the distribution of forecasts by a honest sender ($p_t = f_t$).

For a general stochastic process, an informed sender can pass the calibration test by reporting honestly but it becomes intractable to characterize the set of feasible outcomes that passes the calibration test. We avoid such situations by focusing on the class of stationary and ergodic processes. The following example illustrates the importance of our assumption.

Example 8 *Let μ be a finite Markov chain that is aperiodic and irreducible. Let Q and π^* denote the transition matrix and the invariant distribution respectively. From the Ergodic theorem, we know that the limiting distribution of states converges to the invariant distribution π^* for any initial distribution π_0.*

$$\lim_{n\to\infty} \sum_{i\in\Omega} \pi_0(i) Q^n(i,j) = \pi^*(j) \quad \forall \pi_0 \in \Delta\Omega.$$

The long-run relative frequency of visiting a state i is $\pi^*(i)$ along all sample paths and induces the conditional distribution $Q(\cdot \mid i)$. Thus, the distribution of conditionals C_μ exists and is constant μ-a.s. Each support point $p_i = Q(\cdot \mid i)$ has mass $\lambda_i = \pi^*(i)$ respectively, where $i \in \Omega$. An informed sender, who knows the transition matrix, knows the (constant) distribution of conditionals C_μ.

Recall that a stochastic process $\{\omega_n\}_{n\in\mathbb{N}}$ is stationary if, for any $k \in \mathbb{N}$, the joint distribution of the $k-$tuple $(\omega_n, \omega_{n+1}, ..., \omega_{n+k-1})$ does not depend on n. Let $\mu \in \Delta\Omega^\infty$ denote the distribution of the stationary process. Let $T : \Omega^\infty \to \Omega^\infty$ be the shift transformation given by $T(\omega)_n = \omega_{n+1}$ for all $n \in \mathbb{Z}$. Let $\mathcal{I}$ denote the $\sigma-$algebra of all invariant Borel sets for the transformation T. The stationary process $\{\omega_n\}_{n\in\mathbb{N}}$ is ergodic if $\mathcal{I}$ is trivial, that is, $\mathbb{P}(A) \in \{0,1\}$ for all $A \in \mathcal{I}$. All the statistical properties can be deduced from a single, long realization of the process. This makes the search for the optimal forecasting strategy tractable.

Recall that the indirect utility $\hat{u}_i(f)$ is the expected payoff when the states are drawn according to f and the receiver plays the optimal action $\hat{a}(f)$. Given a calibrated strategy, the empirical distribution of states conditional on forecast f exactly match with the forecast f. The receiver plays the action $\hat{a}(f)$ on (almost) all periods when the forecast f was announced. Thus, for a calibrated strategy, the long run average payoff for player i when the forecast is f equals $\hat{u}_i(f)$. Now, we state our main result that characterizes the optimal forecasting strategy.

Theorem 9 *For a stationary ergodic process μ, the optimal forecasting strategy is given by the optimal signaling policy of the persuasion problem $(C_\mu, \hat{u}_S)$.*

Proof. First, we characterize the set of feasible outcomes for the class of stationary ergodic processes. To do so, we show that the distribution of conditionals C_μ converges and is constant for almost every play (see Lemma 18 in the appendix). Thus, an informed sender knows C_μ, which corresponds to the prior measure in the persuasion problem. Using this result we present, in Lemma 10, a sufficient and necessary condition for a forecasting strategy to be calibrated. The feasible distributions of forecasts are precisely the set of garblings of the distribution of conditionals C_μ.

Lemma 10 *For a stationary ergodic process μ, if a forecasting strategy σ is calibrated then $F_{\mu,\sigma} \in G(C_\mu)$. On the other hand, for any $Q \in G(C_\mu)$, we can construct a calibrated strategy σ such that $F_{\mu,\sigma} = Q$.*

Lemma 10 outlines the set of feasible distribution of forecasts given the forecasting strategy has to pass the calibration test. Given the distribution of conditionals C_μ, the sender can implement any garbled measure $Q \in G(C_\mu)$. This includes the solution to the persuasion problem $(C_\mu, \hat{u}_S)$. Infact, any $Q \in G(C_\mu)$ can be implemented using a stationary strategy that depends on the conditional of that round. This ensures any sequence of error bounds $\{\epsilon_T\}_{T=1}^\infty$ chosen by a honest sender would also work for any calibrated strategy. The sender only fails the ϵ_T-calibration test and faces punishment in finitely many rounds.

We now construct the optimal forecasting strategy. Given the distribution of conditionals C_μ , let $\tau^* : \Delta\Omega \to \Delta F$ and $Q^* = (\mu, f)$ denote the optimal signaling policy and optimal garbled measure for the persuasion problem $(C_\mu, \hat{u}_S)$. Consider the forecasting strategy $\sigma_t^*(f_t = f \mid p_t = p) = \tau^*(f \mid p)$ for all t. At any period t, the strategy only depends on the current conditional p_t and thus is stationary. We show that it achieves the optimal persuasion payoff.

$$= \liminf_{T\to\infty} \frac{\sum_{t=1}^{T} u_S(\omega_t, a_t)}{T} \tag{13}$$

$$= \liminf_{T\to\infty} \frac{\sum_{f\in Supp(Q^*)} \sum_{p\in D} \tau^*(f \mid p) \sum_{t=1}^{T} \mathbf{1}_{\{p_t=p\}} u_S(\omega_t, \hat{a}(f))}{T} \tag{14}$$

$$= \sum_{p\in D} C_\mu(p) \sum_{\omega} p(\omega)\tau^*(f \mid p) u_S(\omega, \hat{a}(f)) \tag{15}$$

$$= Per(C_\mu, \hat{u}_S) \tag{16}$$

■

Finally, we provides a sufficient and necessary condition for a forecasting strategy to pass the calibration test for any stochastic process μ.

Lemma 11 *A forecasting strategy σ passes the calibration test if and only if*

$$\limsup_{T\to\infty} \sum_{f\in F} \frac{|N_T[f]|}{T} \|f - \sum_{p\in D} p\mu_T(f,p)\| = 0 \quad \mu\text{-a.s. where, } \mu_T(f,p) = \frac{\sum_{t=1}^{T} 1_{\{p_t=p, f_t=f\}}}{\sum_{t=1}^{T} 1_{\{f_t=f\}}} \tag{17}$$

The term $\mu_T(f,p)$ can be interpreted as relative weight on conditional p given forecast f. In the long run, combining all conditionals p with their respective weights $\mu_T(f,p)$ equals forecast f. This resembles the definition of a garbling (mean-preserving contraction). But even though equation (17) holds for any stochastic process, the distributions C_μ and $F_{\mu,\sigma}$ might not be well defined. Due to this, it becomes intractable to characterize the set of feasible outcomes for a calibrated forecasting strategy.

3.3 Uninformed Expert

In this section, we consider an *uninformed* expert who does not know the data-generating process. Foster and Vohra [1998] show that an uninformed expert can come up with a calibrated strategy for *any* stochastic process. We ask: what is the maximum payoff that the expert can guarantee? How does this utility compare with that of an informed expert? First, we tackle the situation where nature acts as an adversary. Then, we analyze the situation when the play of nature is stationary and ergodic but still unknown to the sender.

The sender, who interacts with nature, has the objective of maximizing her long run average payoff while passing the calibration test. As is standard in the literature, we first model nature as an adversary trying to prevent the sender from doing so. The receiver's behavioural assumption constrains the sender's ability to persuade. Formally, it implies that the sender has to pass the (asymptotic) calibration test. Given the sender passes the calibration test, the receiver plays according to her forecast almost surely.

At every period t, the sender and nature simultaneously choose $f_t \in F$ and $\omega_t \in \Omega$ respectively. A strategy τ for nature is a mapping from the set of all possible past histories to the set of mixed states, i.e, $\tau : \bigcup_{t\geq 0}(F \times \Omega)^{t-1} \to \Delta\Omega$.

Let $\overline{\omega}_T \in \Delta\Omega$ denote the *empirical distribution* of states by time T,

$$\overline{\omega}_T = \frac{1}{T}\sum_{i=1}^{T}\delta_{\omega_i} \tag{18}$$

If the sender knew the empirical distribution $\overline{\omega}_T$ will equal to $p \in \Delta\Omega$ in advance, she could repeatedly send the fixed forecast p and pass the calibration test. In particular, the sender would be able to achieve $\hat{u}_S(p)$.

In the online learning literature (see Mannor et al. [2009]), this is referred as the *reward-in-hindsight function*: the highest reward (or payoff) the sender could have achieved, while satisfying the constraints, had she known natures choices in advance. [8]Our characterization of what a sender can attain will be as a function of the empirical distribution of states $\overline{\omega}_T \in \Delta\Omega$.

Definition 12 *A function $a : \Delta\Omega \to \mathbb{R}$ is attainable by the sender if there exists a forecasting strategy σ such that for every strategy τ of nature:*

1. $\liminf_{T\to\infty}\left(\frac{\sum_{t=1}^{T} u_S(\omega_t, a_t)}{T} - a(\overline{\omega}_T)\right) \geq 0 \quad a.s$
2. $\limsup_{T\to\infty}\sum_{f\in F}\frac{|\mathbb{N}_T[f]|}{T}\|\overline{\omega}_T[f] - f\| \leq 0. \quad a.s$

where the almost sure convergence is with respect to the probability measures induced by the strategies σ and τ.

The first condition states that the long-run average payoff is higher than the function $a(p)$ almost surely, where p is the limit empirical distribution of states. While the second condition states that the forecasting strategy passes the calibration test.

In the next theorem, we show that it possible to attain the closed convex hull of the reward-in-hindsight function. Denote by $\overline{Co}(f)$ the closed convex hull[9]of function f. We have $\overline{Co}\ \hat{u}_S(p) \leq \hat{u}_S(p)$ for all $p \in \Delta\Omega$. Furthermore, the function $\overline{Co}\ \hat{u}_S(p)$ is continuous on $\Delta\Omega$. In contrast to the concave envelope, which corresponds to the best garbling

[8]Mannor et al. [2009] show that the reward-in-hindsight function is not attainable *in general*. However, using approachability for closed and convex sets, they show that the closed convex hull of the reward-in-hindsight function is attainable.

of a perfectly informed sender, the closed convex hull corresponds to the sender's worst garbling in the persuasion problem. We show it is the highest function that an uninformed sender can attain.

