The Core of Bayesian Persuasion*

Laura Doval† Ran Eilat‡

July 27, 2023

Abstract

An analyst observes the frequency with which an agent takes actions, but not the frequency with which she takes actions conditional on a payoff relevant state. In this setting, we ask when the analyst can rationalize the agent's choices as the outcome of the agent learning something about the state before taking action. Our characterization marries the obedience approach in information design (Bergemann and Morris, 2016) and the belief approach in Bayesian persuasion (Kamenica and Gentzkow, 2011) relying on a theorem by Strassen (1965) and Hall's marriage theorem. We apply our results to ring-network games and to identify conditions under which a data set is consistent with a public information structure in first-order Bayesian persuasion games.

KEYWORDS: *Bayes correlated equilibrium, Bayesian persuasion, information design, stochastic choice, distributions with given marginals, cooperative games, set functions, core*

1 Introduction

Given a primitive payoff structure, information design provides a framework for rationalizing outcomes as the result of non-cooperative play *without* having to specify the players' information structure. For this reason, the seminal work of Bergemann and Morris

*We are grateful to Yaron Azrieli, Denniz Kattwinkel, Emir Kamenica, Shengwu Li, Elliot Lipnowski, Meg Meyer, Stephen Morris, Jacopo Perego, Andrea Prat, John Rehbeck, and Vasiliki Skreta for thought-provoking questions and insightful discussions. We owe special thanks to Quitzé Valenzuela-Stookey, who suggested ideas that facilitated the proof of Theorem 1.

†Columbia Business School and CEPR. Email: laura.doval@columbia.edu

‡Department of Economics, Ben-Gurion University. E-mail: eilatr@bgu.ac.il

(2016) has spurred renewed interest among empirical scholars wishing to obtain identification and estimation results under a weaker set of information assumptions (see, for instance, Syrgkanis et al., 2017; Magnolfi and Roncoroni, 2019; Koh, 2022).

However, the weaker set of assumptions on the information structure comes at the cost of increasing demands on the data set available to the analyst. Indeed, the usual assumption in the literature is that the analyst is given a *joint* distribution over payoff relevant states and action profiles. For instance, the literatures on rational inattention and stochastic choice usually assume that the analyst observes an agent's choices conditional on the state of the world (e.g., Caplin and Dean, 2015; Aguiar et al., 2018). Given this data set, Bayes correlated equilibrium provides an easy to test set of conditions that the joint distribution over states and action profiles must satisfy in order to be consistent with the outcome of non-cooperative play under some information structure.

Oftentimes, however, the analyst's data set is more limited. The analyst may observe the distribution over the payoff relevant states of the world and the distribution over action profiles, but not the distribution over action profiles conditional on the state of the world.[1] We can then ask, given the primitive payoff structure, which marginal distributions can be rationalized as the outcome of non-cooperative play under some information structure. We refer to such marginals as *BCE consistent* because they satisfy that a joint distribution over states and action profiles exists that is consistent with the marginals *and* is a Bayes correlated equilibrium. Characterizing the set of BCE-consistent marginal distributions can only increase the practical applicability of Bayes correlated equilibrium.

The set of BCE-consistent marginal distributions is of interest for two other reasons. First, the analyst oftentimes is not just interested in the existence of an information structure that rationalizes the (marginal) distribution of play, but one that satisfies certain properties. For instance, the analyst may want to test whether the agents have private information. As we explain below, our characterization result provides us with a test for the existence of a public information structure that rationalizes the observed distribution of play. The second reason is related to reduced-form implementation in mechanism design (Matthews, 1984; Border, 1991). Whenever the information designer only cares about the agents' action profiles, but not the state of the world, the information designer's problem can be expressed as the choice out of the set of BCE-consistent marginals.

In this paper, we take the first step towards characterizing the set of BCE-consistent

[1]Whereas state-dependent stochastic choice data is useful to guide the design and interpretation of experiments, this data is oftentimes hard to come by outside the experimental setting. Dardanoni et al. (2020) provides an eloquent discussion of the *data voracity* of stochastic choice.

marginals by considering the single-agent case. Theorem 1 provides a characterization of the set of BCE-consistent marginals building on a theorem in Strassen (1965). Furthermore, marrying the obedience approach in information design with the belief approach in Kamenica and Gentzkow (2011), Proposition 1 characterizes the Bayes plausible distributions over posteriors that implement a given marginal over actions. We provide two network-based proofs of Proposition 1. Relying on recent extensions of Hall's marriage theorem in Barseghyan et al. (2021) and Azrieli and Rehbeck (2022), the first characterization uncovers a connection between BCE-consistency and the *core* of the game induced by loosely speaking, some (Bayes plausible) posterior distribution (see Remark 2 and Grabisch et al., 2016). The second proof relies on the demand problem of Gale (1957). We show that one can interpret BCE-consistency problem as a supply-demand problem in a persuasion economy, in which the marginal action distribution describes the demand and a Bayes plausible posterior distribution describes the supply. We then rely on the results in Gale (1957) to determine when the demand is feasible given the supply.

Section 4 illustrates how Theorem 1 already allows us to study multi-agent games. Section 4.1 applies Theorem 1 to the first-order Bayesian persuasion setting of Arieli et al. (2021) to characterize the subset of BCE-consistent marginals that are consistent with a *public* information structure. Instead, Section 4.2 applies Theorem 1 to characterize BCE-consistent marginals in ring-network games as in Kneeland (2015).

Related literature The two closest papers to ours are Rehbeck (2023) and Azrieli and Rehbeck (2022). Rehbeck (2023) studies the same question as us, but when the analyst has access to a decision maker's unconditional stochastic choices, possibly out of different menus. For the case of a single menu, the characterization in Rehbeck (2023) is different from that in Theorem 1 and is stated in terms of the non-existence of a possibly mixed deviation. Azrieli and Rehbeck (2022) study a similar question to ours in the context of stochastic choice. In their setting, the analyst has access to a marginal distribution over a decision maker's choices and a marginal distribution over the menus out of which the decision maker made her choices. Azrieli and Rehbeck (2022) show that the marginal distributions are consistent if and only if the marginal over choices is in the core of the game induced by the marginal over menus.

A literature in decision theory and experimental economics studies when choices can be rationalized via costly information acquisition and whether the choices can be used to identify the information acquisition costs (see, e.g., Caplin and Dean (2015), Caplin et al. (2017), Chambers et al. (2020), Dewan and Neligh (2020), Denti (2022)). Like we do, many of these papers assume that the decision maker's utility is known. More recently, assuming that the analyst has access to state-dependent stochastic choice data, Caplin et al. (2023) study when choices can be rationalized as if the agent has access to some information before choosing her actions. Whereas their analyst has access to a richer data set, they require consistency of the information structure across a family of decision problems.

Arieli et al. (2021) and Morris (2020) characterize joint distributions over posterior beliefs that are consistent with some information structure.[2] Both papers cast the problem as one of distributions with given marginals: they take as given a profile of marginal distributions over posterior beliefs with the same mean and characterize when a joint distribution with the given marginals exists that is consistent with information.

Finally, Vohra et al. (2023) study reduced-form implementation in a Bayesian persuasion in which the sender and the receiver care only about the posterior mean of the states. They leverage the mean preserving spread property to write a linear programming problem for the sender that only depends on the marginal distribution over actions. Beyond the posterior mean setting, they do not provide a characterization of the set of implementable marginal action distributions.

2 Model

Anticipating our multi-agent results in Section 4, our notation below presumes multiple agents. We then specialize it to the single-agent case in Section 3:

Base game: An incomplete information *base game*, G, is defined as follows. We are given a set of N players, $[N]$=$\{1,\ldots,N\}$. Each player $i \in [N]$ chooses an action from the finite set A_i. Payoffs $u_i(a,\theta)$ depend on the action profiles $a \in A \equiv \times_{i\in[N]} A_i$ and the state of the world, θ, an element of the finite set Θ.[3] The players share a common prior $\mu_0 \in \Delta(\Theta)$ over the state of the world. That is, $G = \langle \Theta, (A_i, u_i)_{i\in[N]}, \mu_0 \rangle$.

Bayes correlated equilibrium: An *outcome* is a joint distribution over action profiles and states of the world, $\pi \in \Delta(A \times \Theta)$. We are concerned with those outcomes that are consistent with non-cooperative play of the base game, where the solution concept is Bayes Nash equilibrium. The notion of Bayes correlated equilibrium in Bergemann and Morris (2016) captures the set of outcomes that are consistent with (Bayes Nash) equilibrium of the base game under *some* information structure:

[2] Whereas Arieli et al. (2021) study the binary-state case, the characterization in Morris (2020) requires no such assumption.

