Implementation with Gradually Informed Agents

Dario Gori[*] and Kourosh Khounsari[†]

September 14, 2026

Abstract

This paper studies the implementation of a social choice function (SCF) when a single agent learns about his type through a designed sequence of private signals and reports. Early reports commit the agent to a continuation menu before further learning occurs, allowing dynamic implementation of SCFs that cannot be implemented statically. When disclosure can depend on both past signals and reports, we recursively characterize priors for which the SCF is implementable, and obtain an equivalent linear programming characterization. Static implementation suffices with two states or two outcomes. For injective social choice functions and strict preferences, a recursive condition on the envy relation characterizes whether implementation is possible for some compatible cardinal utilities. We also give a sufficient condition for implementation when disclosure depends only on reports.

1 Introduction

In many strategic settings, agents frequently update their private information as they interact with a designer. A customer learns about his preferences while exploring

[*]dario.gori@tse-fr.eu

[†]kourosh.khounsari@tse-fr.eu

This project received funding from the European Research Council (ERC, HWS, grant n° 101098319) and the French National Research Agency under the Investments for the Future (Investissements d'Avenir) program (grant ANR-17-EURE-0010).

products, and the seller can influence what he learns and when. Moreover, the seller may have the power to elicit early signals before the customer knows enough to make a completely informed decision. How does the interplay between designing the learning process and dynamically extracting information change the social choice functions that can be implemented?

We study a single agent who is ex-ante uninformed about their fixed payoff type, but privately learns it over the course of the interaction. The designer commits to a sequence of information disclosures and requests a report after each disclosure. The final allocation depends only on these reports. We ask which social choice functions can be implemented by truthful reporting and how a disclosure process can be constructed. The following example illustrates why alternating reports with learning can achieve an allocation that is impossible under static disclosure.

Example 1 *A real estate firm offers three houses: a small house x_1, a medium house x_2, and a large house x_3. A customer is initially uncertain about his preferences. His type is $\theta \in \{1, 2, 3\}$, with probabilities $\mu_1, \mu_2, \mu_3 > 0$. Preferences satisfy*

$$\begin{aligned} &u(x_1; 1) > \max\{u(x_2; 1), u(x_3; 1)\}, \\ &u(x_2; 2) > \max\{u(x_1; 2), u(x_3; 2)\}, \\ &u(x_2; 3) > u(x_3; 3) > u(x_1; 3). \end{aligned}$$

The firm wants to implement $f(i) = x_i$. The customer has committed to purchasing one of the three houses, and the mechanism determines which one. Assume

$$\mu_1\big[u(x_1; 1) - u(x_2; 1)\big] \geq \mu_3\big[u(x_2; 3) - u(x_3; 3)\big]. \tag{1}$$

This SCF cannot be implemented statically. Because the three types must receive different houses, a one-round implementing signal must reveal the customer's type. Type 3 *would then report so as to obtain x_2 instead of x_3.*

Consider the following two-round procedure. An initial inspection reveals whether the customer is type 2 *or belongs to $\{1, 3\}$, without distinguishing types* 1 *and* 3*. The customer reports which signal he received. A report of $\{2\}$ assigns house x_2. A report of $\{1, 3\}$ leads to a further inspection that reveals his type, after which he chooses between x_1 and x_3.*

After signal $\{2\}$*, truthful reporting gives the customer's preferred house. After signal* $\{1, 3\}$*, truthful reporting yields expected utility*

$$\frac{\mu_1 u(x_1; 1) + \mu_3 u(x_3; 3)}{\mu_1 + \mu_3}.$$

Reporting $\{2\}$ *instead assigns* x_2 *in both possible states, giving expected utility*

$$\frac{\mu_1 u(x_2; 1) + \mu_3 u(x_2; 3)}{\mu_1 + \mu_3}.$$

Inequality (1) *makes truthful reporting optimal. In the second round, type* 1 *chooses* x_1 *and type* 3 *chooses* x_3*. Thus, the procedure implements* f*. The first report commits the customer to a smaller menu before further information arrives.*

In Example 1, inspections provide private signals about the customer's preferences. The disclosure technology can represent house visits, photographs, videos, or reports about different features of a property.

