Private Posterior Implementation in Collective Decision Problems*

Justice C. Harasha†

March 15, 2024

Abstract

Posterior implementation is a sparsely studied solution concept for mechanism design when there are interdependent agent types. In posterior equilibrium, it is required that each agent's strategy is optimal with respect to the strategies played by their fellow agents for each possible message profile. There are two main considerations of posterior implementation in the current literature. First, Green and Laffont (1987) offer a geometric characterization of posterior implementable social choice functions in two agent, binary collective decision problems. Then, Niemeyer (2022) generalizes this analysis by considering binary collective decision problems with any number n of finitely many agents, with the main insight being that posterior implementable social choice functions are posterior implementable by score voting mechanisms. In both cases, it is assumed that all messages sent by agents are publicly observable. In this paper, we examine cases where only some aspects of agent messages are observable. Namely, we consider a case where agents submit their messages to a a central agent, or collector, who then uses these reports to make a public choice. Agents, therefore, form posterior beliefs regarding the types of their fellow agents based on this public choice, not on the granular message reports of their fellow agents. This, in turn, creates coarser agent posterior beliefs. We thus define an amended notion of posterior implementation, which we denote *private posterior implementation*, and for this, obtain a complete characterization of the set of privately posterior implementable decision rules in n-person binary collective decision problems. We also consider non-binary collective decision problems, where the public choice is a parameter, such as a price vector, and discuss the challenges that arise in such settings.

JEL CLASSIFICATION: C72, D70, D71, D82.

KEYWORDS: posterior equilibrium, non-transferable utility, interdependent values, extensive form mechanisms

---

*I acknowledge support from the Yale University Department of Economics and invaluable advisement from Dirk Bergemann. My work was aided by helpful comments from Kai Hao Yang and Harry Pei. All errors are, of course, my own.

†Yale University, justice.harasha@yale.edu.

Contents

1 Introduction

The use of collective decision-making procedures are rife in all spheres of human interaction.

1.1 Related Literature

2 Preliminaries

In the forthcoming section, we formalize a standard binary collective decision problem with n agents.

2.1 The Classical Model

A group of n agents are deciding whether to accept or reject a given alternative. We index the set of agents by $i \in N = \{1, 2, .., n\}$. Each agent i has a value function, $v_i(\theta)$, where $\theta \in \Theta$ is some unknown state. The state θ has n elements $(\theta_1, \theta_2, ...\theta_n)$ and each agent i observes only θ_i, known as agent i's type. This is to say that each agent has partial information about the payoff-relevant state. We assume that θ_i is some real number drawn from a compact interval Θ_i, an interval that we normalize to $[0, 1]$, without loss of generality. This yields the following construction of the normalized state space:

$$\Theta = \prod_{i=1}^{n} \Theta_i = [0, 1]^n. \tag{1}$$

We assume that states $\theta \in \Theta$ are distributed according to a probability measure $\mu(\cdot) \in \Delta(\Theta)$ that it has a continuously differentiable and strictly positive density function f.

To make this collective binary choice, a mechanism is introduced. The mechanism of choice is one without transfers. That is, the agents use a mechanism composed of a collection of measurable message spaces M_i and a measurable allocation or outcome function $\psi : M \to [0, 1]$ which assigns an acceptance probability to each message profile $m \in M$, where we define M:

$$M = \prod_{i=1}^{n} M_i. \tag{2}$$

Taken together, a mechanism without transfers is the pair (M, ψ) and it induces a game of incomplete information in which agents attempt to maximize their expected utilities. Before proceeding with considerations of equilibria in this setting, we recount the technical assumptions that are taken as given in both Green and Laffont (1987) and Niemeyer (2022).

2.2 Technical Assumptions

In the two primary works on posterior implementability by Green and Laffont (1987) and Niemeyer (2022), the following three assumptions are employed:

(i) *Monotonicity*: For each agent i, the gradient of $v_i(\theta)$, denoted $\nabla v_i(\theta) : \Theta \to \mathbb{R}^n$ is strictly positive.

(ii) *Heterongeneity*: $v_i(\theta) = 0 \; \forall i \in N \implies \Big(\nabla v_i(\theta)\Big)_{i \in N}$ are not collinear.

(iii) *Affiliation*: $f(\theta)f(\theta') \leq f(\theta \wedge \theta')f(\theta \vee \theta')$, where $\wedge$ and $\vee$ denote the component-wise maximimum and minimum, respectively.

There are intuitive underpinnings and implications to the above assumptions that we summarize in Appendix A. In addition, at each point during the remainder of the paper where an above assumption is employed, we will make a clear note of this.

2.3 Public Posterior Equilibrium in Two-Agent Problems

As was made clear in Section 1.1, considerations of public[1]posterior implementation differ greatly as the number agents in a given setting changes from $n = 2$ to $n \geq 3$. In this section,

[1]We note that the existing notions of posterior implementation will henceforth be referred to as *public posterior implementation*, which accounts for the fact that they are solution concepts that allow for the public observation of individual agent messages. This, of course, is a facet of posterior implementation that we do away with throughout most of the paper, when we introduce our novel notion of *private posterior implementation*, where agent messages are no longer publicly observable.

we consider notions of posterior implmentation in $n = 2$ agent settings, first considered by Green and Laffont (1987). Note, the general model given above considers the case of n agents. The function of this section is to plot the development of the concept of posterior implementation, so as to better orient the introduction of private posterior implementation in later sections. For the most part, the analysis done in this paper will concern settings with $n \geq 3$ agents.