[3] As we explain in Section 3 our single-agent characterization extends to the case in which Θ and A are infinite (see Remark 1). However, the set of finitely many states and actions allows us to provide a *sharper* characterization.

Definition 1 (Bayes correlated equilibrium). *An outcome distribution* $\pi \in \Delta(A \times \Theta)$ *is a* Bayes correlated equilibrium *of base game* $G = \langle \Theta, (A_i, u_i)_{i\in[N]}, \mu_0 \rangle$, *if for all agents* $i \in [N]$, *actions* $a_i, a_i' \in A_i$, *the following holds*

$$\sum_{(a_{-i},\theta)} \pi(a_i, a_{-i}, \theta) \left[u_i(a_i, a_{-i}, \theta) - u_i(a_i', a_{-i}, \theta) \right] \geq 0, \tag{O}$$

and for all $\theta \in \Theta$

$$\sum_{a \in A} \pi(a, \theta) = \mu_0(\theta). \tag{M$_\Theta$}$$

Let $\mathrm{BCE}(\mu_0)$ *denote the set of Bayes correlated equilibria.*

In words, a Bayes correlated equilibrium is an outcome distribution that satisfies a series of *obedience* constraints (O) and a *martingale* condition (M$_\Theta$). The first ensures each player's best response condition under *some* information structure, whereas the second ensures the existence of an information structure that is consistent with the players' prior information. Note that any Bayes correlated equilibrium $\pi \in \Delta(A \times \Theta)$ induces two marginal distributions, $(\pi_\Theta, \pi_A) \in \Delta(\Theta) \times \Delta(A)$. The definition of Bayes correlated equilibrium implies that the primitive base game G pins down π_Θ, but not necessarily π_A.

Information Design with Given Marginals: We take the point of view of an analyst who knows the base game, but not the information structure under which the base game is played. The analyst is also endowed with information about the actions taken by the players. The analyst's goal is to determine whether this information is consistent with non-cooperative play of the base game under some information structure.

We consider two kinds of information the analyst may have about the players' actions, which are equivalent in the single-agent setting. In the first case, the analyst is endowed with a distribution over action profiles, $\nu_0 \in \Delta(A)$. In the second case, the analyst is endowed with a profile of action distributions, one for each player, that is, $\overline{\nu}_0 = (\nu_{0,1}, \ldots, \nu_{0,N}) \in \times_{i\in[N]} \Delta(A_i)$.

In each of these cases, the analyst wants to ascertain whether a Bayes correlated equilibrium $\pi \in \mathrm{BCE}(\mu_0)$ exists such that π_A coincides with the analyst's information about the players' actions (i.e., $\pi_A = \nu_0$ or $\times_{i\in[N]} \pi_{A_i} = \overline{\nu}_0$). In this case, we say that the marginals (μ_0, ν_0) are BCE-consistent or that the profile of marginal distributions $(\mu_0, \overline{\nu}_0)$ are M-BCE-consistent. Definition 2 records this for future reference:

Definition 2 (BCE- and M-BCE-consistent marginals). *Say that* (μ_0, ν_0) *are* BCE-consistent *if a Bayes correlated equilibrium* $\pi \in \mathrm{BCE}(\mu_0)$ *exists such that* $\pi_A = \nu_0$.

Similarly, we say that $(\mu_0, \overline{\nu}_0)$ *are* M-BCE-consistent *if a Bayes correlated equilibrium* $\pi \in \mathrm{BCE}(\mu_0)$ *exists such that for all players* $i \in [N]$, $\pi_{A_i} = \nu_{0,i}$.

Note that if (μ_0, ν_0) are BCE-consistent, then letting $\nu_{0,i}$ denote the marginal of ν_0 over A_i, we have that $(\mu_0, \nu_{0,1}, \ldots, \nu_{0,N})$ are M-BCE-consistent.

Constrained Optimal Transport We close this section by noting a connection with optimal transport. Given (μ_0, ν_0), let $\Pi(\mu_0, \nu_0)$ denote the set of joint distributions $\pi \in \Delta(A \times \Theta)$ withmarginals (μ_0, ν_0), i.e., $(\pi_\Theta, \pi_A) = (\mu_0, \nu_0)$. Note that $\Pi(\mu_0, \nu_0)$ is always nonempty, e.g., the joint distribution $\pi(a, \theta) = \nu_0(a)\mu_0(\theta)$ satisfies the marginal constraints. Instead, the subset $\Pi_{\mathrm{O}}(\mu_0, \nu_0)$ of $\Pi(\mu_0, \nu_0)$ that satisfies the obedience constraints (O) may be empty. Thus, the characterization of the set of BCE-consistent marginals (μ_0, ν_0) is equivalent to the characterization of when the feasible set of a *constrained* optimal transport problem–in this case $\Pi_{\mathrm{O}}(\mu_0, \nu_0)$–is nonempty.[4]

3 Single-agent case

In this section we characterize the set of BCE-consistent marginals in the case of a single agent, that is, $N = 1$. For this reason, in what follows we remove the index $i = 1$ from the action set and the utility function.

Distributions over posteriors and stochastic choice An outcome distribution $\pi \in \Delta(A \times \Theta)$ with marginals (μ_0, ν_0) induces two conditional probability systems: The first, $\{\mu(\cdot|a) \in \Delta(\Theta) : a \in A\}$, describes the agent's beliefs conditional on action a and satisfies for all actions $a \in A$,

$$\nu_0(a)\mu(\theta|a) = \pi(a, \theta).$$

In this case, one can view ν_0 as a distribution over posteriors and the belief system $(\mu(\cdot|a))_{a \in A}$ as its support.

The second, $\{\sigma(\cdot|\theta) \in \Delta(A) : \theta \in \Theta\}$, describes the agent's actions conditional on state θ and satisfies for all states $\theta \in \Theta$,

$$\mu_0(\theta)\sigma(a|\theta) = \pi(a, \theta).$$

The collection $\{\sigma(\cdot|\theta) : \theta \in \Theta\}$ is what the stochastic choice literature dubs the agent's stochastic choice rule.

[4] In their study of credible Bayesian persuasion, Lin and Liu (2022) characterize the set of credible outcome distributions by noting a connection with optimal transport. In their case, to check whether a given message distribution λ_M is implementable, it must be that no other joint distribution over states and messages that respects the given marginals exists and is preferred by the sender to λ_M.

The analysis that follows characterizes the set of BCE-consistent marginals relying on the belief system, $\{\mu(\cdot|a) \in \Delta(\Theta) : a \in A\}$. Instead, the stochastic choice rule $\sigma(\cdot|\theta)$ is the focus of the analysis in Section 3.1.

The action marginal as a distribution over posteriors Given marginals (μ_0, ν_0), the goal is to determine whether a belief system $\{\mu(\cdot|a) : a \in A\}$ exists that satisfies for all states $\theta \in \Theta$

$$\sum_{a \in A} \nu_0(a)\mu(\theta|a) = \mu_0(\theta), \tag{BP$_{\mu_0}$}$$

and for all $a, a' \in A$,

$$\sum_{\theta \in \Theta} \nu_0(a)\mu(\theta|a)\left[u(a,\theta) - u(a',\theta)\right] \geq 0. \tag{O$_\mu$}$$

For an action a, let $\Delta^*(a)$ denote the set of beliefs under which a is optimal.[5] Then, Equations BP$_{\mu_0}$ and O$_\mu$ require that (i) ν_0 induces a Bayes plausible distribution over posteriors and (ii) for all actions a, the *posterior belief* $\mu(\cdot|a)$ is an element of $\Delta^*(a)$. Under this interpretation, the action distribution ν_0 describes the frequency with which inducing beliefs in $\Delta^*(a)$ is necessary. Unsurprisingly, some of the conditions in Theorem 1 below also check that ν_0 satisfies a version of the martingale condition (Aumann et al., 1995; Kamenica and Gentzkow, 2011).