Theorem 13 *For an uninformed sender, $\overline{Co}\ \hat{u}_S(p)$ is the highest attainable function, where $p \in \Delta\Omega$ is the empirical distribution of states.*

Sketch of Proof. We use the techniques of approachability to prove the theorem. First, we show the function $\overline{Co}\ \hat{u}_S(p)$ is attainable and then show it is the highest attainable function. At every period, the actions of the sender and nature result not only in sender's payoff but also a *calibration cost*. Given forecast $f \in F$ and state $\omega \in \Omega$ the calibration cost is given by:

$$c(f,\omega) = (\underline{0}, ..., f - \delta_\omega, ..., \underline{0}) \in \mathbb{R}^{\infty|\Omega|}. \tag{19}$$

It is a vector of infinite elements of size $\mathbb{R}^{|\Omega|}$ with one non-zero element (at the position for f) while the rest are equal to $\underline{0} \in \mathbb{R}^{|\Omega|}$. The calibration condition (2) can be rewritten as follows: the average of the sequence of vector-valued calibration costs $c(f_t, \omega_t)$ converges to the set E almost surely, where

$$E = \{x \in \mathbb{R}^{\infty|\Omega|} : \sum_{f \in F} ||\underline{x}_f|| \leq 0\} \tag{20}$$

Thus, the goal of the sender it to maximize the average payoff , such that the average calibration cost converge to E pathwise, i.e., $\limsup_{t\to\infty} \text{dist}(\frac{\sum_{t=1}^{T} c(f_t,\omega_t)}{T}, E) \to 0$ a.s. where, $dist(.)$ is the Euclidean distance. The crux of the proof is to combine the sender's payoff and calibration cost to form a vector-valued payoff. Then, we use the dual condition of approachability to show the closed and convex function $\overline{Co}\ \hat{u}_S(p)$ is attainable. The idea of the proof is rather straightforward, but the proof is rather technical. We first prove the result for $\epsilon-$calibration and then use the "doubling trick" (see Mannor and Stoltz [2010]) for calibration. Convergence rate results follow from approachability theory and are used to determine the sequence of error margins $\{\epsilon_T\}_{T=1}^{\infty}$ (see appendix for the complete proof).

To show it is also the highest attainable function, we construct a strategy for nature that prevents the sender from attaining any function higher than $\overline{Co}\ \hat{u}_S(p)$) without failing the calibration test.

Let α_l and p_l denote the weight and support of the closed convex hull, i.e., $\overline{Co}\ \hat{u}_S(p) = \sum_{l=1}^{k} \alpha_l \hat{u}_S(p_l)$). Nature plays in a sequence of k blocks, where the relative size of each

[9] Given a function $f : X \to \mathbb{R}$, over a convex domain X, its closed convex hull is the function whose epigraph is $\overline{Co}(\{(x,r)) : r \geq f(x)\})$ where, $\overline{Co}(X)$ is the closed convex hull of set X.

block l is α_l. In block l, nature plays the fixed action $p_l \in \Delta\Omega$. Given any i.i.d process with distribution p_l, the only calibrated strategy for sender is to repeatedly forecast q_l almost surely. The crucial step in the proof is to show that the sender also needs to pass the calibration test within each block to pass the overall calibration criterion. If not, then nature can use a punishment strategy that ensures that the sender fails the overall calibration test. ■

So far, we have considered nature as an adversary preventing the uninformed sender from passing the calibration test. We assume nature can condition her actions based on the sender's previous forecasts. Can an uninformed sender do better when nature is non-adversarial and states are drawn according to a fixed process? We show that this indeed possible. If the stochastic process is stationary and ergodic, the sender is able to (approximately) attain the reward-in-hindsight function $\hat{u}_S(p)$. This allows us to compare the attainable payoffs of an informed and uninformed sender.

For this result, we assume the set of forecasts F is finite. Fix $\epsilon > 0$, we assume the set of feasible forecasts F forms a regular ϵ-grid:

$$F = \{\sum_{\omega\in\Omega} n_\omega \delta_\omega \in \Delta\Omega \mid n_\omega \in \{0, \frac{1}{L}, ..., 1\} \text{ and } \sum_{\omega\in\Omega} n_\omega = 1\} \tag{21}$$

where, $L = \lceil \frac{\sqrt{|\Omega|-1}}{2\epsilon} \rceil \in \mathbb{N}$. [10] This ensures (generically) for any $p \in \Delta\Omega$ there exists a unique pure forecast $f \in F$ such that $\|f - \sum_{\omega\in\Omega} p(\omega)\delta_\omega\| \leq \epsilon$. Let us denote by $f^*(p)$ as the pure forecast that belongs to the ϵ-neighbourhood of $p \in \Delta(\Omega)$.

Lemma 14 *For a stationary ergodic process, the function $\hat{u}_S(f^*(p))$ is attainable, where $p \in \Delta\Omega$ is the empirical distribution of states.*

Sketch of Proof. The proof uses the concept of *opportunistic approachability* (see Bernstein et al. [2014]). A set that is not approachable in general can be approached if nature plays favourably or in a non-adversarial manner. Nature can no longer use a punishment strategy that forces the sender to fail the calibration test on deviation. Given the process is stationary and ergodic, any play is empirically restricted to a neighbourhood around p. The sender comes up with a forecasting strategy that learns the (limit) empirical distribution and thus attains the function $\hat{u}_S(f^*(p))$. ■

Note, the sender does not need to know the (limit) empirical distribution p, nor that the process is stationary and ergodic beforehand to implement the forecasting strategies. If nature's strategy turns out to be favourable the sender attains the favourable payoff, if not she still attains the lower benchmark of the general setting. [11] To summarize, an uninformed sender can attain $\overline{Co}\ \hat{u}_S(p)$ in general and approximately $\hat{u}_S(p)$ if the process is stationary and ergodic, where p is the empirical distribution of states.

[10] where $\lceil x \rceil$ denotes the smallest integer greater than or equal to x.

[11] Bernstein et al. [2014] show this property holds even in general settings.

3.4 Application: Financial App

In this section, we consider a financial app (sender) that provides forecasts about a binary state to a receiver over time. The receiver can also sequentially acquire signals on the app by paying a cost (see Wald [1945]). Our focus is on the app's utility and what it can attain. We solve the app's forecasting problem subject to the calibration constraint and show the app's optimal strategy is to provide accurate but less precise forecasts as compared to honest reporting.

Each day, a receiver has to decide whether to invest in the financial market or not. The market can be in two states: $\Omega = \{H, L\}$. The receiver only wants to invest if the state is H. The state changes through time and follows a Markov chain with the transition matrix: $T(H \mid H) = T(L \mid L) = 0.95$.

The timing and structure of the forecasting and the information acquisition is as follows. At the start of each day $n \in \mathbb{N}$, the app (sender) forecasts the state of that day $f_n \in \Delta\Omega$. Let f_n simply denote $\mathbb{P}(\omega_n = H) \in [0, 1]$. The receiver checks for the credibility of the app using the calibration test. If the app fails the test, he does not use the app. Else, he uses the app's forecast as his *prior belief* $p_n \in \Delta\Omega$ for the information acquisition process.

The receiver sequentially acquires signals until he is certain that the state is either H or L. The receiver continuously observes a process $(Z_t)_{t \geq 0}$. The change of the process Z is the sum of the state plus a noise term which is the increment of a Brownian motion $(W_t)_{t \geq 0}$.

$$dZ_t = \theta dt + \sigma dW_t. \tag{22}$$

The receiver pays a constant cost c_0 to the app for every instant he observes the signal. The cost can be interpreted as the time spent on the app. The amount of signal a receiver acquires corresponds to his action. Morris and Strack [2019] characterize the total cost for sequentially acquiring signals. They show that the ex-ante cost of information is equal to the expected change in the prior and posterior log-likelihood ratio.

$$C(G) = \begin{cases} \frac{c_0 \sigma^2}{2} \int_0^1 L(q) dG(q) - L(p) & \text{if } \mathcal{B}(G) = p \\ \infty & \text{else} \end{cases}$$

where, $L(q) = q \log(\frac{1}{1-q}) + (1-q) \log(\frac{1-q}{q})$ corresponds to the log-likelihood ratio for belief q. Also, let p and G denotes the prior belief and the posterior distribution respectively.

Let us denote by e_L and e_H ($0 \leq e_L \leq e_H \leq 1$) the threshold beliefs for state L and H respectively. The receiver acquires signals till his belief crosses either threshold. Let G_p^* denote the probability distribution with support $e_L = 0.05$ and $e_H = 0.95$ and barycenter p (whenever feasible). Starting with prior belief p, the app's utility from signal acquisition (see Figure 1) is given by :

$$\hat{u}_S^{sig}(p) = \begin{cases} C(G_p^*) & e_l \leq p \leq e_H \\ 0 & \text{otherwise} \end{cases}$$

The app also gains utility from announcing precise forecasts, which is interpreted as reputation. The utility gain is higher when the forecast is more precise. The app's utility gain in reputation from announcing forecast p (see Figure 2) is given by:

$$\hat{u}_S^{rep}(p) = \kappa(p - 0.5)^2 \tag{23}$$

for some constant $\kappa > 0$. The app's total utility $\hat{u}_S(p)$ is the sum of these two components:

$$\hat{u}_S(p) = \hat{u}_S^{sig}(p) + \hat{u}_S^{rep}(p) \tag{24}$$

The receiver always prefers accurate and precise forecasts to take the optimal action and minimize the information acquisition cost. But the app's preference is non-trivial. On the one hand, revealing precise information increases the app's reputation. While on the other, a precise forecast reduces the amount of signals the receiver acquires on the app. On average, the receiver spends lesser time on the app. The app's goal is to find the optimal forecasting strategy that maximizes its long run average payoff subject to passing the calibration test.