To begin, recall our construction of a mechanism without transfers from Section 2.1. Suppose we now restrict the collection of measurable message spaces M to reflect the two-player environment. That is:

$$M = M_1 \times M_2, \tag{3}$$

where M_1 and M_2 correspond to the message spaces for player 1 and player 2, respectively. We can therefore explicitly construct the two-player mechanism to be the the following triple (M_1, M_2, ψ), where ψ is the mapping $\psi : (M_1 \times M_2) \to [0, 1]$, and, as above, yields an acceptance probability for each message profile $m \in (M_1 \times M_2)$.

From here, we also formalize the notion of a strategy in this environment. We say that a *strategy* of agent i is a measurable function from Θ_i to the family of distributions over M_i. We express this idea as the conditional distributions $s_i(m_i|\theta_i)$. Using this understanding of a strategy in the observable-message environment, we give the recount Green and Laffont's notion of a social choice function, which allows us to then begin considering their formalized definitions of posterior optimality and implementation:

Definition 1 *(Green and Laffont, 1987) Suppose s_1 and s_2 are stragtegies used by agents in the mechanism without transfers given by (M_1, M_2, ψ). Then, the social choice function $\phi : (\Theta_1 \times \Theta_2) \to [0, 1]$ is given explicitly as:*

$$\phi(\theta_1, \theta_2) = \int_{M_1 \times M_2} \psi(m_1, m_2) ds_1(m_1|\theta_1) ds_2(m_2, \theta_2), \tag{4}$$

which yields the acceptance probability of a given alternative when (θ_1, θ_2) is the information *received by the agents.*

From this, we can begin a gradual construction of Green and Laffont's definition of posterior equilibrium and implementation, followed by the main characterization theorem for the two-agent setting. We proceed with an general notion of optimality:

Definition 2 *(Green and Laffont, 1987) Suppose the given strategy s_2 is fixed. We say that the message, or report, $m_1 \in M_1$ is optimal for Player 1 if:*

$$m_1 \in \arg\max \int_{\Theta_2} \int_{M_2} v_1(\theta_1, \theta_2)\psi(m_1, m_2) ds_2(m_2|\theta_2)\mu(d\theta_2|\theta_1) \tag{5}$$

Then, a strategy for Player 1 is optimal if for almost everywhere (a.e.) $\theta_1 \in \Theta_1$, Player 1's strategy $s_1(\cdot|\theta_1)$ assigns a probablility of zero to the set of non-optimal messages.

Using the above notion of optimal messages and strategies, we say that for a given strategy s_1, if s_2 is an optimal strategy for Player 2 and s_1 is an optimal strategy for Player 1, then (s_1, s_2) constitutes a *Bayesian equilibrium* of the mechanism (M_1, M_2, ψ). Further, we say that the social choice function ϕ is *Bayesian incentive compatible*. Finally, when there exists a mechanism (M_1, M_2, ψ) and a Bayesian equilibrium (s_1, s_2), we say that that the social choice function ϕ is *implementable*. We refine these notions of optimality and implementation to arrive at Green and Laffont's definition of posterior optimality.

Definition 3 *(Green and Laffont, 1987) Suppose ϕ is implemented via the mechanism (M_1, M_2, ψ) in Bayesian equilibrium (s_1, s_2). Let $\mu(\theta_2|m_2, \theta_1)$ and $\mu(\theta_1|m_1, \theta_2)$ denote the conditional distributions that the two players hold about each other's types after observing the other's choice of m_i. Then, we say that the pair of strategies (s_1, s_2) is posterior optimal if:*

$$m_1 \in \arg\max \int_{\Theta_2} v_1(\theta_1, \theta_2)\psi(m_1', m_2)\mu(d\theta_2|m_2, \theta_1) \tag{6}$$

$$m_2 \in \arg\max \int_{\Theta_1} v_2(\theta_1, \theta_2)\psi(m_1, m_2')\mu(d\theta_1|m_1, \theta_2) \tag{7}$$

over $m_1' \in M_1$ and $m_2' \in M_2$, respectively.

We further say that the social choice function ϕ is *posterior implementable* if it is implementable via the mechanism (M_1, M_2, ψ) with posterior optimal strategies (s_1, s_2).

This, taken with the previous collection of definitions, characterizes Green and Laffont's two-agent construction. From here, we recall the central characterization theorem from their work, which amounts to a constructive geometric analysis of the two-agent environment. The result is given below, without proof:

Theorem 1 *(Green and Laffont, 1987) Any posterior-implementable social choice function ϕ is such that there exists a step function ξ and $\phi(\theta_1, \theta_2) = \phi_+$ if (θ_1, θ_2) lies above ξ and $\phi(\theta_1, \theta_2) = \phi_-$ if (θ_1, θ_2) lies below ξ, where ϕ_+ and ϕ_- are two constant values of ϕ.*

Simply, Green and Laffont's geometric characterization of posterior implementability in $n = 2$ agent binary collective decision problems is as follows: Any posterior implementable social choice function is such that there exists a decreasing step function that partitions the type space into two distinct regions, on each of which the social choice is a constant value. We offer a graphical example of this in the figure below:

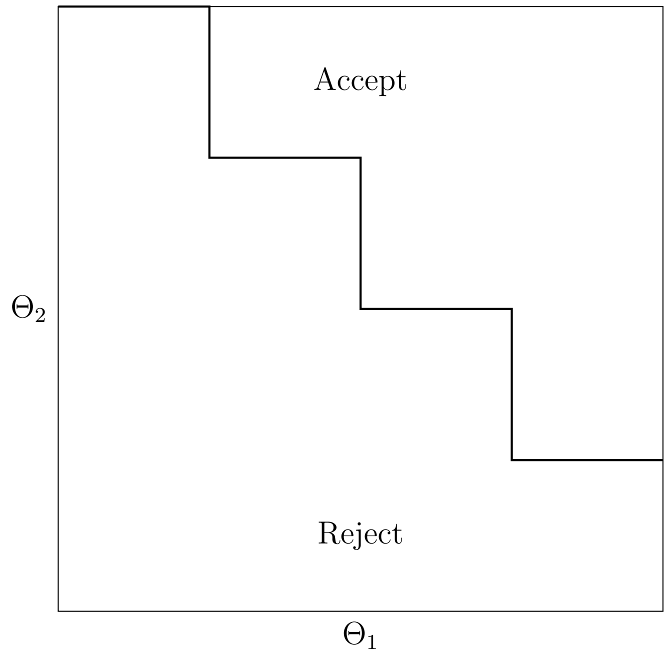

Figure 1: An illustration of Green and Laffont's central geometric result. The step function partitions the type space into two regions, on which the social choice function is constant. This graphically represents a posterior-implementable social choice function.

This concludes our consideration of the two-agent agent environment, offered by Green and Laffont (1987). Their analysis is one that relies heavily upon graphical intuition and geometric arguments that are unique to this refined two-player setting. We will proceed, now, to recount the general n agent case, which will complete our work in grounding the reader in the more classical notions of public posterior implementation, before beginning our discussion of *private posterior implementation.*

2.4 Public Posterior Equilibrium in n Agent Problems

We now return to the more general setting described by the model in Section 2.1. Instead of confining the number of agents to strictly two, we consider the case for any finite number n agents. Niemeyer (2022) generalizes the work of Green and Laffont to the n agent setting, completing the characterization of classical (public) posterior implementation in binary collective decision problems.

In this section, we will revisit a collection of definitions that were offered in Section 2.4 in a manner that may appear redundant, but it is only in an effort to make clear the difference in approaches between the two-player and n-agent settings. It is important to note that, in our introduction and analysis of private posterior implementation, we use an analytical framework most similar to the Niemeyer construction we are about to revisit. These two separate constructions, the Green-Laffont and the Niemeyer frameworks, serve to show how unique the geometric characterization of the two-agent case is, relative to the analysis of the n agent case. We proceed with Niemeyer's construction below.

Once more, we revisit the notion of a strategy in this more general environment. A *strategy* for agent i in the mechanism (M, ψ), where $M = \prod_{i=1}^{n} M_i$, is the map[2] $\sigma_i : \Theta_i \to \Delta(M_i)$. Next, let $\mu(\cdot \mid \theta_i) \in \Delta(\Theta_{-i})$ be the belief that some agent i holds about the types of other players when her own type is θ_i. Then, when agent i observes the strategies of her fellow agents σ_{-i} and thus their messages $m_{-i} \in M_{-i}$, she holds a posterior belief $\mu(\cdot \mid \theta_i, m_{-i}) \in \Delta(\Theta_{-i})$. It is important to note that her beliefs about other agents' types are conditioned on her own type *and* the messages submitted by her fellow agents. We also assume that an agent's posterior beliefs as constructed above are derived via Bayes' rule in almost every case.

From this, we get a definition that aids in our reconstruction of Niemeyer's n-person notion of posterior equilibrium:

Definition 4 *The posterior expected valuation of agent i is given as:*

$$V_i(\theta_i | m_{-i}) = \int_{\Theta_{-i}} v_i(\theta_i, \theta_{-i}) \mu(d\theta_{-i} | \theta_i, m_{-i}) \tag{8}$$

conditioned on m_{-i} and when θ_i is her type.

From this, we can offer Niemeyer's notion of (public) posterior equilibrium in n-person

[2] In the spirit of technicality, a strategy σ is a Markov kernel, which we define in the mathematical portion of the appendix.

collective decision probelms:

Definition 5 *A strategy profile $\sigma = (\sigma)_{i \in N}$ in the mechanism (M, ψ), where M is a measurable message space and ψ is a measurable outcome function $\psi : M \to [0,1]$, is a public posterior equilibrium if for all $\imath \in N$, $\theta_i \in \Theta_i$, $m_{-i} \in M_{-i}$ and $\tilde{m}_i \in M_i$*

$$V_i(\theta_i|m_{-i})\psi(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})) \geq V_i(\theta_i|m_{-i})\psi(\tilde{m}_i, \sigma_{-i}(\theta_{-i})).$$

Simply, a strategy profile σ is in a posterior equilibrium if each σ_i is optimal against the strategies σ_{-i} of other agents for every possible message profile m_{-i}.

Using this grounding in Niemeyer's conception of posterior optimality, we begin building towards the central characterization theorem of the n-person environment. To do this, we begin by recalling the general notion of score voting mechanisms:

Definition 6 *A mechanism (M, ψ) is a score voting mechanism if*

1. *for each $i \in N$, $M_i = \{1, ..., |M_i|\}$ is a set of consecutive integers*

2. *there is some quota $q \in \mathbb{Z}$ and real numbers $0 \leq r < a \leq 1$ such that*

$$\psi(m) = \begin{cases} a & \text{if } \sum_{i=1}^{n} m_i > q \\ r & \text{otherwise} \end{cases} \tag{9}$$

3. *each agent $i \in N$ has at most one veto message $\bar{m}_i \in M_i$ such that $\psi(\bar{m}_i) = a$ and at most one veto message $\underline{m}_i \in M_i$ such that $\psi(\underline{m}_i) = r$.*

There exist a few notable classes of score voting mechanisms, namely i-dictatorship, unanmity for acceptance or rejection, simple majority, and sub-or supermajority mechanisms. We offer a formal definition of these, and others, in Appendix A.