Theorem 1 characterizes the set of BCE-consistent marginals:

Theorem 1 (BCE-consistency). *The pair (μ_0, ν_0) is BCE-consistent if and only if for all states $\theta \in \Theta$,*

$$\sum_{a \in A} \nu_0(a) \min_{\mu \in \Delta^*(a)} \mu(\theta) \leq \mu_0(\theta), \tag{1}$$

and for all pairs of actions $a', a'' \in A$,

$$\sum_{a \in A} \nu_0(a) \max_{\mu \in \Delta^*(a)} \sum_{\theta \in \Theta} \mu(\theta)\left[u(a',\theta) - u(a'',\theta)\right] \geq \sum_{\theta \in \Theta} \mu_0(\theta)\left[u(a',\theta) - u(a'',\theta)\right]. \tag{2}$$

The proof is in Appendix A. In what follows, we provide intuition for the statement in Theorem 1 and review the main steps of its proof.

Equation 1 can be interpreted through the lens of the martingale property of beliefs. As discussed before Theorem 1, the action distribution ν_0 describes the frequency with

[5] Formally, $\Delta^*(a) = \{\mu \in \Delta(\Theta) : (\forall a' \in A) \sum_{\theta \in \Theta} \mu(\theta)\left(u(a,\theta) - u(a',\theta)\right) \geq 0\}$.

which beliefs in $\Delta^*(a)$ must be induced to satisfy (BP_{μ_0}). For a given state $\theta \in \Theta$, the term

$$\underline{\mu}_a(\theta) \equiv \min_{\mu \in \Delta^*(a)} \mu(\theta),$$

describes the smallest probability that the agent can assign to state θ and action a be optimal. Thus, Equation 1 states that for (μ_0, ν_0) to be BCE-consistent, it must be that the average under ν_0 of these minimum probabilities, $\underline{\mu}_a(\theta)$, are below the prior probability of θ, $\mu_0(\theta)$. It is immediate that if for some state θ, Equation 1 does not hold, then (μ_0, ν_0) cannot be BCE-consistent.

As we argue next, Equation 2 can be interpreted through the lens of a martingale property for the utility differences, $u(a', \theta) - u(a'', \theta)$. That is, for all pairs of actions, a', a'', the agent's expected ranking over a' and a'' under the experiment that rationalizes (μ_0, ν_0) has to coincide with the agent's ex ante ranking over these actions, which is the right-hand side of Equation 2. Indeed, because Equation 2 must hold when we exchange the roles of a' and a'', we obtain that (μ_0, ν_0) must also satisfy that

$$\sum_{a \in A} \nu_0(a) \min_{\mu \in \Delta^*(a)} \sum_{\theta \in \Theta} \mu(\theta) \left[u(a', \theta) - u(a'', \theta)\right] \leq \sum_{\theta \in \Theta} \mu_0(\theta) \left[u(a', \theta) - u(a'', \theta)\right]. \quad (3)$$

That is, the ranking at the prior between a' and a'' must be in between the worst and best rankings under the "distribution over posteriors" ν_0.

This is most easily seen in the simple case that a'' is strictly optimal at the prior and $\{a', a''\}$ are the only actions in the support of ν_0. Because a' is in the support of ν_0, under a BCE π that satisfies the marginal constraints the agent must sometimes find it optimal to take action a' instead of a''. Note, however, that *on average* it must be the case that the agent finds action a'' better than a'. Consequently, under π, when the agent takes a'', the agent must prefer a'' over a' (weakly) more than at the prior. Because the left-hand side of Equation 2 selects beliefs in favor of a', it is immediate that if Equation 2 fails one cannot find an experiment in which the agent would take action a' with sufficiently high probability so as to match ν_0.

So far, we have argued that the conditions in Theorem 1 are necessary for (μ_0, ν_0) to be BCE-consistent. To explain why they are also sufficient, it is useful to review the main steps in the proof of Theorem 1. Key to our proof is the following result from Strassen (1965), which we record in present notation:

Observation 1 (Strassen (1965, Theorem 3 and Corollary 1)). *A conditional probability system $\{\mu(\cdot|a) \in \Delta(\Theta) : a \in A\}$ exists such that*

1. *For all actions $a \in A$, $\mu(\cdot|a) \in \Delta^*(a)$, and*
2. *For all states $\theta \in \Theta$, BP_{μ_0} holds,*

if and only if for all directions $c \in \mathbb{R}^{|\Theta|}$,

$$\sum_{a \in A} \nu_0(a) \max\{c^T \mu : \mu \in \Delta^*(a)\} \geq c^T \mu_0. \tag{4}$$

Whereas Theorem 3 in Strassen (1965) requires that Equation 4 holds for *all* directions in $\mathbb{R}^{|\Theta|}$, Theorem 1 states that verifying Equation 4 holds for *finitely* many directions is enough to conclude that (μ_0, ν_0) are BCE-consistent. To see this, note that Equations 1 and 2 correspond to Equation 4 for specific directions $c \in \mathbb{R}^{|\Theta|}$. Indeed, Equation 1 corresponds to $c = -e_\theta \in \mathbb{R}^{|\Theta|}$, where e_θ is the vector with a 1 in the θ-coordinate and 0 otherwise. Instead, Equation 2 corresponds to the direction $c = -d_{a',a''}$, where $d_{a',a''}$ is the vector with θ-coordinate $d_{a',a''}(\theta) = u(a', \theta) - u(a'', \theta)$.

To see why verifying that Equation 4 holds for directions $\{(-e_\theta)_{\theta \in \Theta}, (-d_{a',a''})_{a',a'' \in A}\}$ is enough to determine that Equation 4 holds for all directions $c \in \mathbb{R}^{|\Theta|}$, note the following. First, for a fixed action a', the directions $\{(-e_\theta)_{\theta \in \Theta}, (-d_{a',a''})_{a'' \in A}\}$ are the normal vectors that define the polyhedron $\Delta^*(a')$. Indeed, the directions $(-e_\theta)_{\theta \in \Theta}$ correspond to the condition that the elements of $\Delta^*(a')$ are non-negative, whereas the directions $(-d_{a',a''})_{a'' \in A}$ correspond to the condition that action a' is optimal for all beliefs in $\Delta^*(a')$. Second, it is immediate that in each of the maximization problems on the left hand side of Equation 4, the maximum is attained at an extreme point of $\Delta^*(a)$. Standard results in convex analysis then imply that if Equation 4 holds at all normal directions defining the polyhedra $\{\Delta^*(a) : a \in A\}$, then it holds for all directions (cf. Hiriart-Urruty and Lemaréchal, 2004).

We close Section 3 with a remark on the generality of the results in Strassen (1965). It can be skipped with no loss of continuity.

Remark 1 (Strassen, 1965)**.** *Theorem 3 and Corollary 1 in Strassen (1965) hold more generally than our current assumptions. In present notation, Corollary 1 applies whenever (i)* Θ *and* A *are compact metric spaces and the mapping* $a \mapsto \Delta^*(a)$ *from* A *to subsets of* $\Delta(\Theta)$ *is such that* $\cup_{a \in A}\{a\} \times \Delta^*(a)$ *is closed within* $A \times \Delta(\Theta)$ *endowed with the weak*-topology.*[6]

In other words, under the aforementioned assumptions, (an integral version of) Equation 4 characterizes the set of BCE-consistent marginals.[7] *The finite model allows us to provide*

[6]Instead, Strassen (1965, Theorem 3) requires that Θ is Polish, A be a convex compact topological vector space, and an appropriate measurability condition on the mapping $a \mapsto \sup\{\int c(\theta)\mu(d\theta) : \mu \in \Delta^*(a)\}$ for any continuous function c on Θ.

[7]To be precise, Equation 4 now becomes for all continuous functions $c : \Theta \mapsto \mathbb{R}$,

$$\int_\Theta c(\theta)\mu_0(d\theta) \leq \int_A \sup\left\{\int_\Theta c(\theta)\mu(d\theta) : \mu \in \Delta^*(a)\right\} \nu_0(da)$$

.

a sharper characterization by reducing the number of directions one needs to consider.

3.1 The core of Bayesian Persuasion

In this section we provide a different perspective on Theorem 1. Together with the marginal distributions, (μ_0, ν_0), we are given a distribution over posteriors $\tau \in \Delta(\Delta(\Theta))$ with mean equal to the prior μ_0. Proposition 1 below characterizes the set of such distributions over posteriors that can *implement* the marginal ν_0. Whereas this characterization does not substitute that in Theorem 1, it allows us to illustrate how one would go about constructing an information structure that implements ν_0. Along the way we also establish formal connections with the literature on stochastic choice. For this reason, we work with the agent's stochastic choice rule $\{\sigma(\cdot|\theta) : \theta \in \Theta\}$ instead of the belief system $\{\mu(\cdot|a) : a \in A\}$.