The example illustrates the role of making continuation menus contingent on reports. An early report determines which outcomes remain available. The customer chooses this menu while uncertain about his type and receives further information only after making that choice. The information ultimately needed to identify the prescribed outcome is therefore disclosed after the report that removes the tempting alternative.

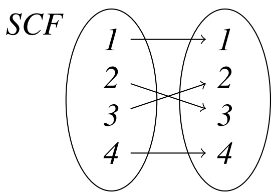
We distinguish mechanisms according to the histories on which future signals can depend. In the example, disclosure depends only on reports. Our main characterization allows the signal-generating technology to condition on both realized signals and reports. A discrepancy between these histories can then change future disclosure, although the allocation rule still depends only on reports.

In this more flexible class, no further information need be supplied after a false report. Withholding information limits the deviating agent to choosing an outcome from the available menu under his current belief. This gives a simple set of incentive constraints. We show that an implementing mechanism can be chosen with finitely many rounds and strictly decreasing posterior supports until the prescribed outcome is determined. We then recursively characterize priors for which the SCF is implementable: the prior must be decomposable into posteriors that satisfy the current incentive constraints and admit implementing continuations. An equivalent linear programming formulation gives an alternative test. With two states or two prescribed outcomes, dynamic disclosure adds nothing to static implementation; this conclusion also applies to the more restricted classes of disclosure.

We next ask which preference structures rule out implementation for every compatible choice of cardinal utilities and full-support prior. For injective social choice functions and strict preferences, the relevant information is the envy relation: state θ envies state θ' when the agent at θ prefers $f(\theta')$ to $f(\theta)$. A recursive condition characterizes whether implementation is possible for some compatible utilities. It yields impossibility results that can be checked without specifying utility magnitudes. Passing this test establishes potential implementability; the cardinal characterization is still needed for a fixed utility function and prior. We also provide a sufficient condition for mechanisms whose disclosure depends only on reports.

Dynamic disclosure can also permit non-monotone social choice functions under preferences satisfying single crossing. The next example shows how a preliminary report can remove the deviation responsible for failure of static implementation.

Example 2 *Suppose there are four equally likely states* $\{1, 2, 3, 4\}$ *and four outcomes* $\{1, 2, 3, 4\}$. *The agent wants to match the state with the outcome,* $u(x; \theta) = -(x - \theta)^2$, *where* x *is the outcome and* θ *is the state. Consider a non-monotone social choice function that wants to match 1 to 1, 2 to 3, 3 to 2, and 4 to 4:*



This social choice function is not statically implementable: type 2 *prefers the outcome assigned to type* 3*, and conversely. On the other hand, it is possible to implement it with the following dynamic signal structure:*

$$s_1 = \begin{cases} odd & \text{if the state is 1 or 3} \\ even & \text{if the state is 2 or 4} \end{cases} \qquad \text{and} \qquad s_2 = \begin{cases} small & \text{if the state is 1 or 2} \\ large & \text{if the state is 3 or 4} \end{cases}.$$

The allocation rule is:

$$a = \begin{cases} 1 & \text{if reports } odd, small \\ 2 & \text{if reports } odd, large \\ 3 & \text{if reports } even, small \\ 4 & \text{if reports } even, large \end{cases}$$

The first round signal reveals whether the state belongs to $\{1, 3\}$ *or* $\{2, 4\}$*, and the two signals together reveal the state. The signals are independent of reports. After an odd signal, the agent assigns equal probability to states 1 and 3. Truthful continuation gives expected utility* $-1/2$*. A false first report of even restricts the allocation to outcomes 3 and 4. Even after observing the second signal, the agent's best deviation is to choose outcome 3 in either state, giving expected utility* -2*. The same calculation applies after an even signal, with outcomes 1 and 2 in the false branch. Within the truthful branch, the prescribed outcome is preferred after each second-round signal. Truthful reporting is therefore optimal.*

Allowing the designer to choose disclosure also gives an upper bound on what can be achieved when information arrives exogenously. If a social choice function cannot be implemented in our most flexible class, no particular disclosure process within that class can implement it.