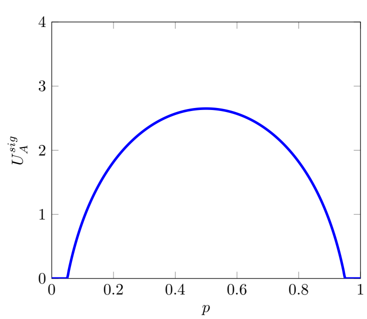


Figure 1: $\hat{u}_S^{sig}(p)$

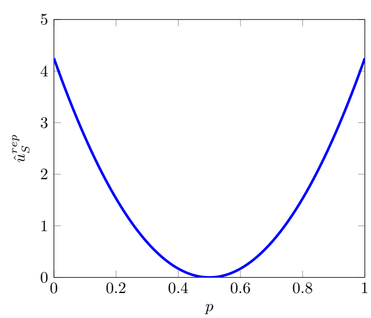


Figure 2: $\hat{u}_S^{rep}(p)$

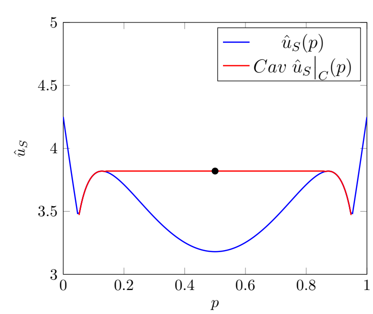


Figure 3: $\hat{u}_S(p)$

Using Theorem 9, we know the optimal calibrated strategy corresponds to the solution to the persuasion problem $(C_\mu, \hat{u}_S)$. Given $Supp(C_\mu) = \{0.05, 0.95\}$ is affinely independent, we can use Proposition 2 to find the optimal garbled measure and the optimal forecasting strategy (see Figure 3, where $C = [0.05, 0.95]$ and $\mathcal{B}(C_\mu) = 0.5$). The solution is given by:

$$Per(C_\mu, \hat{u}_S) = Cav\ \hat{u}_S\big|_{[0.05,0.95]}(0.5) \tag{25}$$

The measure C_μ and the optimal garbled measure Q^* are given by:

$$C_\mu = \begin{bmatrix} \lambda \\ p \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{5}{100} & \frac{95}{100} \end{bmatrix} \qquad Q^* = \begin{bmatrix} \mu \\ q \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{15}{100} & \frac{85}{100} \end{bmatrix}$$

The optimal (stationary) forecasting strategy σ_*, for all n, is given by:

$$\sigma_*(f_n = 15\% \mid p_n = 5\%) = \frac{8}{9} \qquad \sigma_*(f_n = 85\% \mid p_n = 5\%) = \frac{1}{9} \tag{26}$$

$$\sigma_*(f_n = 15\% \mid p_n = 95\%) = \frac{1}{9} \qquad \sigma_*(f_n = 85\% \mid p_n = 95\%) = \frac{8}{9} \tag{27}$$

If the app was perfectly informed and knew the state $\omega \in \{H, L\}$, it would want to reveal it honestly at the start of each day. This corresponds to the standard Bayesian persuasion solution (Kamenica and Gentzkow [2011]), i.e., the concave envelope of the utility function (with no restriction on the domain). Even though the receiver spends no time on the app, the utility from reputation is unmatched. But given the distribution C_μ, the sender cannot announce the true states without failing the calibration test. Thus, for a partially informed app ($p_n = \{5\%, 95\%\}$), the optimal strategy is to announce accurate but garbled forecasts ($f_n = \{15\%, 85\%\}$).

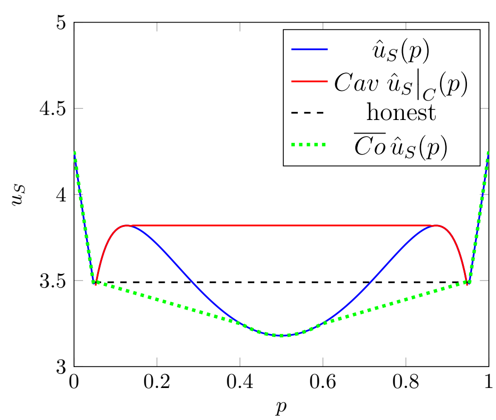


Figure 4: Financial App

Our main results are represented in Figure 4. What could the app achieve if it was uninformed? The app would be able to attain the indirect utility function $\hat{u}_S(p)$ (blue line). This corresponds to the no information signaling policy of the persuasion problem.

What could the app guarantee if the stochastic process was not stationary and ergodic? For a general process, an informed app could always pass the calibration test by reporting honestly (black line). On the other hand, an uninformed app would be able to approximately attain the closed convex hull of the utility function, i.e., $\overline{Co}\ \hat{u}_S(p)$ (green line). This corresponds to the worst garbling of a perfectly informed sender in the persuasion problem. The difference between these two functions gives us the value of information. It quantifies what an uninformed sender is willing to pay to become informed.

4 Forecasting against regret minimizers

In this section, we consider receivers who use regret minimization as a heuristic for decision-making instead of the calibration test. Apart from that, the setting remains the same as in Section 2. At each period, the sender sends a forecast and the receiver uses a regret minimizing strategy to take actions. Regret is used as an exogenous criterion to evaluate a strategy in non-Bayesian environments. It measures whether a sequence of predictions is accurate or not. The goal of the sender is to maximize her long-run average payoff. Given the close connection between regret and calibration (see Perchet [2014]), we compare what the sender can guarantee in each case. We show that the sender can always guarantee the persuasion benchmark as in the case of the calibration test. We show that, in some cases, the sender can guarantee much more when the receiver uses a mean-based learning algorithm, a natural class of regret minimizing strategies.

A key distinction when playing against a no-regret learner is that the forecast f is no longer a probabilistic statement but simply a message. It acts as *context* for the regret minimizer. Given our goal is to compare the analysis with the calibration test, we assume the receiver minimizes his regret with respect to each forecast f. The receiver has no regret with respect to forecast f if on the set of stages when the forecast was f he cannot shift to a fixed action $a^* \in A$ repeatedly and achieve a higher payoff. This is a special case of *contextual regret*, where the forecast acts as context or additional information the receiver has in each sequential game (the receiver takes action after the sender has sent her forecast). Let us denote by $\overline{u}_{R,T}[f]$ the average receiver's payoff upto stage T when the forecast was f, i.e.,

$$\overline{u}_{R,T}[f] = \frac{\sum_{t \in \mathbb{N}_T[f]} u_R(\omega_t, a_t)}{|\mathbb{N}_T[f]|}. \tag{28}$$

Definition 15 *The receiver has no regret with respect to forecast f if*

$$\limsup_{T \to \infty} \frac{|\mathbb{N}_T[f]|}{T} \left(\max_{a^* \in A} u_R(\overline{\omega}_T[f], a^*) - \overline{u}_{R,T}[f] \right) \leq 0. \tag{29}$$

The following proposition provides a justification for using the calibration test as a heuristic.

Proposition 3 *The receiver has no regret with respect to any forecast if he follows the recommendations of a calibrated forecasting strategy.*

Proof. Fix any calibrated strategy and assume the receiver plays according to the forecast $a_t = \hat{a}(f_t)$ almost surely. The receiver's regret with respect to the forecast f is given by:

$$=\limsup_{T\to\infty} \frac{|\mathbb{N}_T[f]|}{T}\Big(\max_{a^*\in\Delta A} u_R(\overline{\omega}_T[f],a^*) - u_R(\overline{\omega}_T[f],\hat{a}(f))\Big) \tag{30}$$

$$=\limsup_{T\to\infty} \frac{|\mathbb{N}_T[f]|}{T}\Big(\max_{a^*\in\Delta A} \sum_{\omega\in\Omega} f(\omega)[u_R(\omega,a^*) - u_R(\omega,\hat{a}(f))]\Big) \leq 0 \quad \text{(using calibration)} \tag{31}$$

■

For a stationary ergodic process, the optimal persuasion payoff is the sender's upper bound when she has to pass the calibration test. We show that she can also guarantee this benchmark when facing a regret minimizer.

Proposition 4 *For a stationary ergodic process μ, the sender can guarantee $Per(C_\mu, \hat{u}_S)$ against a regret minimizer.*

For any $Q \in G(C_\mu)$, the sender can repeatedly use the signaling policy $\sigma(f \mid p)$ that implements Q in the persuasion problem. Given any forecast f, if the receiver uses any action $a \in \hat{a}(f)$ on non-negligible fraction of rounds then he has positive regret.

We now show there are examples where the sender can guarantee much more. We follow the approach used by Deng et al. [2019] and Braverman et al. [2018] by focusing on the natural class of *mean-based learning algorithms*. This class of no-regret strategies includes Multiplicative Weights algorithm, the Follow-the-Perturbed-Leader algorithm, and the EXP3 algorithm. Intuitively, mean-based strategies play the action that historically performs the best. For the next result, we assume the receiver uses a mean-based learning algorithm for a T−period game, where $T >> 0$.

Definition 16 *Let $\sigma_{a,t} = \sum_{s=1}^{t} u_R(\omega_s, a)$ be the cumulative reward for action a for the first t rounds. An algorithm is mean-based if whenever $\sigma_{a,t} < \sigma_{b,t} - \gamma T$ for some $b \in A$, the probability to play action a on round t is at most γ. An algorithm is mean-based if it is γ-mean-based for some $\gamma = o(1)$.*

The next theorem provides an instance where a sender can attain a payoff greater than the optimal persuasion payoff against a mean-based learner.

Theorem 17 *There exists a game where the sender can guarantee $V > Per(C_\mu, \hat{u}_S)$ against a mean-based learner.*

Proof. Consider the following payoff matrix, which represents $u_S(\omega, a)$ and $u_R(\omega, a)$ respectively.

	a_1	a_2	a_3	a_4
ω_1	$(2,8)$	$(0,7)$	$(4,3)$	$(2,0)$
ω_2	$(2,0)$	$(4,3)$	$(0,7)$	$(2,8)$

The receiver's optimal action depends on the distribution of states (let $p = \mathbb{P}(\omega_2)$). It is optimal to play a_1 when $p \le 0.25$, to play a_2 when $0.25 \le p \le 0.5$, to play a_3 when $0.5 \le p \le 0.75$ and to play a_4 otherwise (see Figure 5). This induces the indirect utility $\hat{u}_R$ and $\hat{u}_S$ (see Figure 5 and 6).