Using this notion of score voting mechanisms, we can now proceed to offer the central characterization theorem of the n agent environment, from Niemeyer:

Theorem 2 *(Niemeyer, 2022) Let $n \geq 3$. A responsive*[3]*social choice function is posterior implementable if and only if it is posterior implementable by score voting in pure surjective*[4] *strategies.*

We offer a reconstruction of this proof, as well as valuable intuition regarding the equivalence between score voting implementation and posterior implementation in Appendix B.

This central result from Niemeyer's characterization of the n-person environment provides the basis upon which we begin our novel framework of *private posterior implementation*.

3 Introducing Private Posterior Implementation

In the forthcoming section, we introduce our novel concept of *private posterior implementation*. To begin with, the amended setting, including the notion of a *central agent* or *collector*, is formalized. From this new model, we introduce the formal definition of private posterior equilibrium, and state our first result that is confined to the two agent setting.

3.1 The Central Agent Model

Once more, a group of n agents are deciding whether to accept or reject a given alternative. The formalities, notation, and assumptions from the public message model are maintained in this setting. See Section 2.1 for explicit delineations of these things.

Now, suppose we introduce an additional agent or piece of collecting technology (either of which, clearly, do not participate in the game of incomplete information induced by the mechanism) to which each of the participating agents privately submit their messages. This means that each agent no longer observes the reports or messages of their fellow agents, as they all simply make their reports privately and directly to this *central agent.* We consider this addition concretely, given in the timing of this amended setting:

[3]A social choice function ϕ is *responsive* if it can only be implemented by giving each agent at least two messages.

[4]A strategy profile $\sigma : \Theta \to M$ is said to be *surjective* if for every message profile $m \in M$ there exists a strategy profile $\theta \in \Theta$ such that $\sigma(\theta) = m$.

1. The designer commits to the mechanism (M, ψ).

2. Each agent i submits their message m_i to the central agent. The central agent and the central agent alone observes these messages.

3. The central agent uses these message submissions to select a public choice (i.e. to accept or reject the given alternative).

From this construction, agents learn information and draw inferences about the types of their fellow agents based on the central agent's selection, not on the individual messages submitted. For this reason, we classify posterior implementation in this setting as *private posterior implementation.* As a result of not directly observing the messages, agents form less refined posterior distributions about the types of their fellow agents in this environment. We formalize the concept of private posterior implementation and discuss these coarser posterior distribution next.

3.2 On the Concept of Private Posterior Implementation

Prior to defining private posterior implementation, we begin by discussing how to specify the posterior beliefs of agents after observing the equilibrium behavior of their fellow agents.

As above, we maintain that $\mu(\cdot | \theta_i) \in \Delta(\Theta_{-i})$ denotes the belief some given agent i holds about the types of other players when her own type is θ_i. Then, recall that we define strategies to be the map $\sigma_i : \Theta_i \to M_i$. In the public message environment, an agent observes these strategies and thus messages of her fellow agents, forming posterior beliefs $\mu(\cdot \mid \theta_i, m_{-i})$, about the types of her fellow agents, conditioned on her own type and the messages reported by her fellow agents. With the introduction of the central agent and private messages, agents lose the ability to condition their posterior beliefs about the types of their fellow agents on messages. Instead, the central agent (henceforth denoted CA) privately observes each agent's individual message m_i and uses them to make a public choice, which we denote $CA(m)$. Therefore, in this novel private message environment, agents condition their posterior beliefs about their fellow agents' types on their own type and the public choice $CA(m)$ made by the central agent, which is given explicitly as $\mu\Big(\cdot \,\Big|\, \theta_i \; CA(m)\Big)$. We offer it explicitly in definitional form for ease of reference below:

Definition 7 *The private posterior belief that an agent i holds about the types of her fellow agents is given by:*

$$\mu\Big(\cdot \,\Big|\, \theta_i,\; CA(m)\Big), \tag{10}$$

where μ is the probability measure which θ_i is distributed according to, θ_i is agent i's own type, and $CA(m)$ is the public choice made by the central agent after observing the private messages of all agents.[5]

From this, we also get an amended notion of posterior expected valuation, given below:

Definition 8 *Let:*

$$V_i\Big(\theta_i \mid CA(m)\Big) = \int_{\Theta_{-i}} v_i(\theta_i, \theta_{-i})\mu\Big(d\theta_{-i}\,\Big|\, \theta_i,\; CA(m)\Big) \tag{11}$$

define the private posterior expected valuation of an agent i, where $v_i(\cdot)$ denotes the value function of a given agent i.