Our question is related to the implementation literature, including Maskin (1999), because it concerns the restrictions that incentives impose on social choice functions. Our implementation concept is weak: truthful reporting must be optimal and produce the prescribed allocation, but other optimal strategies need not produce the same allocation. It therefore differs from full implementation, which requires every equilibrium outcome to satisfy the social choice rule. We focus on how the timing of information changes truthful implementability for a single agent.

Dynamic mechanism design also studies incentive constraints as private information evolves. Baron and Besanko (1984) analyze regulation in a continuing relationship, and Pavan et al. (2014) characterize incentive-compatible dynamic allocations and the associated information rents. Much of this literature takes the private-information process as given and designs allocations and transfers. We instead fix a terminal social choice function and jointly design disclosure and reporting. Endogenous disclosure is itself an established instrument: Eső and Szentes (2007) study optimal disclosure and auctions. Our question concerns the feasibility of a prescribed state-contingent allocation when reports are alternated with learning.

The paper also relates to sequential information design. Doval and Ely (2020) characterize outcomes attainable by designing information and the extensive form of interaction for a given base game. de Oliveira and Lamba (2023) characterize dynamic choices that can be rationalized by Bayesian learning without observing the underlying information. More directly, de Oliveira and Lamba (2025) study successive actions and signals with terminal utility depending on the action sequence and state, and characterize feasible distributions using deviation rules. Reports in our setting are also successive actions affecting a terminal payoff. Our analysis focuses on implementing a fixed state-contingent social choice function, distinguishes the histories available to the disclosure technology, and develops recursive belief and envy characterizations for this implementation problem.

Section 2 presents the model. Section 3 gives results common to the disclosure classes. Section 4 characterizes implementation when disclosure can depend on signals and reports. Section 5 studies ordinal restrictions, and Section 6 considers disclosure based only on reports. Section 7 discusses multiple agents and stochastic social choice functions. Section 8 concludes. Proofs, including a revelation principle formulated for multiple agents, appear in the appendix.

2 Model

There is an agent who is initially uninformed about his type $\theta \in \Theta = \{1, 2, \cdots, n\}$, where $n < \infty$. The agent's type is distributed according to $\mu(\theta)$, with $\mu(\theta) > 0 \ \forall \theta \in \Theta$. The agent's payoff is $u(x; \theta) \in \mathbb{R}$, where $x \in X$ is the outcome, and X is an arbitrary outcome set. A social choice function is a mapping $f : \Theta \to X$. Denote $X_f = f(\Theta) \subseteq X$, the image of f. The number of (distinct) outcomes of f is $|X_f|$. For instance, when the outcome of an SCF is binary, we have $|X_f| = 2$.

A mechanism M is a vector (T^M, S^M, π^M, a^M). The superscript M is dropped whenever it doesn't create ambiguity. $T \in \mathbb{N} \cup \{\infty\}$ is the number of rounds.[1] The space of signal sequences is $S = \prod_{t=1}^{T} S_t$, where each S_t is finite. At each round $t \leq T$, the agent receives a signal $s_t \in S_t$, which is informative about θ and reports $\hat{s}_t \in S_t$.[2] We denote the sequence of the first t signals received $s^t = (s_1, s_2, \cdots, s_t) \in S^t$ and the sequence of the first t reports made $\hat{s}^t = (\hat{s}_1, \hat{s}_2, \cdots, \hat{s}_t) \in S^t$. For ease of notation we define $S^0 = \{s_0\}$, a singleton set. Moreover, we use the notation $s^{t_2} \supset s^{t_1}$ whenever $t_2 > t_1$ we want to emphasize that the first t_1 elements of s^{t_2} are the ones contained in s^{t_1}. A similar notation is used for reports. The signal structure is a sequence $\pi = (\pi_t)_{t=1}^{T}$, and π_t is a mapping from $S^{t-1} \times S^{t-1} \times \Theta$ to $\Delta(S_t)$. In particular, for any t,

$$\pi_t(s_t \mid s^{t-1}, \hat{s}^{t-1}, \theta)$$

is the probability that the signal received by the agent is s_t given that his previous signals have been s^{t-1}, the earlier reports of the agent were $\hat{s}^{t-1}$, and his type is θ. Finally, the allocation rule $a : S^T \to X$ maps the sequence of reports into an outcome.