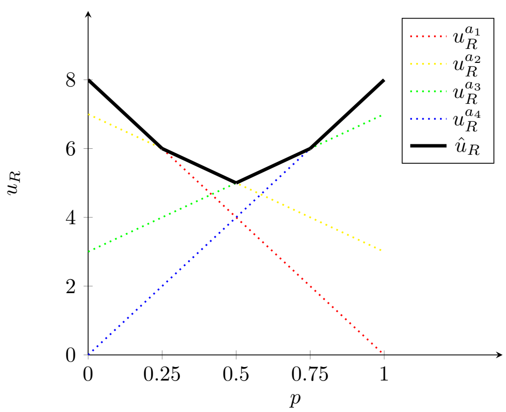


Figure 5: Receiver's indirect utility: $\hat{u}_R$

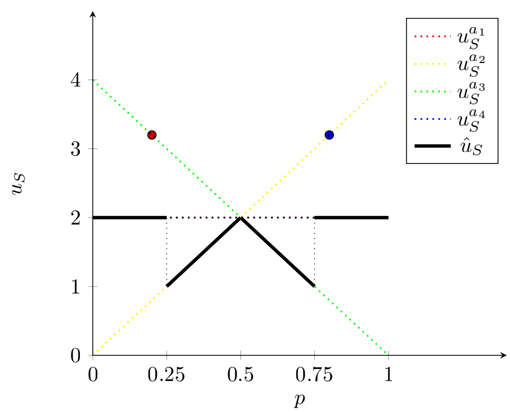


Figure 6: Sender's indirect utility: $\hat{u}_S$

Consider a Markov chain with transition matrix $T(\omega_1 \mid \omega_1) = T(\omega_2 \mid \omega_2) = 0.8$. Thus, the distribution of conditionals C_μ has equal weights on the support 20% and 80%. The optimal garbled measure $Q^* \in G(C_\mu)$ is to either report honestly or babble, which implies $Per(C_\mu, \hat{u}_S) = 2$ (see Figure 6). We now describe a strategy that will guarantee the sender a utility $V > 2$.

Notice, given the receiver uses mean-based strategy with respect to forecast f, we have

$$\sigma^f_{a,t} = \left[u_R(\overline{\omega}_t[f], a)\right] \times t. \tag{32}$$

As the receiver is a mean-based learner, he (with a high probability) plays the best response to the current distribution of states $\overline{\omega}_t[f]$ at stage t.

The strategy comprises of two forecasts (or messages) l and h. For the first $\frac{T}{2}$ stages, the sender announces l when $p_t = 20\%$ and announces h when $p_t = 80\%$. Based on the mean-based algorithm, the receiver responds by playing action a_1 and a_4 (with a high probability) on forecasts l and h respectively. As $T >> 0$, the sender's average payoff ≈ 2 in the first $\frac{T}{2}$ stages.

Now, for the remaining $\frac{T}{2}$, the sender switches the forecast, i.e., she announces l when $p_t = 80\%$ and announces h when $p_t = 20\%$. This forces the empirical distribution $\overline{\omega}_t[l]$ starting from 20% to increase till it finally reaches 50%. But as soon as the empirical distribution crosses 25%, action a_2 has the highest cumulative reward and so the receiver switches to playing action a_2 (with a high probability). Similarly for forecast h, as the empirical distribution decreases from 75%, the receiver starts playing action a_3. This guarantees an average payoff $\approx \frac{173}{55}$ in the last $\frac{T}{2}$ rounds. This is because, for most of the $\frac{T}{2}$ stages, the receiver plays action a_2 while the true distribution of states under l is 80% . The receiver takes an action that is sub-optimal with respect to the distribution of states and so the average payoff does not lie on the graph of indirect utilities $\hat{u}_S$ and $\hat{u}_R$ (see Figure 6). Overall, across the T period the sender is able to guarantee an average payoff $V \approx \frac{283}{110} > 2$. Thus, we have shown that against a mean-based regret learner, the sender is able to guarantee a utility higher than the solution to the persuasion problem. ■

This example holds even in the case when the receiver does not observe the state $\{\omega_t\}_{t\in\mathbb{N}}$ and/or only observes the reward from the action he chose. Using standard tools from bandits problem, such as inverse propensity score estimator, it is possible to build unbiased estimator of the reward of unplayed actions [Bubeck and Cesa-Bianchi, 2012]. The main idea consists in slightly perturbing the decision, say by playing with some small probability at random. This perturbation can actually be quite small, with probability of order $T^{-1/3}$ and even $T^{-1/2}$, so that the actual cost of estimation is negligible as T increases. Those techniques are now rather standard and very well documented [Lattimore and Szepesvári, 2020].

5 Conclusion and Future Work

We studied a dynamic forecasting game, where the sender maximizes her utility given she has to pass the calibration test. For the class of stationary and ergodic processes, we determined the optimal calibrated forecasting strategy. We did this by reducing the dynamic forecasting game in terms of a static persuasion problem. In essence, we showed that the dynamic interaction of a sender and a receiver performing the calibration test substitutes for ex-ante commitment in persuasion models. We compared what an informed and uninformed expert can attain. We also compared regret minimization and the calibration test as heuristics.

Many problems remain open for the setting that we study. In particular, the optimal calibrated strategy for *any* stochastic process. Given this problem might be intractable, we could investigate, for what class of stochastic processes, we can (or cannot) find a calibrated strategy that does better than honest reporting. Also, the complete characterization for attainable payoffs against no-regret learners remains open.

There are natural extensions of the model: we can consider other statistical tests to verify the credibility of the sender. This raises the question of finding tests under which honest reporting is optimal or that minimize the extent of strategic misreporting.

References

Itai Arieli, Yakov Babichenko, Rann Smorodinsky, and Takuro Yamashita. Optimal persuasion via bi-pooling. In *Proceedings of the 21st ACM Conference on Economics and Computation*, EC '20, page 641, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450379755. doi: 10.1145/3391403.3399468. URL https://doi.org/10.1145/3391403.3399468.

Andrey Bernstein, Shie Mannor, and Nahum Shimkin. Opportunistic approachability and generalized no-regret problems. *Mathematics of Operations Research*, 39(4): 1057–1083, 2014.

James Best and Daniel Quigley. Persuasion for the long run. 2020. URL https://ssrn.com/abstract=2908115orhttp://dx.doi.org/10.2139/ssrn.2908115.

David Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 24(2):265–272, 1953. ISSN 00034851.

Mark Braverman, Jieming Mao, Jon Schneider, and Matt Weinberg. Selling to a no-regret buyer. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, EC '18, page 523538, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450358293. doi: 10.1145/3219166.3219233. URL https://doi.org/10.1145/3219166.3219233.

Sébastien Bubeck and Nicolò Cesa-Bianchi. Regret analysis of stochastic and non-stochastic multi-armed bandit problems. *Foundations and Trendső in Machine Learning*, 5(1):1–122, 2012. ISSN 1935-8237. doi: 10.1561/2200000024. URL http://dx.doi.org/10.1561/2200000024.

Vincent P. Crawford and Joel Sobel. Strategic information transmission. *Econometrica*, 50(6):1431–1451, 1982. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/1913390.

A Philip Dawid. The well-calibrated bayesian. *Journal of the American Statistical Association*, 77(379):605–610, 1982.

Yuan Deng, Jon Schneider, and Balasubramanian Sivan. Strategizing against no-regret learners. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32.

Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper/2019/file/8b6dd7db9af49e67306feb59a8bdc52c-Paper.pdf.

Piotr Dworczak and Giorgio Martini. The simple economics of optimal persuasion. *Journal of Political Economy*, 127(5):1993–2048, 2019. ISSN 00223808.

Federico Echenique and Eran Shmaya. You wont harm me if you fool me. Working Papers 1281, California Institute of Technology, Division of the Humanities and Social Sciences, 2007. URL https://EconPapers.repec.org/RePEc:clt:sswopa:1281.

J Elton and T.P Hill. On the basic representation theorem for convex domination of measures. *Journal of mathematical analysis and applications*, 228(2):449–466, 1998. ISSN 0022-247X.

Dean Foster and Rakesh Vohra. Calibration: Respice, adspice, prospice. In *Advances in Economics and Econometrics: Tenth World Congress*, volume 1, pages 423–442. Cambridge University Press, 2013.

Dean P. Foster and Rakesh V. Vohra. Asymptotic calibration. *Biometrika*, 85(2):379–390, 1998. ISSN 00063444. URL http://www.jstor.org/stable/2337364.

Dean P Foster, Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. Complexity-based approach to calibration with checking rules. In *Proceedings of the 24th Annual Conference on Learning Theory*, pages 293–314. JMLR Workshop and Conference Proceedings, 2011.

Ronen Gradwohl and Yuval Salant. How to buy advice with limited instruments. *Available at SSRN 1791750*, 2011.

Yingni Guo and Eran Shmaya. Costly miscalibration. *Available at Theoretical economics*, 2021. URL https://econtheory.org/ojs/index.php/te/article/viewForthcomingFile/3991/28098/1.

Chirag Gupta and Aaditya Ramdas. Top-label calibration and multiclass-to-binary reductions. In *International Conference on Learning Representations*, 2021.

Kallenberg. *Foundations of Modern Probability*. Springer New York, 2002. ISBN 1-4419-2949-5.

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, October 2011. doi: 10.1257/aer.101.6.2590. URL https://www.aeaweb.org/articles?id=10.1257/aer.101.6.2590.

Tor Lattimore and Csaba Szepesvári. *Bandit algorithms*. Cambridge University Press, 2020.

Shie Mannor and Gilles Stoltz. A geometric proof of calibration. *Mathematics of Operations Research*, 35(4):721–727, 2010.

Shie Mannor, John N Tsitsiklis, and Jia Yuan Yu. Online learning with sample path constraints. *Journal of Machine Learning Research*, 10(3), 2009.

Stephen Morris and Philipp Strack. The wald problem and the relation of sequential sampling and ex-ante information costs. 2019. URL https://ssrn.com/abstract=2991567orhttp://dx.doi.org/10.2139/ssrn.2991567.

Wojciech Olszewski. Calibration and expert testing. In *Handbook of Game Theory with Economic Applications*, volume 4, pages 949–984. Elsevier B.V, 2015. ISBN 044453766X.

Wojciech Olszewski and Marcin Pęski. The principal-agent approach to testing experts. *American Economic Journal: Microeconomics*, 3(2):89–113, 2011. ISSN 19457669, 19457685. URL http://www.jstor.org/stable/41237186.