We can thus define the notion of public posterior equilibrium using Definition 8:

Definition 9 *A strategy profile $\sigma = (\sigma)_{i \in N}$ in the mechanism (M, ψ), where M is a measurable message space and ψ is a measurable outcome function $\psi : M \to [0, 1]$, is a private posterior equilibrium if for all $i \in N$, $\theta_i \in \Theta_i$, $m_{-i} \in M_{-i}$ and $\tilde{m}_i \in M_i$*

$$V_i\Big(\theta_i \mid CA(m)\Big)\psi(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})) \geq V_i\Big(\theta_i \mid CA(m)\Big)\psi(\tilde{m}_i, \sigma_{-i}(\theta_{-i})).$$

In contrast to Green and Laffont and Niemyer's classical (public) posterior equilibrium which requires that the strategy σ_i of each agent i is optimal against the strategies σ_{-i} of other

[5] In keeping with the n-person public message environment from Niemeyer (2022), we assume private posterior beliefs are derived via Bayes' rule when it is possible to do so.

agents for every possible message profile m_{-i}, the above notion of public posterior equilibrium only requires that the strategy σ_i of each agent i is optimal against the strategies σ_{-i} of other agents for the public choice $CA(m)$ made by the central agent. In our general characterization of this notion of private posterior equilibrium in Section 5, it will become apparent that it can be viewed as the most extreme, or most private version of an equilibrium condition that is the result of a *coarsening* procedure, where the message space becomes less and less publicly observable by the introduction of lotteries over available information. See Section 5.1.

Further, we let the map given by $\phi : \Theta \to [0,1]$ be called the *social choice function*, which assigns an acceptance probability to every state $\theta \in \Theta$. Taken together with the mechanism (M, ψ), the collection (M, ψ, σ) is called an *implementation* when $\psi \circ \sigma = \sigma$ almost everywhere. Most particularly, when σ constitutes a public posterior equilibrium, we call (M, ψ, σ) a *private posterior implementation*, and ϕ *private posterior implementable.*

Remark 1 *As is true in both of the public settings (two agents and n agents), the revelation principle, in a standard sense, does not hold in cases of private posterior implementation.*

The above analysis is the result of the following argument, and was first made by Niemeyer. It, however, also holds in the case of public posterior implementability. Suppose some private posterior implementable social choice function is private posterior implementable via a direct revelation mechanism where truth-telling is optimal. This sort of mechanism in the given setting perfectly reveals the types of each agent to every other agent, such that truth-telling is a private posterior equilibrium if and only if truth-telling is an ex-post equilibrium. From Feng, Niemeyer, and Wu (2022), we know that ex-post implementable social choice functions are constant, which completes the restructuring of the argument for why the revelation principle does not generally hold in private posterior implementation in the given setting.

4 Coarse Information Structures

We proceed with our general characterization of private posterior implementation. To do this, we consider a process of *informational coarsening*, which allows us to use the classical notion of public posterior implementation as our starting point and gradually show a progression towards private implementation. That is, we interpret public (classical) posterior implementation and private posterior implementation as the relative extremes on a spectrum of informational publicity. On the one hand, the classical notion of public posterior implementation can be characterized by its *perfectly public* message space (i.e. each granular and individual message is publicly observable by each agent), while private posterior implementation is characterized by its *perfectly private* message space (i.e. individual messages are not observable to any agents, as is the case in the central agent environment). The region between these two notions of posterior implementation can be characterized by *coarse information structures*, which can be interpreted as lotteries over available information that are included in the mechanism. We formalize these notions in the forthcoming subsections, but consider the diagram below as a helpful overview of this:

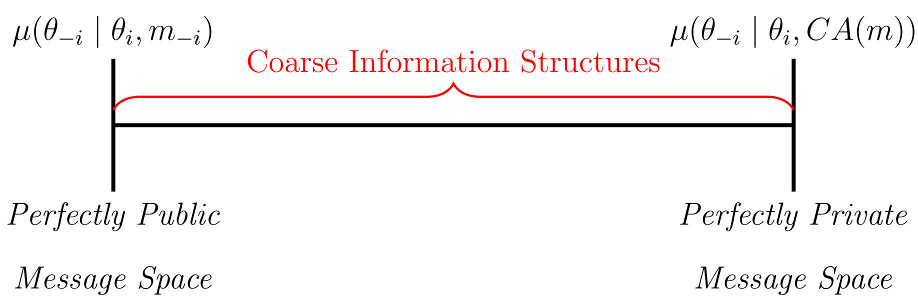

Figure 2: A depiction of private versus public posterior implementation, with the coarse information structures characterizing the region between the two.

From Figure 2, we get a preview of the arguments and characterizations that we will make in the forthcoming sections, namely how we will introduce informational coarseness to the perfectly public message space through lotteries over available information, and inevitably, through this coarsening procedure, arrive at a characterization of private posterior implementation.

4.1 Formalization

We begin by formalizing some of the concepts mentioned above, particularly the notion of public posterior implementation and private posterior implementation existing as the extrema on a spectrum of informational publicity. We characterize them as extrema using the terms *perfectly public* and *perfectly private*, respectively. We offer their formal definitions below, beginning with the notion of perfectly public:

Definition 10 *A conditional distribution of posterior beliefs $\mu(\cdot \mid \theta_i, m_{-i}) \in \Delta(\Theta_{-i})$ is called perfectly public if each posterior belief can be conditioned on all reported messages. That is, the entirety of the message space is publicly observable to each agent.*

Next, we consider the notion of perfectly private:

Definition 11 *A conditional distribution of posterior beliefs $\mu(\cdot \mid \theta_i, CA(m)) \in \Delta(\Theta_{-i})$ is called perfectly private if no aspect of the message space is publicly observable. That is, the only available information upon which agents can condition their posterior beliefs are their own types and the public choice.*