Obedient stochastic choice To understand the results that follow, it is useful to state the obedience and marginal conditions in terms of the stochastic choice rule: Given (μ_0, ν_0), we want a stochastic choice rule that satisfies for all actions $a \in A$

$$\sum_{\theta \in \Theta} \mu_0(\theta)\sigma(a|\theta) = \nu_0(a), \tag{M$_A$}$$

and for all $a, a' \in A$,

$$\sum_{\theta \in \Theta} \mu_0(\theta)\sigma(a|\theta)\left[u(a, \theta) - u(a', \theta)\right] \geq 0. \tag{O$_\sigma$}$$

Distributions over posteriors and stochastic choice Given a Bayes plausible distribution over posteriors $\tau \in \Delta(\Delta(\Theta))$, constructing a state-dependent stochastic choice rule is almost at hand. Almost because a Bayes plausible distribution over posteriors does not specify how the agent breaks ties when indifferent. Indeed, to a Bayes plausible distribution over posteriors, $\tau(\mu)$, we can associate a decision rule $\alpha : \Delta(\Theta) \mapsto \Delta(A)$, describing the probability $\alpha(a|\mu)$ with which the agent takes action a when her belief is μ. The pair (τ, α) determines a stochastic choice rule $\{\sigma(\cdot|\theta) : \theta \in \Theta\}$ as follows:

$$\sigma(a|\theta) = \sum_{\mu \in \Delta(\Theta)} \tau(\mu)\frac{\mu(\theta)}{\mu_0(\theta)}\alpha(a|\mu). \tag{5}$$

Equation 5 suggests that conditions under which a stochastic choice rule that satisfies M_A and O_σ are intimately related to the existence of a Bayes plausible distribution τ and a decision rule α that satisfy certain properties. In fact, the analysis that follows identifies conditions on Bayes plausible distributions over posteriors under which a decision rule exists that induces a stochastic choice rule–and hence a joint distribution $\pi \in \Delta(A \times \Theta)$–that satisfies all the constraints.

Distributions over posteriors as distributions over menus Given a Bayes plausible $\tau \in \Delta(\Delta(\Theta))$, one can construct a measure over subsets B of the set of actions A as follows. For each $\mu \in \Delta(\Theta)$, let $a^*(\mu)$ denote the agent's best response when her belief is μ. That is, $a^*(\mu) = \arg\max_{a \in A} \mathbb{E}_{\theta \sim \mu}[u(a, \theta)]$. For each $B \subseteq A$, define $\tau_A(B)$ as

$$\tau_A(B) = \tau\{\mu \in \Delta(\Theta) : a^*(\mu) = B\}. \tag{6}$$

In words, each action subset B has mass equal to the probability that τ induces a belief under which B is optimal.

Proposition 1 characterizes when the distribution over posteriors τ implements ν_0:

Proposition 1. *Suppose (μ_0, ν_0) are BCE-consistent. A Bayes plausible distribution over posteriors, $\tau \in \Delta(\Delta(\Theta))$, implements ν_0 if and only if for all $B \subseteq A$, the following holds*

$$\sum_{a \in B} \nu_0(a) \geq \sum_{C \subseteq B} \tau_A(C). \tag{7}$$

To interpret Equation 7, note the following. The left-hand side of Equation 7 is the probability under which the agent takes *some* action a in the set B. Instead, the right-hand side of Equation 7 is the probability under which the agent finds optimal *some* action in the set B (but no action that is not in B). Equation 7 then says that the frequency with which the agent takes actions in B has to be at least the frequency with which an action in B is optimal.

Remark 2 (A core interpretation). *Equation 7 implies that ν_0 is in the* core of the game *induced by the measure τ_A.*[8] *Indeed, given τ_A, define the* cooperative game (A, w_{τ_A}) *as follows. The set function $w_{\tau_A} : 2^A \mapsto \mathbb{R}$ is given by $w_{\tau_A}(B) = \sum_{C \subseteq B} \tau_A(C)$. Because $w_{\tau_A} \geq 0$, the core of the game (A, w_{τ_A}) is given by*

$$\text{Core}(w_{\tau_A}) = \left\{ p \in \Delta(A) : (\forall B \subseteq A) \sum_{a \in B} p(a) \geq w_{\tau_A}(B) \right\}.$$

The proof of Proposition 1 is based on the following graphical representation of the BCE-consistency problem depicted in Figure 1. Consider the following graph. Nodes are (i) the actions $a \in A$, (ii) the (non-empty) action subsets $B \subseteq A$ (i.e., the elements of $2^A \setminus \{\emptyset\}$), (iii) a source node s, and (iv) a sink node t. Edges are as follows. There is an edge of weight one between $a \in A$ and $B \subseteq A$ if and only if $a \in B$. There is an edge with weight $\nu_0(a)$ between the source s and a. Finally, there is an edge between

[8] Azrieli and Rehbeck (2022) also note the connection between stochastic menu choice and cooperative games.

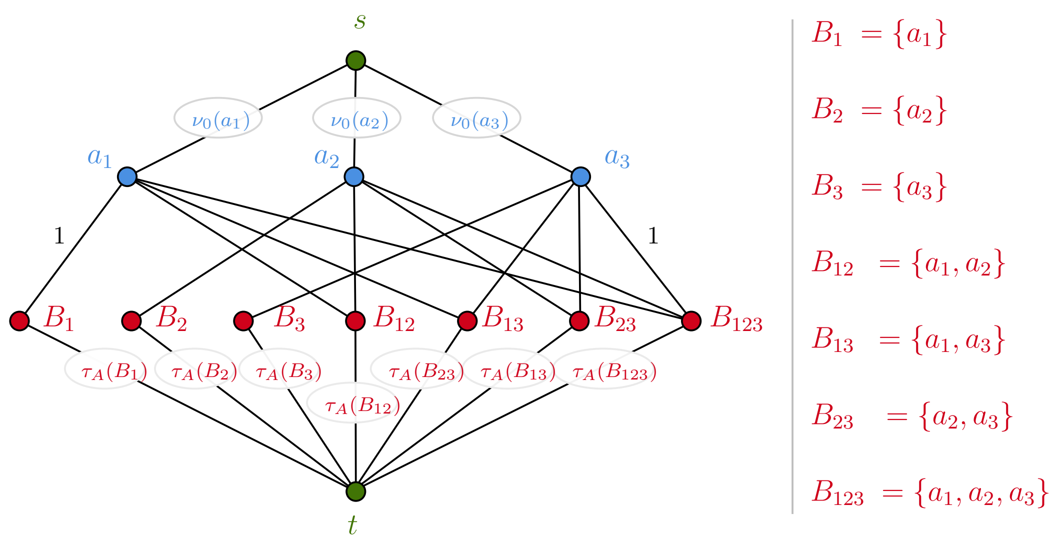

Figure 1: Graphical representation of the BCE-consistency problem with 3 actions.

$B \subseteq A$ and the sink t with weight $\tau_A(B)$. The condition in Equation 7 ensures that a feasible flow exists throughout the network.[9]

Proof of Proposition 1. It is immediate to show that if (μ_0, ν_0) are BCE-consistent, then a Bayes plausible distribution over posteriors τ exists such that Equation 7 holds.

Suppose that Equation 7 holds for all $B \subseteq A$. Azrieli and Rehbeck (2022, Proposition 9) implies that a conditional probability system $\alpha' : 2^A \mapsto \Delta(A)$ exists such that for all $a \in A$

$$\nu_0(a) = \sum_{B:a\in B} \tau_A(B)\alpha'(a|B). \tag{8}$$

The slight abuse of notation in the definition of the conditional probability system is justified since α' below plays the role of the decision rule in Equation 5.

We use the conditional probability system to create a s stochastic choice rule $\sigma : \Theta \mapsto \Delta(A)$ as follows:

$$\sigma(a|\theta) = \sum_{B:a\in B} \sum_{\mu:a^*(\mu)=B} \frac{\mu(\theta)}{\mu_0(\theta)}\tau(\mu)\alpha'(a|B).$$

The experiment has an intuitive explanation: We first draw a subset of actions B using the measure τ_A and then recommend to the agent which particular action she must take using the conditional probability system $\alpha'(\cdot|B)$.

[9]That is, a flow f such that $\sum_{a\in A} f(s,a) = \sum_{B\in 2^A} f(B,t) = 1$.