Lionel Page and Robert T. Clemen. Do Prediction Markets Produce WellCalibrated Probability Forecasts? *The Economic Journal*, 123(568):491–513, 12 2012. ISSN 0013-0133. doi: 10.1111/j.1468-0297.2012.02561.x. URL https://doi.org/10.1111/j.1468-0297.2012.02561.x.

Vianney Perchet. Approachability, regret and calibration: Implications and equivalences. *Journal of Dynamics and Games*, 1(2):181–254, 2014.

Ariel Rubinstein. *Modeling Bounded Rationality*. The MIT Press, 01 1998. ISBN 9780262282758. doi: 10.7551/mitpress/4702.001.0001. URL https://doi.org/10.7551/mitpress/4702.001.0001.

Ran Spiegler. *Bounded Rationality and Industrial Organization*. Number 9780199334261 in OUP Catalogue. Oxford University Press, 2014. ISBN ARRAY(0x4d867780). URL https://ideas.repec.org/b/oxp/obooks/9780199334261.html.

A. Wald. Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics*, 16(2):117–186, 1945. ISSN 00034851. URL http://www.jstor.org/stable/2235829.

Joseph Whitmeyer and Mark Whitmeyer. Mixtures of mean-preserving contractions. *Journal of Mathematical Economics*, 94:102450, 2021. ISSN 0304-4068. doi: https://doi.org/10.1016/j.jmateco.2020.11.006. URL https://www.sciencedirect.com/science/article/pii/S0304406820301270.

A Omitted Proofs

A.1 Proof of Proposition 1

Fix $c \geq \bar{c}$, where $\bar{c} = \max_{p \in \Delta\Omega, a \in A} \mathbb{E}_p[u_S(\omega, a) - u_S(\omega, \hat{a}(p))]$. Let us assume the optimal forecasting strategy is not calibrated. This implies that there exists an $\epsilon > 0$ and a possible play such that for infinitely many rounds T, we have

$$\sum_{f \in F} \frac{|\mathbb{N}_T[f]|}{T} \|\overline{\omega}_T[f] - f\| > \epsilon \tag{33}$$

This implies that the sender pays the punishment cost c in infinitely many rounds. Thus, her long run average payoff equals $-c$. Recall, $\hat{a}(p)$ is the receiver's optimal action given his belief over state is p. The punishment cost $c \geq \bar{c}$ ensures that for any conditional over states $p \in \Delta\Omega$, the sender prefers to report honestly and pass the calibration test than pay the punishment cost c.

We now show, in Proposition 5, that there exists a sequence of desirable error margins $\{\epsilon_T\}_{T=1}^{\infty}$ such that a honest sender only pays the punishment cost c in finitely many rounds.

Proposition 5 *There exists a sequence of error margins $\{\epsilon_T\}_{T=1}^{\infty}$ such that $\lim_{T\to\infty} \epsilon_T = 0$ and an honest sender only fails the ϵ_T-calibration test finitely many times almost surely.*

Proof.

Let

$$\overline{x}_T^f = \frac{1}{T}\sum_{t=1}^{T} \overline{d}_T^f = \frac{1}{T}\sum_{t=1}^{T} 1_{\{f_t = f\}}[\delta_{\omega_t} - f]$$

If $\sum_{f \in F} \|\overline{x}_T^f\| > \epsilon_T$, then the sender fails the calibration test in round T. Let's consider an honest sender who forecasts the true conditional distribution of states, i.e., $f_t = p_t$ for all t. Then, $\overline{d}_T^f$ is a martingale difference sequence adapted to the process μ. We have $\mathbb{E}[\overline{d}_T^f] = \mathbf{0}$ and that $\|\overline{d}_T^f\| \leq 1$ a.s.. Using Lemma 2 of Foster et al. [2011], we have

$$\mathbb{P}(\|\overline{x}_T^f\| \geq \epsilon_T) \leq 2e^{\frac{T\epsilon_T^2}{8c}}$$

for some constant $c > 0$.

The Borel-Cantelli lemma states that if the sum of the probability of events, in our case when the sender fails the calibration test ($\sum_{f \in F} \|\overline{x}_T^f\| > \epsilon_T$), is finite then the probability that infinitely many of them occur is zero. Given the realized forecasts exactly match with the finite set of conditionals D, we put a bound on the event $E_T =$

$(\mathbb{P}(\max_{f\in D})\|\overline{x}_T^f\| \geq \epsilon_T)$. So, we have $\sum_{T=1}^{\infty}\mathbb{P}(E_T) = \sum_{T=1}^{\infty}\mathbb{P}(\max_{f\in D}\|\overline{x}_T^f\| > \epsilon_T) \leq \sum_{T=1}^{\infty} e^{\frac{-T\epsilon_T^2}{8c}}$. Choosing $\epsilon_T = o(T^{-\frac{1}{3}})$ suffices to complete the proof. ■

Thus, assuming the error margins satisfy the property, the sender can always avoid the punishment cost c and improve her payoff by honest forecasting.

A.2 Proof of Proposition 2

The proof relies on the simple characterization of $G(P)$, when $Supp(P)$ is affinely independent. A probability measure is a simple mean-preserving contraction if and only if the barycenter remains fixed $\mathcal{B}(P) = \mathcal{B}(Q)$ and $q \in Co(Supp(P)) \forall q \in Supp(Q)$. This feasibility condition only holds when the support is affinely independent. Using Proposition 6, the solution to the persuasion problem is given by the concave envelope restricted to $Co(Supp(P))$. The feasibility condition for measure Q simply boils down to Bayes plausibility in the restricted domain.

Proposition 6 *Suppose $Supp(P)$ is affinely independent, then*

$$Q \in G(P) \iff Supp(Q) \subset Co(Supp(P)) \text{ and } \mathcal{B}(P) = \mathcal{B}(Q) \tag{34}$$

Proof. ($\Leftarrow$) Choose a (finite) probability measure Q such that $Supp(Q) \subset Co(Supp(P))$ and $\mathcal{B}(P) = \mathcal{B}(Q)$. For all $q \in Supp(Q)$, we can write q as a convex combination of $Supp(P)$.

$$q_j = \sum_{i=1}^{n} \alpha_{ij} p_i \text{ such that } \alpha_{ij} \geq 0, \sum_{i=1}^{n} \alpha_{ij} = 1 \quad \forall i \in \{1, ..., n\}, j \in \{1, ..., m\} \tag{35}$$

$$\sum_{j=1}^{m} \mu_j q_j = \sum_{j=1}^{m}\sum_{i=1}^{n} \mu_j \alpha_{ij} p_i \tag{36}$$

Under our assumption, we have $\mathcal{B}(P) = \mathcal{B}(Q)$, i.e.,

$$\sum_{i=1}^{n} \lambda_i p_i = \sum_{j=1}^{m} \mu_j q_j \tag{37}$$

Combining this with the above equation we get

$$\sum_{i=1}^{n} \Big(\lambda_i - \sum_{j=1}^{m} \mu_j \alpha_{ij}\Big) p_i = 0 \tag{38}$$

As $Supp(P)$ is affinely independent, we have

$$\Rightarrow \lambda_i = \sum_{j=1}^{m} \mu_j \alpha_{ij} \quad \forall i \in \{1,..,n\} \tag{39}$$

Let $G_{ij} = \frac{\mu_j \alpha_{ij}}{\lambda_i}$. G is a row-stochastic matrix. We show that this satisfies definition (**??**)

$$\sum_{i=1}^{n} \lambda_i G_{ij} = \sum_{i=1}^{n} \mu_j \alpha_{ij} \tag{40}$$

$$= \mu_j \tag{41}$$

Similarly, we have

$$\sum_{i=1}^{n} \lambda_i p_i G_{ij} = \sum_{i=1}^{n} \mu_j p_i \alpha_{ij} \tag{42}$$

$$= \mu_j q_j \tag{43}$$

($\Rightarrow$) Given each q_j is constructed by merging weights of $Supp(P)$, it is necessary that $q_j \in Co(Supp(P))$. We only need to show that $\mathcal{B}(P) = \mathcal{B}(Q)$. Assume $Q \in G(P)$, so we know that there exists a row-stochastic matrix G such that:

$$\mu_j q_j = \sum_{i=1}^{n} \lambda_i p_i G_{ij} \quad \forall j \in \{1,...,m\} \tag{44}$$

$$\mathcal{B}(Q) = \sum_{j=1}^{m} \mu_j q_j = \sum_{j=1}^{m} \sum_{i=1}^{n} \lambda_i p_i G_{ij} \tag{45}$$

$$= \sum_{i=1}^{m} \lambda_i p_i = \mathcal{B}(P) \tag{46}$$

■

A.3 Proof of Lemma 11

First, we show that if the forecasting strategy is calibrated, the equation (17) holds. We have

$$\frac{|N_T[f]|}{T}\|f-\sum_{p\in D}p\mu_T(f,p)\| \tag{47}$$

$$\leq\frac{|N_T[f]|}{T}\|f-\frac{\sum_{t=1}^{T}1_{\{f_t=f\}}\delta_{\omega_t}}{\sum_{t=1}^{T}1_{\{f_t=f\}}}\|+\|\frac{\sum_{t=1}^{T}\sum_{p\in D}1_{\{f_t=f,p_t=p\}}[\delta_{\omega_t}-p]}{T}\| \tag{48}$$

Both the terms converge to zero as $T\to\infty$. The first vanishes from the property of calibration. For the second term, we apply Azuma-Hoeffding inequality. Denote

$$\overline{x}_T=\|\frac{\sum_{p\in D}\sum_{t=1}^{T}1_{\{f_t=t,p_t=p\}}[\delta_{\omega_t}-p]\|}{T}$$

We have $\mathbb{E}[\overline{x}_T]=0$ and that $\|\overline{x}_T-\overline{x}_1\|\leq 1$. Using the Azuma-Hoeffding inequality, we have $\mathbb{P}(\|\overline{x}_T\|>\eta)\leq\exp^{-2T\eta^2}$. Choosing $\eta=o(T^{-\frac{1}{3}})$ suffices in our case.

For the converse, we can use the same procedure. We apply the triangle inequality to show that the if equation (17) holds, then the forecasting strategy is calibrated.

$$\frac{|N_T[f]|}{T}\|f-\frac{\sum_{t=1}^{T}1_{\{f_t=f\}}\delta_{\omega_t}}{\sum_{t=1}^{T}1_{\{f_t=f\}}}\| \tag{49}$$

$$\leq\frac{|N_T[f]|}{T}\|f-\sum_{p}p\mu_T(f,p)\|+\|\frac{\sum_{t=1}^{T}\sum_{p}1_{\{f_t=t,p_t=p\}}[\delta_{\omega_t}-p]}{T}\| \tag{50}$$

The first term goes to zero from assumption and the second terms goes to zero from Azuma-Hoeffding inequality.