We have thus formalized the concept of a perfectly public versus a perfectly private environment. Simply, the perfectly public environment corresponds to the Niemeyer setting, while the perfectly private environment corresponds to the central agent setting discussed in Section 3.1. Inevitably, we obtain a general characterization for private posterior implementable social choice functions, i.e. we obtain a general characterization of the environment with a perfectly private message space. However, in order to begin our characterization of the perfectly private case, we proceed by considering a series of interim settings, namely when the message spaces are partially observable in a public sense. To do this, we introduce the concept of *informational coarseness*, which presents a means by which we can randomize the subset of the message space that is publicly observable, and thus engage in an analysis of the interim dynamics of posterior implementation. This interim analysis using coarse information structures founds the basis upon which we consider private posterior implementation:

Definition 12 *A coarse information structure consists of a collection of measurable message spaces* $M = \prod_{i=1}^{n} M_i$ *and a measurable function* $\varphi : M \to [0,1]^n$*, which maps the collection of message spaces to a publicity probability, which corresponds to the probability that a given message space is publicly observable. Together, the pair* (M, φ) *are called a coarse information structure.*

From this, a natural course of action to begin our general characterization is to define two fairly trivial coarse information structures, namely by ascribing a sort of deterministic coarse information structure to both the perfectly public and perfectly private message space environments. Not only does this provide a constructive baseline upon which we will begin making more general arguments regarding private implementation, it provides an opportunity to explicitly examine how coarse information structures function and operate within a given mechanism.

Consider, first, the case of a perfectly public message space environment with n agents, the Niemeyer construction, where every message submitted is publicly observable to each agent. The mechanism (M, ψ) is used to make the collective binary choice and the coarse information structure (M, φ) describes the publicity probability of each message profile, where φ is the measurable function given by the mapping $\varphi : M \to [0,1]^n$. Suppose m denotes an arbitrary message profile. Explicitly, in the case of a perfectly public message space, the coarse information structure is described by the following result:

$$M = \prod_{i=1}^{n} M_i \qquad \varphi(m) = \begin{bmatrix} 1.0 \\ 1.0 \\ \vdots \\ 1.0 \end{bmatrix}, \tag{12}$$

where the output of the mapping given by φ holds for each $m \in M$. Simply, in a perfectly public message space environment, the coarse information structure is characterized by a $\varphi(\cdot)$ function that maps each message profile to a publicity probability equal to 1, which is to say that each message profile is always and entirely public and observable to every agent. As the reader can infer, a correspondingly trivial coarse information structure can describe the perfectly private message space environment i.e. the *central agent model*, which we give below:

$$M = \prod_{i=1}^{n} M_i \qquad \varphi(m) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{13}$$

where, once more, the $\varphi(\cdot)$ mapping holds for each possible message profile $m \in M$. Here, (13) describes the perfectly private nature of the central agent mode, where each message profile is entirely and always unobservable to every agent, which is quantified by the zero probability that is assigned to every possible message profile.

Taken together, we have offered an explicit characterization of either ends of the publicity spectrum offered in Figure 2. They are, as noted, the trivial cases of informational coarseness–the binary cases of when a message space is either entirely observable or entirely not. In the interim cases of informational coarseness that exist between these two extrema, the entries in the output column vector of the $\varphi(\cdot)$ mapping are probabilities $p_1, p_2, ..., p_n \in (0, 1)$ that describe the interim probabilities of a given message profile being publicly observable to each agent.

We are now equipped with a formal construction that allows us to deviate from the perfectly public nature of Niemeyer's environment, and analyze how characterizations of posterior implementation evolve as we privatize portions of the message space through probabilistically limiting the publicity of message profiles to each agent. In order to begin this analysis, we first consider informational coarseness in the context of a more familiar, but closely related solution concept: ex post implementation. In fact, a recent result on the limitations of ex post implementation is crucial in Niemeyer's proof of his score voting characterization of n-person public posterior implementation. Thus, our analysis of informational coarseness as it relates to ex post implementation makes a path towards a characterization of private posterior implementation much clearer.

5 On Ex Post Implementation

We proceed by recounting crucial work on the impossibility of ex post implementation by Jehiel et al. (2006) and Feng, Niemeyer, and Wu (2022) in collective decision-making environments.

6 Appendix

The forthcoming appendix includes a collection of omitted proofs and technical remarks from the main text, as well as a series of examples, with visualizations of implementation in two-agent settings.

6.1

6.2 On the Derivation of Private Posterior Beliefs

6.3 Notes on Ex-Post Impossibility Results

6.4 Example

To aid in the development of intuition, we revert our attention back to the two-agent setting, and consider a particular example. In this way, we familiarize the reader with the practical workings of our novel notion of private posterior equilibrium, and perhaps more importantly, how it is situated relative to other more familiar solution concepts, like Bayesian, ex-post, and public posterior equilibrium. This provides a valuable sort of intuition prior to our discussion of results specific to private posterior implementation.

6.4.1 The Setting

Suppose that $n = 2$ agents are deciding whether to accept or reject some given alternative. Assume that $\forall i \in N$, the type space is defined according to $\Theta_i = [-1, 1]$ and the individual types θ_i are uniformly distributed along the same interval: $\theta_i \sim U[-1, 1]$. Further, the valuations for each agent i are given by:

$$v_i(\theta_i, \theta_j) = \alpha\theta_i + (1 - \alpha)\theta_j \tag{14}$$

where $\alpha \in (\frac{1}{2}, 1)$. We note that agent i can observe θ_i but not θ_j, meaning that each agent only partial information about the realized payoff-relevant state. We can thus interpret α as the parameter that captures the degree to which a given agent's valuation depends on the private information of his fellow agent(s). We call environments of the form given in (12) an *interdependent value* environment.