Define the information structure, $\pi \in \Delta(A \times \Theta)$ by letting $\pi(a, \theta) = \mu_0(\theta)\sigma(a|\theta)$. To see that it has the desired properties, note first that

$$\sum_{a \in A} \sigma(a|\theta) = \sum_{a \in A} \sum_{B: a \in B} \sum_{\mu: a^*(\mu) = B} \frac{\mu(\theta)}{\mu_0(\theta)} \tau(\mu)\alpha'(a|B)$$
$$= \sum_{B \subseteq A} \left(\sum_{a \in B} \alpha'(a|B) \right) \sum_{\mu: a^*(\mu) = B} \tau(\mu) \frac{\mu(\theta)}{\mu_0(\theta)} = \sum_{B \subseteq A} \sum_{\mu: a^*(\mu) = B} \tau(\mu) \frac{\mu(\theta)}{\mu_0(\theta)} = 1$$

Second, note that

$$\sum_{\theta \in \Theta} \pi(a, \theta) = \sum_{\theta \in \Theta} \mu_0(\theta)\sigma(a|\theta) = \sum_{\theta \in \Theta} \sum_{B: a \in B} \sum_{\mu: a^*(\mu) = B} \mu(\theta)\tau(\mu)\alpha'(a|B)$$
$$= \sum_{B: a \in B} \sum_{\mu: a^*(\mu) = B} \left(\sum_{\theta \in \Theta} \mu(\theta) \right) \tau(\mu)\alpha'(a|B) = \nu_0(a),$$

by Equation 8.

Finally, note that the experiment is obedient: If a is recommended with positive probability, then a set B exists such that $a \in B$ and μ such that $a^*(\mu) = B$ is in the support of τ, under which a is optimal. Because $\sigma(a|\theta)$ is obtained by averaging over beliefs in which a is optimal, it remains optimal. □

Appendix B provides an alternative proof of Proposition 1 using Gale's network flow theorem.

Connection to stochastic choice: The proof of Proposition 1 connects two sets of conditional distribution over choices that arise in the stochastic choice literature: stochastic choices conditional on a state of the world–denoted by σ in the proof–and stochastic choices out of a menu–denoted by α' in the proof. Indeed, the measure τ_A can be interpreted as the frequency with which the agent faces different menus–action subsets in this case–whereas the measure ν_0 represents the frequency with which the agent makes different choices. In other words, the pair (τ_A, ν_0) is analogous to the data set in Azrieli and Rehbeck (2022). Our ultimately goal, however, is to obtain the agent's stochastic choice rule, which we obtain relying on the Bayes' plausibility of τ.

4 Applications

We consider in this section two applications of Theorem 1 to simple multi-agent settings. Section 4.1 studies under what conditions a pair of marginal distributions (μ_0, ν_0)

can be rationalized by a *public* information structure. Section 4.2 shows that Theorem 1 characterizes the set of M-BCE-consistent marginals.

4.1 When is information public?

We consider in this section the following multiplayer game. We assume $N \geq 1$ and that each player's utility function depends only her own action and the state of the world.[10] That is, for all players $i \in \{1, \ldots, N\}$, all action profiles $(a_i, a_{-i}) \in A$, and states of the world $\theta \in \Theta$,

$$u_i(a_i, a_{-i}, \theta) = u_i(a_i, \theta).$$

The analyst, who knows the base game G and the marginal distribution of play $\nu_0 \in \Delta(A)$, wants to ascertain whether the distribution of play ν_0 can be rationalized by a *public* information structure (i.e., the players publicly observe the realization of a common signal structure before play).

As we show next, Theorem 1 can be applied to address this question. In what follows, we rely on the following definition:

Definition 3 (Public BCE-consistency). *The pair* (μ_0, ν_0) *is* public BCE-consistent *if: (i)* (μ_0, ν_0) *are BCE-consistent, and (ii) a BCE* $\pi \in \mathrm{BCE}(\mu_0) \cap \Pi(\mu_0, \nu_0)$ *exists, whose information structure uses public signals alone.*

Consider now an auxiliary single-agent base game $\bar{G} = \langle \Theta, (A, \bar{u}), \mu_0 \rangle$. In this game, a player with payoff $\bar{u}(a) = \sum_{i=1}^{N} u_i(a_i, \theta)$ chooses an action $a \in A = \times_{i \in N} A_i$ under incomplete information about θ.

The following result is an immediate corollary of Theorem 1 and the focus on public signals:

Corollary 1. (μ_0, ν_0) *are public BCE-consistent if and only if* (μ_0, ν_0) *are BCE-consistent in base game* $\bar{G}$.

Because of the focus on public signal structures, the analysis of the multi-agent game reduces to the analysis of a single-agent problem. To see this, in a slight abuse of notation, let $A^*(\mu)$ denote the set of actions that the agent with payoff $\bar{u}$ finds optimal when their belief is μ. It is immediate that $A^*(\mu) = \times_{i \in N} a_i^*(\mu)$, where for each player i, $a_i^*(\mu)$ denotes the set of actions player i finds optimal when her belief is μ. That is, the profile $a = (a_1, \ldots, a_N) \in A$ is optimal for the agent with payoff $\bar{u}$ if and only if action a_i is optimal for agent i, for all $i \in [N]$. And, given a posterior belief μ, any distribution of (optimal) action profiles that the agent with payoff $\bar{u}$ can generate, can

[10] Arieli et al. (2021) dub this setting first-order Bayesian persuasion.

also be generated by the players using a public correlation device or by duplicating signal realization, and vice versa.[11] Notice that this equivalence no longer holds if either information is not public, or the players' utilities are interdependent.

4.2 Ring-network games

We consider here ring-network games as in Kneeland (2015), extended to account for incomplete information. A ring-network game is a base game G in which player's payoffs satisfy the following:

$$\begin{aligned} u_1(a,\theta) &= \tilde{u}_1(a_1,\theta) \\ (\forall i \geq 2) u_i(a,\theta) &= \tilde{u}_i(a_{i-1},a_i). \end{aligned} \tag{RN-P}$$

In words, player 1 cares about their action and the state of the world, whereas for $i \geq 2$ player i cares about their action and that of player $i-1$. Ring-network games are used in the experimental literature that measures players' higher order beliefs to identify departures from Nash equilibrium.

The analyst knows the ring-network base game and for each player i, player i's action distribution, $\nu_{0,i} \in \Delta(A_i)$. The analyst wants to ascertain whether $(\mu_0, \overline{\nu}_0)$ is M-BCE-consistent. Relying on Theorem 1 and the ring-network structure, Proposition 2 characterizes the set of M-BCE-consistent marginals:

Proposition 2 (M-BCE-consistency in ring-network games). *The profile $(\mu_0, \overline{\nu}_0)$ is M-BCE-consistent for the ring-network game $(\tilde{u}_i)_{i=1}^N$ if and only if the following holds:*

1. *$(\mu_0, \nu_{0,1})$ are BCE-consistent in the base game $\langle \Theta, A_1, \tilde{u}_1, \mu_0 \rangle$,*
2. *For all $i \geq 2$, $(\nu_{0,i-1}, \nu_{0,i})$ are BCE-consistent in the base game $\langle A_{i-1}, A_i, \tilde{u}_i, \nu_{0,i-1} \rangle$.*

Similar to Corollary 1, Proposition 2 exploits the structure of the ring-network game to reduce it to a series of single-agent problems in which except for player 1, the states are given by the actions of the preceding player and the prior distribution over this state space by the marginal over actions of the preceding player. Indeed, for $i \geq 2$, BCE-consistency of $(\nu_{0,i-1}, \nu_{0,i})$ implies that an information structure exists that rationalizes player i's choices as the outcome of some information structure under "prior" $\nu_{0,i-1}$, whereas BCE-consistency of $(\nu_{0,i-2}, \nu_{0,i-1})$[12] guarantees that the "prior" $\nu_{0,i-1}$

[11] For example, suppose that the signal realization s induces the posterior belief μ. Suppose also that under μ, the agent with payoff $\bar{u}$ selects the two optimal action profiles $a, a' \in A^*(\mu)$ with equal probability. The same distribution of actions can be generated by the players: Indeed, one can "split" the signal s into two new signals, s' and s'', such that both new signals induce the same posterior belief μ, and each of them is sent with half the probability of the original signal s. If whenever s' and s'' are realized, each agent acts according to her corresponding optimal action in the profiles a and a', respectively, the distribution over actions will coincide with that of the agent with payoff $\bar{u}$.