A.4 Lemma 18

Lemma 18 *For a stationary ergodic process μ, the distribution of conditionals C_μ exists and is constant μ-a.s.*

Proof. Consider the two-sided extension of the stationary process i.e., $\mu\in\Delta\Omega^{\mathbb{Z}}$. Let $\omega_a^b=(\omega_a,.....,\omega_{b-1})$ where $a<b-1$.

$$f_n=\mu(\omega_0=\cdot\mid\omega_{-n}^{0}) \tag{51}$$

$$f_\infty=\mu(\omega_0=\cdot\mid\omega_{-\infty}^{0}) \tag{52}$$

Using the martingale convergence theorem we have that, μ-a.s., $f_n\to f_\infty$. Given μ is stationary, using the shift transformation T, we have

$$f_n \circ T^n = \mu(\omega_n = \cdot \mid \omega_0^n) = p_n \tag{53}$$

Since f_n and $p_n = f_n \circ T^n$ have the same distribution for all $n \in \mathbb{N}^+$, we can conclude that $p_n \to p_\infty = \mu(\omega_\infty = \cdot \mid \omega_0^\infty)$ μ-a.s. and $\mathbb{E}[\mu(\omega_0 = \cdot \mid \omega_{-\infty}^0)] = \mathbb{E}[\mu(\omega_\infty = \cdot \mid \omega_0^\infty)]$. Now, given μ is stationary and ergodic, we apply the Maker's Ergodic theorem to get:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(\omega_n = \cdot \mid \omega_0^n) = \mathbb{E}[\mu(\omega_\infty = \cdot \mid \omega_0^\infty)] \quad \mu\text{-a.s.} \tag{54}$$

In particular, this implies for any $p \in P$

$$C_\mu(p) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{1}_{\{p_n = p\}} = \mathbb{E}[\mathbf{1}_{\{p_\infty = p\}}] \quad \mu\text{-a.s.} \tag{55}$$

For reference, the Maker's Ergodic Theorem Kallenberg [2002]:

Theorem 19 *(Maker's Ergodic Theorem) Let $\mu \in \Delta\Gamma$ be a stationary distribution and let $f_0, f_1, \ldots : \Gamma \to \mathbb{R}$ be such that $sup_n|f_n| \in L_1(\mu)$ and $f_n \to f_\infty$ μ-a.s. Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f_n \circ T^n \to \mathbb{E}[f_\infty \mid \mathcal{I}] \quad \mu\text{-a.s.} \tag{56}$$

■

A.5 Proof of Lemma 10

First, we show that for a calibrated strategy σ, there will exist a distribution of forecasts $F_{\mu,\sigma}$ such that $F_{\mu,\sigma} \in G(C_\mu)$. Consider

$$\sum_{f\in F} \| \frac{\sum_{t=1}^T \mathbf{1}_{\{f_t = f\}}}{T} f - \sum_p p C_\mu(p) \frac{\sum_{t=1}^T \mathbf{1}_{\{f_t = f, p_t = p\}}}{\sum_{t=1}^T \mathbf{1}_{\{p_t = p\}}} \| \tag{57}$$

$$\leq \sum_{f\in F} \| \frac{\sum_{t=1}^T \mathbf{1}_{\{f_t = f\}}}{T} f - \sum_p p \frac{\sum_{t=1}^T \mathbf{1}_{\{f_t = f, p_t = p\}}}{T} \| + \sum_{f\in F} \sum_p p (\frac{\sum_{t=1}^T \mathbf{1}_{\{f_t = f, p_t = p\}}}{T}) \| 1 - \frac{C_\mu(p) T}{\sum_{t=1}^T \mathbf{1}_{\{p_t = p\}}} \| \tag{58}$$

From Lemma 11, we have that the first term goes to zero as $T \to \infty$. The second term also goes to zero using the definition of the distribution of conditionals C_μ. Thus, we have

$$\limsup_{T\to\infty}\sum_{f\in F}\|\frac{\sum_{t=1}^{T}\mathbf{1}_{\{f_t=f\}}}{T}f-\sum_{p}pC_\mu(p)\frac{\sum_{t=1}^{T}\mathbf{1}_{\{f_t=f,p_t=p\}}}{\sum_{t=1}^{T}\mathbf{1}_{\{p_t=p\}}}\|=0 \tag{59}$$

For any n, let $F_n[f]=\frac{\sum_{t=1}^{n}\mathbf{1}_{\{f_t=f\}}}{T}$. We have $F_n\in G(C_\mu)$ as the row stochastic matrix $G_n(f,p)=\frac{\sum_{t=1}^{n}\mathbf{1}_{\{f_t=f,p_t=p\}}}{\sum_{t=1}^{T}\mathbf{1}_{\{p_t=p\}}}$ satisfies the condition for a simple mean-preserving contraction. As $\Delta(\Delta\Omega)$ is compact, we know there exists a subsequence $n_1,n_2,....$ such that converges to a limit point F, i.e.,$F=\lim_{t\to\infty}F_{n_t}$. Thus, from definition 6, F is a mean-preserving contraction of C_μ.

On the other hand, let $F\in G(C_\mu)$. Given Lemma 18, the sender knows the distribution of conditionals C_μ. Let $\tau:P\to\Delta F$ be the signaling policy that given prior C_μ results in the distribution $F=(\mu,q)$. Consider the stationary forecasting strategy given by $\sigma_t(f_t=f\mid p_t=p)=\tau(f\mid p)$ for all t. We show the forecasting strategy σ is calibrated. For any $f\in Supp(Q)$, we have

$$\frac{|N_T[f]|}{T}\|\frac{\sum_{t=1}^{T}1_{\{f_t=f\}}\delta_{\omega_t}}{\sum_{t=1}^{T}1_{\{f_t=f\}}}-f\| \tag{60}$$

$$\leq\|\frac{\sum_p\sum_{t=1}^{T}1_{\{f_t=f,p_t=p\}}[\delta_{\omega_t}-p]}{T}\|+\|\frac{\sum_{p\in P}\sum_{t=1}^{T}1_{\{f_t=f,p_t=p\}}(p-f)}{T}\| \tag{61}$$

$$\leq\|\frac{\sum_p\sum_{t=1}^{T}1_{\{f_t=f,p_t=p\}}[\delta_{\omega_t}-p]}{T}\|+\|\frac{\sum_p\sum_{t=1}^{T}1_{\{p_t=p\}}\tau(f\mid p)[p-f]]}{T}\| \tag{62}$$

Again, using the Azuma-Hoeffding inequality, the first term vanishes to zero. For the second term, we know that C_μ is well defined and so the overall term converges to zero from the definition of mean-preserving contraction. Hence, we have proved that for a stationary ergodic process, the sender can implement any $F\in G(C_\mu)$.

A.6 Proof of Theorem 13

We first show that the function $\overline{Co}\ \hat{u}_S(f^*(p))$ is attainable. Then, we show that is the highest continuous attainable function.

(a) The main idea of the proof is to combine the payoff and calibration cost function such that the overall set is a closed and convex set (a similar proof can be found in Mannor et al. [2009]). Then, we use the dual condition to show that the set is approachable.

First, we provide the proof for the case of $\epsilon-$calibration. We then use that strategy to construct a calibrated sender by using the "doubling trick". The idea is to partition the play into blocks of exponentially increasing sizes.

Let's assume the set of forecasts F is finite. Given $\epsilon>0$, we assume the set of feasible forecasts F is given by the regular ϵ-grid F_ϵ:

$$F_\epsilon = \{\sum_{\omega\in\Omega} n_\omega \delta_\omega \in \Delta\Omega \mid n_\omega \in \{0, \frac{1}{L}, ..., 1\} \text{ and } \sum_{\omega\in\Omega} n_\omega = 1\} \tag{63}$$

where, $L = \lceil \frac{\sqrt{|\Omega|-1}}{2\epsilon} \rceil \in \mathbb{N}$. [12] This ensures (generically) for any $p \in \Delta\Omega$ there exists a unique pure forecast $f \in F$ such that $\|f - \sum_{\omega\in\Omega} p(\omega)\delta_\omega\| \le \epsilon$. Let us denote by $f^*(p)$ as the pure forecast that belongs to the ϵ-neighbourhood of $p \in \Delta(\Omega)$.

Given forecast $f \in F$ and state $\omega \in \Omega$ the calibration cost c is given by:

$$c(f,\omega) = (\underline{0}, ..., f - \delta_\omega, ..., \underline{0}) \in \mathbb{R}^{|F_\epsilon||\Omega|} \tag{64}$$

It is a vector of infinite elements of size $\mathbb{R}^{|\Omega|}$ with one non-zero element (at the position for f) while the rest are equal to $\underline{0} \in \mathbb{R}^{|\Omega|}$. The ϵ-calibration condition (2) can be rewritten as follows: the average of the sequence of vector-valued calibration costs $c_t = c(f_t, \omega_t)$ converges to the set E_ϵ almost surely, where

$$E_\epsilon = \{x \in \mathbb{R}^{|F_\epsilon||\Omega|} : \sum_{f\in F} \|\underline{x}_f\| \le \epsilon\}$$

For each stage t, the sender and nature simultaneously choose $f_t \in F$ and $\omega_t \in \Omega$ respectively. This results in a reward $r_t = \hat{u}_S(f_t)$ and penalty $c_t = c(f_t, \omega_t)$ for the sender in stage t.

Unlike the (exact) calibration test, the forecast and the limit empirical distribution of states do not have to exactly match but can differ up to an error margin ϵ. This is because we assume the set of feasible forecasts F is a finite grid. Even an informed sender who sends reports honestly can fail the exact calibration test if she was restricted to send forecasts from the finite set F.