Then, we can define the joint welfare $w(\theta_i, \theta_j)$ in the two agent setting as follows:

$$w(\theta_i, \theta_j) = v_i(\theta_i, \theta_j) + v_i(\theta_i, \theta_j) \tag{15}$$

$$= [\alpha\theta_i + (1-\alpha)\theta_j] + [\alpha\theta_j + (1-\alpha)\theta_i] \tag{16}$$

$$= \theta_i + \theta_j. \tag{17}$$

Under a socially optimal outcome, we necessarily have that:

$$w(\theta_i, \theta_j) = \theta_i + \theta_j \geq 0. \tag{18}$$

We thus characterize the socially optimal region as follows:

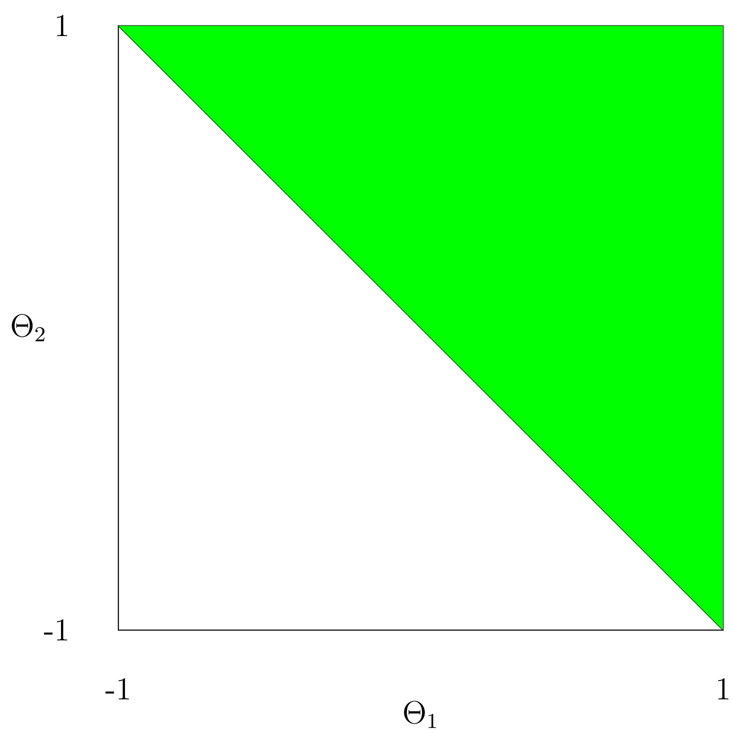

Figure 3: The socially optimal acceptance region in a two-player binary collective decision problem, where the region is given by the inequality $w(\theta_{i,j}) \geq 0$.

In addition, in each of the following iterations of this example, we assume that the designer commits to the following mechanism without transfers: (M, ψ) where $M_i = \{0, 1\}$ for each $i \in N$ and $\psi : M \to [0, 1]$ is given as:

$$\psi(m) = \begin{cases} a & \text{if } \sum_{i=1}^{n} m_i > 1 \\ r & \text{otherwise} \end{cases} \tag{19}$$

where $0 \leq r < a \leq 1$. This sort of mechanism is a type of score voting mechanism known as *unanimity for acceptance (rejection).*

In the forthcoming subsections, we illustrate how our novel notion of private posterior implementation relates to Bayesian, ex-post, and public posterior equilibrium in this two-agent case. It provides a valuable practical lens to bear in mind as we begin our general characterization of the private message (central agent) environment.

6.4.2 Bayes-Nash Equilibrium

We begin by introducing a formal definition:

Definition 13 *A strategy profile $\sigma = (\sigma)_{i \in N}$ in the mechanism (M, ψ), where M is a measurable message space and ψ is a measurable outcome function $\psi : M \to [0, 1]$, is a Bayes-Nash equilibrium if for all $i \in N$, $\theta \in \Theta$, and $\tilde{m}_i \in M_i$, we have*

$$\int_{\Theta_{-i}} v_i(\theta_i, \theta_{-i}) \psi(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})) \mu(d\theta_{-i}|\theta_i) \geq \int_{\Theta_{-i}} v_i(\theta_i, \theta_{-i}) \psi(\tilde{m}_i, \sigma_{-i}(\theta_{-i})) \mu(d\theta_{-i}|\theta_i) \tag{20}$$

where μ is a probability measure that distributes states $\theta \in \Theta$.

As this section is primarily an exercise in computation, we use the above Bayesian equilibrium conditions to construct the following maximization problem:

$$\sup_{\sigma_1(\theta_1), \sigma_2(\theta_2)} \mathbb{E}\Big[w(\theta_1, \theta_2)\Big] \tag{21}$$

subject to:

$$\begin{aligned}&\int_{\Theta_2}[\alpha\theta_1+(1-\alpha)\theta_2]\psi(\sigma_1(\theta_1),\sigma_2(\theta_2))\mu(d\theta_2|\theta_1)\\ &\geq\int_{\Theta_2}[\alpha\theta_1+(1-\alpha)\theta_2]\psi(\tilde{m}_1,\sigma_2(\theta_2))\mu(d\theta_2|\theta_1)\end{aligned}\tag{22}$$

$$\begin{aligned}&\int_{\Theta_1}[\alpha\theta_2+(1-\alpha)\theta_1]\psi(\sigma_1(\theta_1),\sigma_2(\theta_2))\mu(d\theta_1|\theta_2)\\ &\geq\int_{\Theta_1}[\alpha\theta_2+(1-\alpha)\theta_1]\psi(\sigma_1(\theta_1),\tilde{m}_2)\mu(d\theta_1|\theta_2),\end{aligned}\tag{23}$$

where (20) and (21) correspond to each agent's Bayesian incentive compatibility constraint. In this way, (19)-(21) characterize the Bayes-Nash equilibrium optimization problem in this two-agent example.