A (pure) reporting strategy of the agent, $\sigma = (\sigma_t)_{t=1}^{T}$, is a sequence of mappings $\sigma_t : S^t \to S_t$ where, for any t, $\sigma_t(s^t)$ is the report of the agent given that he observed s^t. Notice that the pair (π, σ) induces a unique probability distribution over the set $S^T \times S^T$. The probability for each (finite) t is defined as

$$\mathbb{P}_{\pi,\sigma}(s^t, \hat{s}^t \mid \theta) = \prod_{\tau=1}^{t} \mathbb{1}\big(\sigma_\tau(s^\tau) = \hat{s}_\tau\big)\pi_\tau(s_\tau \mid s^{\tau-1}, \hat{s}^{\tau-1}, \theta),$$

where $s^t \supset s^\tau$ and $\hat{s}^t \supset \hat{s}^\tau \;\; \forall\tau$, and $\mathbb{1}(\cdot)$ is the indicator function. This probability is

$$\mathbb{P}_\pi(s^t \mid \theta) = \prod_{\tau=1}^{t} \pi_\tau(s_\tau \mid s^{\tau-1}, s^{\tau-1}, \theta),$$

when the agent is truthful. If $t = \infty$, standard extension theorems apply.

Denote the conditional probability that the agent assigns to type θ as a function of history by $\mu(\theta \mid s^t, \hat{s}^{t-1})$, and the belief when he is truthful by $\mu(\theta \mid s^t) = \mu(\theta \mid s^t, s^{t-1})$.

[1] It is without loss of generality to assume T is deterministic. If T is stochastic, we can simply set $T' = \sup \text{Support}\, T$. Then whenever $T < t \leq T'$, the signal distribution gives no information to the agent, and the agent's belief remains the same until T'.

[2] The revelation principle is proved in Lemma 4 in the appendix. We state that result for an arbitrary finite number of agents; the main analysis specializes to one agent.

We denote the agent's belief vectors by $\mu(s^t, \hat{s}^{t-1})$ and $\mu(s^t)$. When $t = \infty$, this conditional belief is well-defined using the Radon-Nikodym derivative of the measure $\mathbb{P}_\pi(\cdot \mid \theta)$.[3] We denote the support of a belief μ by:

$$\operatorname{Supp} \mu = \{\theta \mid \mu(\theta) > 0\}.$$

The set of possible mechanisms is denoted by $\mathbb{M}$. We also consider other classes of mechanisms:

- $\mathbb{M}_R \subset \mathbb{M}$: the subset of mechanisms in which the signal structure only conditions on the history of reports:

$$\pi_t(s_t \mid s^{t-1}, \hat{s}^{t-1}, \theta) = \pi_t(s_t \mid \hat{s}^{t-1}, \theta).$$

- $\mathbb{M}_S \subset \mathbb{M}$: the subset of mechanisms in which the signal structure only conditions on the history of signals:

$$\pi_t(s_t \mid s^{t-1}, \hat{s}^{t-1}, \theta) = \pi_t(s_t \mid s^{t-1}, \theta).$$

- $\mathbb{M}_I \subset \mathbb{M}$: the subset of mechanisms in which the signal structure is independent of the history:

$$\pi_t(s_t \mid s^{t-1}, \hat{s}^{t-1}, \theta) = \pi_t(s_t \mid \theta).$$

In the following, we define implementability.

Definition 1 *Given (u, μ), a mechanism M **implements** f if for all $\theta \in \Theta$:*

$$\int_{s^T \in S^T} \mathbb{1}\big(a(s^T) = f(\theta)\big) d\mathbb{P}_\pi(s^T \mid \theta) = 1,$$

and M is incentive compatible (IC), i.e. for every reporting strategy σ:

$$\sum_{\theta \in \Theta} \mu(\theta) u\big(f(\theta); \theta\big) \geq \sum_{\theta \in \Theta} \mu(\theta) \int_{(s^T, \hat{s}^T) \in S^T \times S^T} u\big(a(\hat{s}^T); \theta\big) d\mathbb{P}_{\pi,\sigma}(s^T, \hat{s}^T \mid \theta).$$