We now show the function $\overline{Co}\ \hat{u}_S(f^*(p))$ is attainable. We show this using the dual condition of approachability. For any time period t, consider the vector-valued payoff $m(f,\omega)$ constructed using the sender's payoff and the calibration cost:

$$m_t = m(f_t, \omega_t) = (r_t, c_t, \delta_{\omega_t}) \in D \subseteq \mathbb{R} \times \mathbb{R}^{n|\Omega|} \times \Delta(B) \tag{65}$$

Now, consider the sets:

$$D_1 = \{(r,c,p) \in D : r \ge \overline{Co}\ \hat{u}_S(f^*(p))\} \quad D_2 = \{(r,c,p) \in D : c \in E_\epsilon\} \tag{66}$$

[12] where $\lceil x \rceil$ denotes the smallest integer greater than or equal to x.

and let $D^* = D_1 \cap D_2$. The set D is closed and convex. To show the average of the sequence of the vector m_t approaches D^*, we need to verify the dual condition of approachability, i.e., for any $p \in \Delta(\Omega) \exists \mu \in \Delta(F)$ such that $m(\mu, p) \in D^*$. We have

$$m(\mu,p) = (\sum_{\omega,f} \mu(f)\hat{u}_S(f), \sum_{\omega,f} \mu(f)p(\omega)c(f,\omega), p) \in D^* \tag{67}$$

By definition $f^*(p)$ satisfies Condition (67) and the set D^* is approachable. The set D_1 ensures that the sender can attain $\overline{Co}\ \hat{u}_S(f^*(p))$ while the set D_2 ensures that the ϵ-calibration condition is met.

Moreover, approachability theory provides convergence rates of vector payoff and the calibration condition (see Perchet [2014]). For every strategy of Nature and for every $\delta > 0$, with probability at least $1 - \delta$, we have

$$\sum_{f \in F_\epsilon} \frac{|\mathbb{N}_T[f]|}{T} \|\overline{\omega}_T[f] - f\| \leq \epsilon + \frac{2}{\epsilon^{|\Omega|}} \sqrt{\frac{(|\Omega|-1)}{T} \log(\frac{3}{\delta})} \tag{68}$$

Now, we use the "doubling trick" to show that the sender can pass the exact calibration test. The main idea is to partition time into blocks of exponentially increasing sizes (see Mannor and Stoltz [2010]). Blocks are indexed by $k = 1, 2..$, where the size and error margin of block k is s_k and ϵ_k. The sender uses the ϵ_k-calibrated strategy described before using forecasts f from the grid F_{ϵ_k}. Let b_T denote the block number when the time period is T. For any forecast f, using triangle inequality, we have

$$\sum_{f \in F} \| \sum_{t=1}^{T} \mathbf{1}_{\{f_t = f\}} (\delta_{\omega_t} - f) \| \leq \sum_{f \in F} \sum_{k=1}^{b_T} \| \sum_{t=1}^{T_k} \mathbf{1}_{\{f_t = f\}} (\delta_{\omega_t} - f) \|. \tag{69}$$

We now substitute the bound obtained in the previous section and get that with probability $1 - (\delta_{1,T} + ... + \delta_{b_T,T}) \geq 1 - \frac{1}{T^2}$:

$$\sum_{f \in F} \frac{1}{T} \| \sum_{t=1}^{T} \mathbf{1}_{\{f_t = f\}} (\delta_{\omega_t} - f) \| \leq \sum_{k=1}^{b_T} \frac{T_k}{T} \sum_{f \in F_\epsilon} \frac{|\mathbb{N}_{T_k}[f]|}{T} \|\overline{\omega}_{T_k}[f] - f\|_2 \tag{70}$$

$$\leq \sum_{k=1}^{b_T} \frac{T_k}{T} [\epsilon_k + \frac{2}{\epsilon_k^{|\Omega|}} \sqrt{\frac{(|\Omega|-1)}{T_k} \log(\frac{3}{\delta_{k,T}})}] \tag{71}$$

Now, we choose ϵ_k, T_k and $\delta_{k,T}$ appropriately and use Borel-Cantelli Lemma to obtain the desired convergence. Let $T_k = 2^k$, $\epsilon_k = 2^{\frac{-k}{2(|\Omega|+1)}}$ and let $\delta_{k,T} = \frac{1}{2^k T^2}$. This ensures that ϵ_k and $\frac{1}{\epsilon_k{}^{|\Omega|}\sqrt{T_k}}$ converge to zero while T_k increases exponentially.

(b) Now, we show that the function $\overline{Co}\ \hat{u}_S(p)$ is the highest continuous attainable function. We construct Nature's strategy τ that prevents the sender from attaining $\tilde{r}$. Let $\{p_j\}_{j=1}^k$ denote the support points corresponding to the closed convex hull of $\hat{u}_S$, i.e., $\overline{Co}\ \hat{u}_S(p_0) = \sum_{j=1}^k \alpha_j \hat{u}_S(p_j)$. From Caratheodorys Theorem we can take k to be equal to $|\Omega| + 2$. Consider a game with T stages where Nature plays in a sequence of k blocks, where the size of block l is $\alpha_l T$. In block l, Nature plays i.i.d. according to p_l.

First, we show that for any i.i.d process with distribution p the any calibrated forecasting strategy has to send the truthful forecast p almost surely. In other words, a sender cannot send a forecast $f \neq p$ (with positive probability) and pass the calibration test. From the calibration condition and the law of large numbers, we have

$$\limsup_{T\to\infty} \frac{|\mathbb{N}_T[f]|}{T} \|f - p\| > 0 \quad \forall f \in F \tag{72}$$

Now, consider sender's play in each block l. We claim that for the sender to pass the overall calibration test, she has to pass the test in each block l and thus can only report p_l almost surely. Assume that is not the case and let l denote the first block where the sender's strategy is not calibrated. Then, we have

$$\limsup_{T\to\infty} \sum_{f\in F} \frac{|\mathbb{N}_{n_l T}[f]|}{n_l T} \|\overline{\omega}_{n_l T}[f] - f\| > 0$$

where, $n_l = \sum_{j=1}^l \alpha_j$. If this happens, then let nature can play according to δ_ω for the rest of the game, where $\delta_\omega \neq p_j$ for $j = 1, \dots, l$. Even if the sender repeatedly forecasts δ_ω, the calibration cost will be positive. Even if the sender knew the sequence $p_1, \dots, p_k$ in advance, she could not guarantee a higher payoff without failing the calibration test. Using this block strategy, nature restricts sender to announce forecast p_l in block l. Thus, under nature's strategy τ, the sender's payoff cannot be higher than $\overline{Co}\ \hat{u}_S(p_0) = \sum_{j=1}^k \alpha_j \hat{u}_S(p_l)$.

A.7 Proof of Lemma 14

We use the notion of *opportunistic approachability*, which was developed by Bernstein et al. [2014]. They devise algorithms that in addition to approaching the convex hull of the reward-in-hindsight, seek to approach strict subsets thereof when the opponents play turns out to be restricted in an appropriate sense (either statistically or empirically).

Definition 20 *The play of the opponent is empirically Q-restricted with respect to a partition $\{\tau_m\}$, if there exists a convex subset $Q \subset \Delta(B)$ such that, for the given sample path*

$$\lim_{M\to\infty} \frac{1}{n_M} \sum_{m=1}^{M} \tau_m d(\overline{\omega}_m, Q) = 0 \tag{73}$$

where, τ_m denote the length of block m, $n_M = \sum_{m=1}^{M} \tau_m$ denotes the time at the end of the block M and $\overline{\omega}_m$ denotes the empirical distribution of states in block m.

Bernstein et al. [2014] (see Theorem 5) show that if nature's play is empirically Q-restricted with respect to a partition with subexponentialy increasing blocks, then

$$\lim_{n\to\infty} d(\hat{r}_n, R^+(Q)) = 0 \quad \text{where, } R^+(Q) = \cap_{\epsilon>0} Co\{\hat{u}_S(p) : d(p,Q) \leq \epsilon\} \tag{74}$$

Here, $R^+(Q)$ denotes the closed convex image of the reward-in-hindsight restricted to the set Q. If $Q = \Delta(\Omega)$, then we obtain the same bounds as in the case of an adversarial environment.

Given $\eta > 0$, let $N_\eta(p)$ denote the η–neighbourhood around p where p is the empirical distribution of states. We show that for a stationary ergodic process the play is empirically $N_\eta(p)$-restricted with respect to a partition with finite blocks τ. This results directly from the ergodic theorem: given any $\eta > 0$, there exists $\tau_\omega \in \mathbb{N}$ such that forall $\tilde{\tau} \geq \tau_\omega$ we have

$$\left|\frac{1}{\tilde{\tau}} \sum_{t=1}^{\tilde{\tau}} \mathbf{1}_{\{\omega_t=\omega\}} - p(\omega)\right| \leq \eta \tag{75}$$

The empirical distribution p exists and is constant for a stationary ergodic process. Choosing $\tau^* = \max_{\omega\in\Omega} \tau_\omega$ as the block size ensures that nature's play in each block is empirically restricted to $N_\eta(p)$. This implies that the average reward function $\hat{r}_n$ converges to $R^+(N_\eta(p))$. Thus, choosing an appropriate $\eta > 0$ such that $f^*(q) = f^*(p)$ for all $q \in N_\eta(p)$, the sender can attain the reward-in-hindsight function evaluated at the empirical distribution p. [13] Additionally, Bernstein et al. [2014] show that the results hold without knowing if nature's play is empirically restricted nor knowing the restriction set Q.

A.8 Proof of Proposition 4

Fix $Q \in G(C_\mu)$ and let $\sigma : P \to \Delta F$ denote the signaling policy that results in Q. From Theorem (9), we know that the strategy is calibration and the distribution of forecasts $F_{\mu,\sigma} = Q$. In particular, we have

[13] Note, this is valid as long as $f^*(p_o) \in A$ corresponds to a unique forecast. This condition is met generically.

$$\limsup_{T\to\infty} \frac{\sum_{t=1}^{T} \mathbf{1}_{\{f_t=f,\omega_t=\omega\}} u_R(\omega, a_t)}{T} = f(\omega) \limsup_{T\to\infty} \frac{\sum_{t=1}^{T} \mathbf{1}_{\{f_t=f\}} u_R(\omega, a_t)}{T} \tag{76}$$

This follows as if $\lim_n a_n = a$, then $\limsup_n a_n b_n = a \limsup_n b_n$. Given the receiver minimizes regret, we have

$$\limsup_{T\to\infty} \max_{a^*\in A} \frac{\sum_{t=1}^{T} \mathbf{1}_{\{f_t=f\}} [u_R(\omega_t, a^*) - u_R(\omega_t, \hat{a}(f))]}{T} \leq 0 \tag{77}$$

$$\Rightarrow \limsup_{T\to\infty} \max_{a^*\in A} \frac{\sum_{t=1}^{T} \mathbf{1}_{\{f_t=f\}} \mathbb{E}_f [u_R(\omega_t, a^*) - u_R(\omega_t, a_t)]}{T} \leq 0 \tag{78}$$

If the receiver plays actions $a_t \notin \hat{a}(f)$ on rounds with non-negligible weight, it results in a positive regret. Hence, if the receiver uses any no-regret learning algorithm, it will almost surely play the recommended action $\hat{a}(f)$. Thus, for a stationary ergodic process, the sender can ensure that she guarantees $\mathbb{E}_Q[\hat{u}_S]$ for any $Q \in G(C_\mu)$.