[12] With the understanding that $\nu_{0,0} = \mu_0$.

is consistent with player $i-1$ observing the outcome of some information structure given their belief $\nu_{0,i-2}$.

References

Aguiar, V. H., M. J. Boccardi, N. Kashaev, and J. Kim (2018): "Does random consideration explain behavior when choice is hard? Evidence from a large-scale experiment," *arXiv preprint arXiv:1812.09619*.

Arieli, I., Y. Babichenko, F. Sandomirskiy, and O. Tamuz (2021): "Feasible joint posterior beliefs," *Journal of Political Economy*, 129, 2546–2594.

Aumann, R. J., M. Maschler, and R. E. Stearns (1995): *Repeated games with incomplete information*, MIT press.

Azrieli, Y. and J. Rehbeck (2022): "Marginal stochastic choice," *arXiv preprint arXiv:2208.08492*.

Barseghyan, L., M. Coughlin, F. Molinari, and J. C. Teitelbaum (2021): "Heterogeneous choice sets and preferences," *Econometrica*, 89, 2015–2048.

Bergemann, D. and S. Morris (2016): "Bayes correlated equilibrium and the comparison of information structures in games," *Theoretical Economics*, 11, 487–522.

Border, K. C. (1991): "Implementation of reduced form auctions: A geometric approach," *Econometrica: Journal of the Econometric Society*, 1175–1187.

Caplin, A. and M. Dean (2015): "Revealed preference, rational inattention, and costly information acquisition," *American Economic Review*, 105, 2183–2203.

Caplin, A., M. Dean, and J. Leahy (2017): "Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy," Tech. rep., National Bureau of Economic Research.

Caplin, A., D. J. Martin, and P. Marx (2023): "Rationalizable Learning," Tech. rep., National Bureau of Economic Research.

Chambers, C. P., C. Liu, and J. Rehbeck (2020): "Costly information acquisition," *Journal of Economic Theory*, 186, 104979.

Dardanoni, V., P. Manzini, M. Mariotti, and C. J. Tyson (2020): "Inferring cognitive heterogeneity from aggregate choices," *Econometrica*, 88, 1269–1296.

Denti, T. (2022): "Posterior separable cost of information," *American Economic Review*, 112, 3215–3259.

DEWAN, A. AND N. NELIGH (2020): "Estimating information cost functions in models of rational inattention," *Journal of Economic Theory*, 187, 105011.

GALE, D. (1957): "A theorem on flows in networks," *Pacific J. Math*, 7, 1073–1082.

GRABISCH, M. ET AL. (2016): *Set functions, games and capacities in decision making*, vol. 46, Springer.

HIRIART-URRUTY, J.-B. AND C. LEMARÉCHAL (2004): *Fundamentals of convex analysis*, Springer Science & Business Media.

KAMENICA, E. AND M. GENTZKOW (2011): "Bayesian persuasion," *American Economic Review*, 101, 2590–2615.

KNEELAND, T. (2015): "Identifying higher-order rationality," *Econometrica*, 83, 2065–2079.

KOH, P. S. (2022): "Stable Outcomes and Information in Games: An Empirical Framework," *arXiv preprint arXiv:2205.04990.*

LIN, X. AND C. LIU (2022): "Credible persuasion," *arXiv preprint arXiv:2205.03495.*

MAGNOLFI, L. AND C. RONCORONI (2019): "Estimation of discrete games with weak assumptions on information," *University of Warwick, Department of Economics.*

MATTHEWS, S. A. (1984): "On the implementability of reduced form auctions," *Econometrica: Journal of the Econometric Society*, 1519–1522.

MORRIS, S. E. (2020): "No trade and feasible joint posterior beliefs," Tech. rep., Massachusetts Institute of Technology.

MYERSON, R. B. (1982): "Optimal coordination mechanisms in generalized principal–agent problems," *Journal of Mathematical Economics*, 10, 67–81.

REHBECK, J. (2023): "Revealed Bayesian expected utility with limited data," *Journal of Economic Behavior & Organization*, 207, 81–95.

STRASSEN, V. (1965): "The existence of probability measures with given marginals," *The Annals of Mathematical Statistics*, 36, 423–439.

SYRGKANIS, V., E. TAMER, AND J. ZIANI (2017): "Inference on auctions with weak assumptions on information," *arXiv preprint arXiv:1710.03830.*

VOHRA, A., J. TOIKKA, AND R. VOHRA (2023): "Bayesian persuasion: Reduced form approach," *Journal of Mathematical Economics*, 102863.

A Omitted proofs

A.1 Proof of Theorem 1

Preliminaries Before stating the proof of Theorem 1, we collect some definitions and results from convex analysis that we use in the proof.

Define $C(a') = \{(-e_\theta)_{\theta\in\Theta}, (-d_{a',a''})_{a''\in A}\}$ to be the normal directions to the polyhedron $\Delta^*(a')$, which is implicitly defined as the set of vectors x in $\mathbb{R}^{|\Theta|}$ that satisfy:

$$(\forall\theta \in \Theta)(-e_\theta)^T x \leq 0 \qquad (9)$$
$$(\forall a'' \in A)(-d_{a',a''})^T x \leq 0.$$

We are omitting the condition that $\sum_{\theta\in\Theta} x(\theta) = 1$, but this is irrelevant in what follows.

Recall that for $x \in \mathbb{R}^{|\Theta|}$, the normal cone of $\Delta^*(a)$ at x, $N(x|\Delta^*(a))$, is defined as

$$N(x|\Delta^*(a)) = \{c \in \mathbb{R}^{|\Theta|} : (\forall x' \in \Delta^*(a))c^T x' \leq c^T x\}. \qquad (10)$$

That is, the normal cone of $\Delta^*(a)$ at x is the set of directions c for which x solves $\max\{c^T x' : x' \in \Delta^*(a)\}$. Importantly, the normal cone of a polyhedron, like $\Delta^*(a)$, satisfies the following property. To state it, recall that given a set of points C, the cone of C is defined as $\text{cone}(C) = \{\sum_{j=1}^J \alpha_j c_j : J < \infty, c_j \in C, \alpha_j \geq 0\}$.

Lemma 1 (Hiriart-Urruty and Lemaréchal (2004, Example 5.2.6(b))). *Suppose $\mu \in \Delta^*(a)$ and let $B(\mu) = \{c \in C(a) : c^T\mu = 0\}$. Then, $N(\mu|\Delta^*(a)) = \text{cone}(B(\mu))$.*

Proof of Theorem 1. Necessity of Equations 1 and 2 follows from Strassen (1965, Theorem 3).

We now argue sufficiency. Given Observation 1, it suffices to show that Equations 1 and 2 imply Equation 4 holds for all $c \in \mathbb{R}^{|\Theta|}$.

For fixed c, we can write Equation 4 as follows:

$$\sum_{a\in A} \nu_0(a) \max_{\mu\in\Delta^*(a)} c^T(\mu - \mu_0) \geq 0. \qquad (11)$$

Thus, Equation 4 holds for all directions $c \in \mathbb{R}^{|\Theta|}$ if and only if

$$\min_{c\in\mathbb{R}^{|\Theta|}} \sum_{a\in A} \nu_0(a) \max_{\mu\in\Delta^*(a)} c^T(\mu - \mu_0) \geq 0. \qquad (12)$$

Note that we can replace $\Delta^*(a)$ for the set of extreme points of $\Delta^*(a)$, $\Delta^*_E(a)$ in Equation 12. That is,

$$\min_{c\in\mathbb{R}^{|\Theta|}} \sum_{a\in A} \nu_0(a) \max_{\mu\in\Delta^*_E(a)} c^T(\mu-\mu_0) \geq 0. \tag{13}$$

Now, let $E = \prod\{\Delta^*_E(a) : a \in A\}$. For $\bar{\mu}_e \equiv (\mu_{e,a})_{a\in A} \in E$, let

$$C(\bar{\mu}_e) = \{c \in \mathbb{R}^{|\Theta|} : (\forall a \in A)\ c^T\mu_{e,a} = \max_{\mu\in\Delta^*(a)} c^T\mu\}.$$

Then, we can write the left hand side of Equation 13 as follows:

$$\min_{\bar{\mu}_e\in E} \min_{c\in C(\bar{\mu}_e)} \sum_{a\in A} \nu_0(a) \max_{\mu\in\Delta^*_E(a)} c^T(\mu-\mu_0). \tag{14}$$

Note that for each $\bar{\mu}_e \in E$

$$C(\bar{\mu}_e) = \cap_{a\in A} N\left(\mu_{e,a}|\Delta^*(a)\right), \tag{15}$$

and by Lemma 1, $N(\mu_{e,a}|\Delta^*(a)) \subseteq C(a)$. Thus, Equations 1 and 2 ensure that the term inside $\min_{\bar{\mu}_e\in E}$ is non-negative, so that Equation 4 holds for all $c \in \mathbb{R}^{|\Theta|}$.