6.4.3 Ex-Post Equilibrium

We next consider the familiar solution concept of *ex-post implementation*, which is formally given as follows:

Definition 14 *A strategy profile $\sigma=(\sigma)_{i\in N}$ in the mechanism (M,ψ), where M is a measurable message space and ψ is a measurable outcome function $\psi: M\to[0,1]$, is an ex post equilibrium if for all $i\in N$, $\theta\in\Theta$, $\tilde{m}_i\in M_i$, we have*

$$v_i(\theta_i,\theta_{-i}(\theta_{-i}))\psi(\sigma_i(\theta_i),\sigma_{-i}(\theta_{-i}))\geq v_i(\theta_i,\theta_{-i}(\theta_{-i}))\psi(\tilde{m}_i,\sigma{-}i(\theta_{-i}))$$

In keeping with the previous subsection, we compute the ex-post equilibrium by solving a constrained optimization problem, given below:

$$\sup_{\sigma_1(\theta_1),\sigma_2(\theta_2)}\ \mathbb{E}\Big[w(\theta_1,\theta_2)\Big]\tag{24}$$

subject to:

$$[\alpha\theta_1+(1-\alpha)\theta_2]\psi(\sigma_1(\theta_1),\sigma_2(\theta_2))\geq[\alpha\theta_1+(1-\alpha)\theta_2]\psi(\tilde{m}_1,\sigma2(\theta_2))\tag{25}$$

$$[\alpha\theta_2 + (1-\alpha)\theta_1]\psi(\sigma_1(\theta_1), \sigma_2(\theta_2)) \geq [\alpha\theta_2 + (1-\alpha)\theta_1]\psi(\tilde{m}_1, \sigma 2(\theta_2)) \quad (26)$$

6.4.4 Classical (Public) Posterior Equilibrium

Next, we observe how the classical notion of posterior implementation is situated with respect to the two previous solution concepts. First, we define the following:

Definition 15 *Let*

$$V_i(\theta_i|m_{-i}) = \int_{\Theta_{-i}} v_i(\theta_i, \theta_{-i})\mu(d\theta_{-i}|\theta_i, m_{-i})$$

denote the posterior expected valuation of agent i given messages m_{-i} when she is of type θ_i.

From this, we get the formal notion of classical posterior equilibrium:

Explicitly in the specific environment described above, the posterior optimality condition is written as:

$$\int_0^1 [\alpha\theta_i + (1-\alpha)\theta_j]\mu(d\theta_{-i}|\theta_i, m_{-i})\psi(\sigma_i(\theta_i), \sigma{-i}(\theta_{-i})) \geq \int_0^1 [\alpha\theta_i + (1-\alpha)\theta_j]\mu(d\theta_{-i}|\theta_i, m_{-i})\psi(\tilde{m}_i, \sigma{-i}(\theta_{-i})) \quad (27)$$

$$\implies \int_0^1 [\alpha\theta_i + (1-\alpha)\theta_j]\mu(d\theta_{-i}|\theta_i, m_{-i})\Big[\psi(\sigma_i(\theta_i), \sigma{-i}(\theta_{-i})) - \psi(\tilde{m}_i, \sigma{-i}(\theta_{-i}))\Big] \geq 0 \quad (28)$$

Then, from Niemeyer's derivation of posterior beliefs, we have the general approach, by taking message subsets of positive measure $M'_{-i} \subset M_i$ and computing posterior beliefs as follows:

$$V_i(\theta_i|M'_{-i}) = \int_{M'_{-i}} V_i(\theta_i|m_{-i})\lambda(dm_{-i}|_i, M'_{-i}), \quad (29)$$

where we define $\lambda(\cdot)$ as the distribution over $\Delta(M_i)$ induced by a given strategy profile $\sigma_i(\theta)$. Explicitly, let $\tilde{M} \subset M$ and define as:

$$\lambda(\tilde{M}) = \int_{\Theta} \sigma(\theta)[\tilde{M}]\mu(d\theta). \quad (30)$$

We return to our particular two player environment, described concisely by the posterior optimality condition in (14). Using this derivation of posterior beliefs, we get:

$$\int_0^1 \left[\int_0^1 [\alpha\theta_i + (1-\alpha)\theta_j]\right] \int_0^1 \sigma(\theta)\mu(d\theta) \left[\psi(\sigma_i(\theta_i), \sigma{-}i(\theta_{-i})) - \psi(\tilde{m}_i, \sigma{-}i(\theta_{-i}))\right] \geq 0 \quad (31)$$

$$\implies \frac{1}{2}\left[\psi(\sigma_i(\theta_i), \sigma{-}i(\theta_{-i})) - \psi(\tilde{m}_i, \sigma{-}i(\theta_{-i}))\right] \geq 0, \quad (32)$$

which is simply a constant. This can be visaulized as a horizontal line through the origin in the above diagrams. This fits neatly between the small acceptance region given by ex-post implementation and the larger acceptance region given by Bayesian implementation, as anticipated.

6.4.5 Private Posterior Equilibrium