B Extension: MDP

In this section, we consider an environment where the receiver's action affects how the states evolves. We consider the stochastic process μ is a Markov Decision Process with transition matrix $T(\omega_{t+1} \mid \omega_t, a_t)$. The behavioural assumption of the receiver remains the same: he uses the calibration test to verify the claims of the sender. He plays according to the sender's forecast $a_t = \hat{a}(f_t)$ if she passes the test and punishes her if she fails. [14] While characterizing the set of feasible policies that pass the calibration test, it suffices to consider policies with memory 1, i.e., policies that maps the past forecast and state into a possible random forecast $\sigma : F \times \Omega \to \Delta F$. Any such policy induces a Markov chain over the set $F \times \Omega$, whose transition matrix Q_σ is given by

$$Q_\sigma(f, \omega \mid f', \omega') = T(\omega \mid \omega', f') \sigma(f \mid \omega', f') \tag{79}$$

Let us denote by $\mathcal{F}$ the set of feasible distributions $\mu \in \Delta F$, i.e., μ corresponds to the marginal over F of the invariant distribution of the Markov chain induced by a calibrated reporting policy $\sigma : F \times \Omega \to \Delta F$.

[14] For simplicity, we assume the receiver plays according to the sender's forecasts and only performs the calibration test at the end of the game. This is because the punishment action also affects the state transition and can be difficult to keep track of.

Proposition 7 *The set of feasible distributions $\mathcal{F}$ is a convex polytope. It is the marginal of $\eta \in \Delta(F \times \Omega \times F)$ over F where η satisfies:*

$$\sum_{\omega',f'} \eta(f',\omega',f)T(\omega \mid \omega',f') = \eta(f,\omega) \tag{80}$$

$$\eta(f,\omega) = \eta(f)f(\omega) \tag{81}$$

Equation (80) states that η is the invariant distribution for some reporting policy. Equation (81) states that the invariant distribution η is calibrated.

Proof. Given a reporting policy $\sigma : F \times \Omega \to \Delta F$, we know the (time-averaged) distribution of outcomes (forecasts and states) converges almost surely. For simplicity, we assume the MDP is unichain. This implies that distribution of outcomes is the unique invariant distribution η of the Markov chain induced by σ. [15] As η is the invariant distirbution of the induced Markov chain, we have

$$\eta(f',\omega') = \lim_{T\to\infty} \frac{\sum_{t=1}^{T} 1_{\{\omega_t=\omega',f_t=f'\}}}{T} \tag{82}$$

$$\eta(f,\omega) = \sum_{\omega',f'} \eta(f',\omega')Q_\sigma(f,\omega \mid f',\omega') \tag{83}$$

Also, as we assume the forecasts are calibrated, we have

$$\eta(\omega \mid f) = \lim_{T\to\infty} \frac{\sum_{t=1}^{T} 1_{\{f_t=f,\omega_t=\omega\}}}{\sum_{t=1}^{T} 1_{\{f_t=f\}}} = f(\omega) \tag{84}$$

Equivalently, a distribution $\eta \in \Delta(F \times \Omega \times F)$ is feasible and invariant for some reporting policy σ if

$$\sum_{f''} \eta(f,\omega,f'') = \sum_{\omega',f'} \eta(f',\omega',f)T(\omega \mid \omega',f') \tag{85}$$

$$\eta(f,\omega) = \eta(f)f(\omega) \tag{86}$$

where, the policy (with memory 1) to induce the distribution η is given by

$$\sigma(f \mid \omega',f') = \begin{cases} \frac{\eta(f',\omega',f)}{\eta(f',\omega')} & \text{if } \eta(f',\omega') > 0 \\ \eta(f) & \text{if } \eta(f',\omega') = 0 \end{cases}$$

[15] An MDP is unichain if every pure policy gives rise to a Markov chain with at most one recurrent class.

■

The sender's maximization problem is given by the following linear program:

$$\max_{\mu \in \mathcal{F}} \sum_f \mu(f) \hat{u}_S(f) \tag{87}$$

Thus, we can extend our model to situations where the receiver's action affects the distribution of future states. Furthermore, we can characterize the set of outcomes that result from calibrated strategies and solve for the optimal forecasting strategy.

C Persuasion problem in terms of experiments

In this section, we define an equivalent way of describing the persuasion problem in terms of Blackwell experiments. This is the standard way of modeling the persuasion problem in the case of a perfectly informed sender(see Kamenica and Gentzkow [2011]). We extend it to the case, where the sender is imperfectly informed and can only use experiments less informative than a prior experiment.

Definition 21 *An experiment $F : \Omega \to \Delta T$ is a garbling of the experiment $E : \Omega \to \Delta S$ if there exists a row-stochastic matrix (or mapping) $G : S \to \Delta T$ such that $EG = F$.*

This defines a partial ordering in the set of Blackwell experiments. We say $F \precsim E$ when experiment F is a garbling of experiment E. A prior belief $p_0 \in \Delta\Omega$ and an experiment $E : \Omega \to \Delta S$ give rise to a probability measure $\mathcal{P}(p_0, E) := (\lambda_s, p_s)_{s \in S}$. Conversely, given any probability measure $Q = (\mu, q)$ you can define a prior belief (or barycenter) $\mathcal{B}(Q)$ and an experiment $\mathcal{E}(Q)$.

$$\lambda_s = \sum_{\omega \in \Omega} p_0(\omega) E(s \mid \omega) \qquad \mathcal{B}(Q) = \sum_{i=1}^{n} \mu_i q_i$$

$$p_s(\omega) = \frac{p_0(\omega) E(s \mid \omega)}{\sum_{\omega \in \Omega} p_0(\omega) E(s \mid \omega)} \qquad \mathcal{E}(Q)(s_i \mid \omega) = \frac{\mu_i q_i(\omega)}{\sum_{i=1}^{n} \mu_i q_i(\omega)}$$

Lemma 22 shows the equivalence between the simple mean-preserving contraction of a probability measure (Elton and Hill [1998],Whitmeyer and Whitmeyer [2021]) and the garbling of an experiment (Blackwell [1953]) with a fixed prior belief.

Lemma 22 *A probability measure Q is a mean-preserving contraction of P if and only if $\mathcal{B}(P) = \mathcal{B}(Q)$ and the experiment $\mathcal{E}(Q)$ is a garbling of $\mathcal{E}(P)$.*

$$
\begin{array}{ccc}
P & \text{-----------} & (p_0, E) \\
{\scriptstyle G}\downarrow & & \downarrow{\scriptstyle G} \\
Q & \text{-----------} & (p_0, EG)
\end{array}
\tag{88}
$$

Proof. (⇒) Assume probability measure Q is a mean-preserving contraction of P. First, we show that the barycenter of the two measures is equal.

$$
\begin{aligned}
\mathcal{B}(Q) &= \sum_{j=1}^{m} \mu_j q_j && (89)\\
&= \sum_{j=1}^{m} \sum_{i=1}^{n} \lambda_i p_i G_{ij} && (90)\\
&= \sum_{i=1}^{n} \lambda_i p_i = \mathcal{B}(P) && (91)
\end{aligned}
$$

Now, we show the resulting experiment F is a garbling of E.

$$
\begin{aligned}
F(t_j \mid \omega_n) &= \frac{\mu_j q_j(\omega_n)}{\sum_{j=1}^{m} \mu_j q_j(\omega_n)} && (92)\\
&= \frac{\sum_{i=1}^{n} \lambda_i p_i(\omega_n) G_{ij}}{\sum_{i=1}^{n} \lambda_i p_i(\omega_n)} && (93)\\
&= \sum_{i=1}^{n} E(s_i \mid \omega_n) G_{ij} && (94)
\end{aligned}
$$

(⇐) Assume the Blackwell experiment F is a garbling of E. We show that the resulting measure Q is a simple mean-preserving contraction of P.

$$\mu_j = \sum_{k=1}^{s} p_0(\omega_k) F(t_j \mid \omega_k) \tag{95}$$

$$= \sum_{k=1}^{s} p_0(\omega_k) \sum_{i=1}^{n} E(s_i \mid \omega_k) G_{ij} \tag{96}$$

$$= \sum_{i=1}^{n} \lambda_i G_{ij} \tag{97}$$

$$\mu_j q_j = p_0(\omega_k) F(t_j \mid \omega_k) \tag{98}$$

$$= p_0(\omega_k) \sum_{i=1}^{n} E(s_i \mid \omega_k) G_{ij} \tag{99}$$

$$= \sum_{i=1}^{n} \lambda_i p_i G_{ij} \tag{100}$$

■

Thus, solving the persuasion problem $(P, \hat{u}_S)$ is equivalent to finding the optimal garbling of the experiment $\mathcal{E}(P)$ with prior $\mathcal{B}(P)$.

$$\max_{F \precsim \mathcal{E}(P)} \mathbb{E}_Q[u] \text{ where } Q = \mathcal{P}(\mathcal{B}(P), F) \tag{101}$$

For a perfectly informed sender, there is no constraint on the Blackwell experiment F as $E(\mathcal{P})$ corresponds to full information. Also, $\mathcal{B}(P)$ is simply the common prior belief of the players. Any distribution Q such that $(\mathcal{B}(P) = \sum_j \mu_j q_j$ (Bayes plausibility condition) can be implemented.

Note the optimization problem (101) is in terms of mean-preserving spreads (splittings) while the equivalent problem (8) is in terms of mean-preserving contractions (garblings).