□

A.2 Proof of Proposition 2

Proof of Proposition 2. In the ring-network base game, for a joint distribution $\pi \in \Delta(A\times\Theta)$, the obedience constraints can be written as follows:

$$(\forall a_1, a_1' \in A_1) \sum_{\theta\in\Theta} \pi_{\Theta\times A_1}(a_1,\theta)\left(\tilde{u}(a_1,\theta) - \tilde{u}(a_1',\theta)\right) \geq 0$$

$$(\forall i \in \{2,\dots,N\})(\forall a_i, a_i' \in A_i) \sum_{a_{i-1}\in A_{i-1}} \pi_{A_{i-1},i}(a,\theta)\left(\tilde{u}(a_{i-1},a_i) - \tilde{u}(a_{i-1},a_i')\right) \geq 0,$$

where $\pi_{\Theta\times A_1}$ is the marginal of π over $\Theta\times A_1$ and similarly for $i \geq 2$, $\pi_{A_{i-1}\times A_i}$ is the marginal of π over $A_{i-1}\times A_i$. Thus, it is immediate that the conditions in Proposition 2 are necessary for $(\mu_0, \overline{\nu}_0)$ to be M-BCE-consistent.

For sufficiency, note that Theorem 1 implies that under the conditions of Proposition 2, $\left(\pi_{\Theta\times A_1}, \dots, \pi_{A_{N-1}\times A_N}\right)$ exist each of which satisfy the respective marginal conditions and obedience constraints.

Given these distributions, define $\hat{\pi} \in \Delta(A\times\Theta)$ as follows: for each $(a,\theta) \in A\times\Theta$

$$\hat{\pi}(a,\theta) = \pi_{A_1\times\Theta}(a_1,\theta)\pi_{A_1\times A_2}(a_2|a_1) \times \dots \pi_{A_{N-1}\times A_N}(a_N|a_{N-1}), \tag{16}$$

where abusing notation we let for $i \geq 2$, $\pi_{A_{i-1}\times A_i}(\cdot|a_{i-1})$ denote the distribution $\pi_{A_{i-1}\times A_i}$ conditional on $a'_{i-1} = a_{i-1}$.

Note that $\hat{\pi}(a,\theta)$ satisfies the obedience constraints of player 1 if and only if $\pi_{A_1\times\Theta}(\cdot)$ does. Indeed, for all a_1, a'_1, we have

$$\sum_{a_{-1},\theta} \hat{\pi}(a_1,a_{-1},\theta)\left(\tilde{u}_1(a_1,\theta)-\tilde{u}_1(a'_1,\theta)\right) =$$

$$\sum_{\theta} \pi_{A_1\times\Theta}(a_1,\theta)\left(\tilde{u}_1(a_1,\theta)-\tilde{u}_1(a'_1,\theta)\right) \sum_{(a_2,\ldots,a_N)} \prod_{i=2}^{N} \pi_{A_{i-1}\times A_i}(a_i|a_{i-1}) =$$

$$\sum_{\theta} \pi_{A_1\times\Theta}(a_1,\theta)\left(\tilde{u}_1(a_1,\theta)-\tilde{u}_1(a'_1,\theta)\right). \tag{17}$$

Consider now player $i \geq 2$. For simplicity, fix $i = 2$–the rest of the players follow immediately. Then, let $a_2, a'_2 \in A_2$. We want to check that π satisfies the obedience constraint of player 2 if and only if $\pi_{A_1\times A_2}$ does.

$$\sum_{a_{-2},\theta} \hat{\pi}(a_2,a_{-2},\theta)\left(\tilde{u}_2(a_1,a_2)-\tilde{u}_2(a_1,a'_2)\right) =$$

$$\sum_{a_1,\theta} \pi_{A_1\times\Theta}(a_1,\theta)\pi_{A_1\times A_2}(a_2|a_1)\left(\tilde{u}_2(a_1,a_2)-\tilde{u}_2(a_1,a'_2)\right) \sum_{(a_3,\ldots,a_N)} \prod_{i=3}^{N} \pi_{A_{i-1}\times A_i}(a_i|a_{i-1}) =$$

$$\sum_{a_1\in A_1}\left(\sum_{\theta} \pi_{A_1\times\Theta}(a_1,\theta)\right) \pi_{A_1\times A_2}(a_2|a_1)\left(\tilde{u}_2(a_1,a_2)-\tilde{u}_2(a_1,a'_2)\right) =$$

$$\sum_{a_1\in A_1} \nu_{01}(a_1)\pi_{A_1\times A_2}(a_2|a_1)\left(\tilde{u}_2(a_1,a_2)-\tilde{u}_2(a_1,a'_2)\right) =$$

$$\sum_{a_1\in A_1} \pi_{A_1\times A_2}(a_1,a_2)\left(\tilde{u}_2(a_1,a_2)-\tilde{u}_2(a_1,a'_2)\right),$$

where the third equality follows from the assumption that $\pi_{A_1\times\Theta}$ satisfies the marginal constraints for player 1. □

B A demand-supply interpretation

We provide here an alternative, but still network based, proof of Proposition 1 using the fundamental results of Gale (1957) on demand and supply in a network.

Flows in networks: The problem in Gale (1957) can be described as follows. Given a graph (V, E), suppose that to each node $v \in V$ corresponds a real number $d(v)$. If

$d(v) > 0$ we interpret $|d(v)|$ as the *demand* of node v for some homogenous good. If $d(v) < 0$ we interpret $|d(v)|$ as the *supply* of the good by v. To each edge $(v, v') \in E$ correspond two nonnegative real numbers $c(v, v')$ and $c(v', v)$, the capacity of this edge, which assign an upper bound to the possible flow of the good from v to v' and from v' to v, respectively. The demand $d_{(\nu_0,\tau)}$ is called *feasible* if there is a flow in the graph such that the flow along each edge is no greater than its capacity, and the net flow into (out of) each node is at least (at most) equal to the demand (supply) at that node. The demand problem identifies the conditions under which a given demand $d_{(\nu_0,\tau)}$ is feasible in the graph.

BCE-consistency and the demand problem: Fix a Bayes plausible distribution $\tau \in \Delta(\Delta(\theta))$ and denote its support by $T = \text{supp}\ \tau$. Because A is finite, it is without loss of generality to assume that T is a finite set (Myerson, 1982; Kamenica and Gentzkow, 2011).

The conditions in Gale (1957) can be used to check whether τ implements the action distribution ν_0: That is, that a decision rule α exists that together with τ define an obedient experiment (see Equation 5). To that end, we construct a (bipartite) graph in which posterior beliefs serve as supply nodes, and actions serve as demand nodes. That is, the homogeneous good in our construction can be thought of as probability that "flows" from induced posterior beliefs to actions. The construction of the graph guarantees that if the demand is feasible, we can specify choices for the agent such that the probabilities according to which she breaks ties between optimal actions at each posterior belief induce the (ex-ante) desired action distribution ν_0.

Formally, define the graph $G_P(\tau) = (A \cup T, E)$ as follows. To each action $a \in A$ corresponds a node that demands the marginal probability of a, i.e. $d_{(\nu_0,\tau)}(a) = \nu_0(a)$. To each belief $\mu \in T$ corresponds a node that supplies the probability with which μ is realized in τ, i.e. $d_{(\nu_0,\tau)}(\mu) = -\tau(\mu)$. For any belief-action pair (μ, a), an edge $(\mu, a) \in E$ exists between the nodes μ and a if and only if action a is optimal under posterior belief μ, that is if and only if, $a \in a^*(\mu)$. Finally, for any edge $(\mu, a) \in E$, the edge's flow capacity is given by $c(\mu, a) = \infty$ and $c(a, \mu) = 0$. That is, there is no upper bound on the flow from μ to a, but there cannot be a flow from a to μ. We denote the flow from node μ to node a by $f(\mu, a)$. If there is no edge between μ and a then $f(\mu, a) = 0$. The right-hand side panel of Figure 2 illustrates the graph G_P.

Proposition 3 motivates the connection between our problem and that in Gale (1957).

Proposition 3 (Feasibility and BCE-consistency). *The Bayes plausible distribution over posteriors τ implements ν_0 if and only if $d_{(\nu_0,\tau)}$ is feasible on $G_P(\tau)$.*

The proof of Proposition 3 relies on the following lemma:

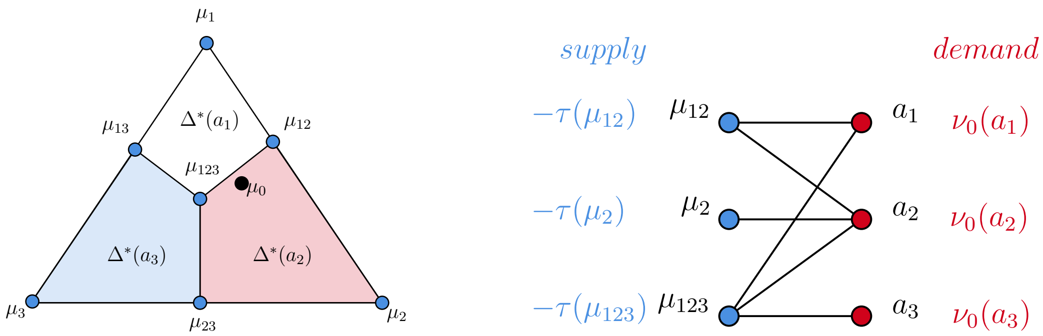

Figure (a) Figure (b)

Figure 2: Illustration of the supply-demand proof of Proposition 1 with $|A| = |\Theta| = 3$. The simplex on the left-hand side depicts the optimal action(s) for each posterior belief. The graph on the right-hand side corresponds to the Bayes plausible distribution over posteriors τ supported on $T = \{\mu_{12}, \mu_2, \mu_{123}\}$.

Lemma 2 (Market clearing). *If $d_{(\nu_0,\tau)}$ is feasible on $G_P(\tau)$, then the flow out of any supply node $\mu \in T$ is exactly $\tau(\mu)$ (and not less), and the flow into any demand node $a \in A$ is exactly $\nu_0(a)$ (and not more).*

Proof of Lemma 2. Suppose that $d_{(\nu_0,\tau)}$ is feasible. We show that the flow into any demand node $a \in A$ is exactly $\nu_0(a)$. Towards a contradiction, suppose that $\sum_{\mu \in T} f(\mu, a) \geq \nu_0(a)$ for all $a \in A$, with strict inequality for some a. Summing over all actions on both sides of the inequality yields

$$\sum_{a \in A} \sum_{\mu \in T} f(\mu, a) > \sum_{a \in A} \nu_0(a) = 1.$$

On the other hand, because $d_{(\nu_0,\tau)}$ is feasible, then the flow out of each $\mu \in T$ is at most $\tau(\mu)$, and therefore for all $\mu \in T$

$$\sum_{a \in A} f(\mu, a) \leq \tau(\mu).$$

Summing again over all actions on both sides yields

$$\sum_{\mu \in T} \sum_{a \in A} f(\mu, a) \leq \sum_{\mu \in T} \tau(\mu) = 1,$$

a contradiction. The proof that the flow out of any supply node μ is exactly $\tau(\mu)$ is analogous and hence omitted. □

Proof of Proposition 3. Suppose first that the Bayes plausible distribution over posteriors τ is such that $d_{(\nu_0,\tau)}$ is feasible on $G_P(\tau)$ and let f denote the feasible flow. Consider a decision rule $\alpha : \Delta(\Theta) \mapsto \Delta(A)$ such that the agent takes action $a \in A$ when the belief is $\mu \in T$ with probability $\sigma(a \mid \mu) = f(\mu, a) / \tau(\mu)$. This correctly defines a decision rule as

$$\sum_{a\in A} \alpha(a \mid \mu) = \frac{\sum_{a\in A} f(\mu, a)}{\tau(\mu)} = 1$$

where the second equality is implied by Lemma 2. Furthermore, α is optimal for the agent because μ and a are connected with an edge only if a is optimal under μ, i.e. $a \in a^*(\mu)$.

To verify that (τ, α) induce ν_0, note that for all $a \in A$

$$\sum_{\mu\in T} \tau(\mu) \alpha(a \mid \mu) = \sum_{\mu\in T} f(\mu, a) = \nu_0(a).$$

where the second equality follows again from Lemma 2. Thus, ν_0 is consistent with τ.

Conversely, suppose that (μ_0, ν_0) are BCE-consistent. Then, by Equation 5, a Bayes plausible distribution over posteriors τ and a decision rule α exists that induce an obedient experiment.[13] Define the graph $G_P(\tau)$ and the demand $d_{(\nu_0,\tau)}$. Note that the demand $d_{(\nu_0,\tau)}$ is feasible on $G_P(\tau)$ by defining the flow $f(\mu, a) = \alpha(a \mid \mu)\tau(\mu)$ for all $(\mu, a) \in T \times A$. ☐

Proposition 3 implies that verifying that τ implements ν_0 is equivalent to verifying the feasibility of the demand $d_{(\nu_0,\tau)}$ for the graph G_P. The main theorem in Gale (1957) provides necessary and sufficient conditions under which $d_{(\nu_0,\tau)}$ is feasible. Adapted to our setting, the conditions in Gale (1957) can be stated as follows:

Proposition 4 (Gale, 1957). *The demand $d_{(\nu_0,\tau)}$ is feasible on graph $G_P(\tau)$ if and only if for every set $B \subseteq A$ a flow f_B exists such that:*

1. $\sum_{a\in A} f_B(\mu, a) \leq \tau(\mu)$ *for all $\mu \in T$, and*
2. $\sum_{a\in B} \sum_{\mu\in T} f_B(\mu, a) \geq \sum_{a\in B} \nu_0(a)$.

[13] Namely, BCE-consistency implies the existence of an obedient experiment from which we can infer the following distribution over posteriors. First, let

$$\mu_a(\theta) = \frac{\mu_0(\theta)\pi(a|\theta)}{\sum_{\theta'\in\Theta} \mu_0(\theta')\pi(a|\theta')},$$

and let $\tau(\{\mu_a\}) = \sum_{\theta\in\Theta} \mu_0(\theta)\pi(a|\theta)$. The decision rule $\alpha(\cdot|\mu_a) = \mathbb{1}[a' = a]$ completes the construction.

In our setting, given a set $B \subseteq A$ items 1 and 2 in Proposition 4 are satisfied for some flow f_B if and only if they are satisfied when the out flow from every supply node that is connected to nodes in B is maximal. Denote the set of all posterior beliefs in T for which some action in B is optimal (and perhaps also actions that are not in B) by $\Delta^*(B) = \{\mu \in T \mid \exists a \in B, a \in a^*(\mu)\}$. Thus, in the graph we constructed, all and only beliefs (i.e., supply nodes) in $\Delta^*(B)$ are connected to actions (i.e., demand nodes) in B. The next corollary follows immediately:

Corollary 2. *The Bayes plausible distribution over posteriors τ implements ν_0 if and only if for every subset $B \subseteq A$,*

$$\sum_{\mu \in \Delta^*(B)} \tau(\mu) \geq \sum_{a \in B} \nu_0(a). \tag{18}$$

To see that the condition in Corollary 2 is equivalent to that in Proposition 1, note first that because τ, ν_0 are measures (and hence add up to 1), Equation 18 can be equivalently written as follows:

$$\sum_{a \in \overline{B}} \nu_0(a) \geq \sum_{\mu \in \overline{\Delta^*(B)}} \tau(\mu), \tag{19}$$

where the upper-bar notation denotes the complement of a set–for instance, $\overline{B} = A \setminus B$.

Note that

$$\overline{\Delta^*(B)} = \{\mu \in T | a^*(\mu) \cap B = \emptyset\} = \bigcup_{C \subseteq \overline{B}} \{\mu \in T | a^*(\mu) = C\}.$$

Hence, we can write Equation 19 as follows

$$\sum_{a \in \overline{B}} \nu_0(a) \geq \sum_{C \subseteq \overline{B}} \sum_{\mu \in T : a^*(\mu) = C} \tau(\mu) \tag{20}$$

which is the equation in Proposition 1.