*For $\mathcal{M} \in \{\mathbb{M}, \mathbb{M}_R, \mathbb{M}_S, \mathbb{M}_I\}$, we say f is **$\mathcal{M}$-implementable** if there exists a $M \in \mathcal{M}$ that implements it. Finally, f is statically implementable if there exists a mechanism M with $T = 1$ that implements it.*

[3] The derivative should be taken with respect to the measure $\sum_{\theta \in \Theta} \mu(\theta) \mathbb{P}_\pi(s^T \mid \theta)$. Moreover, note that $\mu(s^t)$ converges to $\mu(s^T)$, almost surely, as t goes to T.

The implementability definition means that the set of signal histories for implementing something other than $f(\theta)$ should be a probability-zero set conditional on θ. When a section specifies a mechanism class $\mathcal{M}$, the word *implementable* refers to that class.

All allocation rules are measurable. Expected utilities of the mechanisms under consideration are assumed to exist; the constructions below have finite allocation ranges and hence bounded payoffs on the finite state space. Pure reporting strategies suffice: randomization is a mixture over contingent reporting plans and cannot improve upon all pure plans.

3 General Results

All results of this section hold for all definitions of $\mathcal{M}$-implementability, i.e., for all $\mathcal{M} \in \{\mathbb{M}, \mathbb{M}_R, \mathbb{M}_S, \mathbb{M}_I\}$. We first see when a static implementation is enough. Partition Θ according to the equivalence classes $[\theta] = \{\tilde{\theta} \in \Theta : f(\tilde{\theta}) = f(\theta)\}$.

Proposition 1 *Given (u, μ), an SCF f is statically implementable* ***if and only if***

$$\forall \theta : \quad f(\theta) \in \arg\max_{x \in X_f} \sum_{\tilde{\theta} \in [\theta]} \mu(\tilde{\theta}) u(x; \tilde{\theta});$$

i.e., if the agent just learns to be in the equivalence class $[\theta]$, the outcome of this equivalence class is the best available outcome for him.

Suppose, $f(\theta') \neq f(\theta)$ whenever $\theta' \neq \theta$. Then the SCF f is statically implementable ***if and only if***

$$\forall \theta : \quad f(\theta) \in \arg\max_{x \in X_f} u(x; \theta),$$

or equivalently,

$$\forall \theta, \theta' : \quad u(f(\theta); \theta) \geq u(f(\theta'), \theta). \tag{2}$$

The second part of the proposition states that we have a static implementation if and only if the agent's preferences fully align with the SCF, and that, in all types, the agent prefers their outcome to all other outcomes. However, when the SCF assigns the same outcome to some types, then the mechanism does not need to reveal everything to the agent. It can just provide the agent with information about the equivalence class their type belongs to, and we just need the agent to prefer their outcome given that information.

The following proposition states that, if the ex post incentive condition (2) fails, there is a full-support belief for which the SCF is not dynamically implementable.

Proposition 2 *Given u, there exists a (full support) belief μ for which the SCF f is not implementable, if and only if Equation* (2) *doesn't hold, i.e.,*

$$\exists \theta, \theta' : \quad u(f(\theta); \theta) < u(f(\theta'), \theta).$$

When $\mu(\theta)$ is close enough to one, whatever the mechanism is, the agent benefits from reporting in a way that he always gets the outcome of θ'. Together, these two propositions establish a fundamental dichotomy: either an SCF is statically implementable (no matter the prior belief), or there exists at least one belief under which no mechanism, static or dynamic, can implement it.

The following lemma shows that when a mechanism implements the SCF, the posterior belief $\mu(\cdot|s^T)$ puts probability one on a single equivalence class, almost surely. Denote the posterior probability of an equivalence class given the joint signal history s^T as:

$$\mu([\theta]|s^T) = \sum_{\tilde{\theta} \in [\theta]} \mu(\tilde{\theta}|s^T).$$

Lemma 1 *Given (u, μ), in any mechanism that implements the SCF f, $\mu([\theta]|s^T) = 1$ for some type $\theta \in \Theta$, almost surely. Formally, for any $\theta \in \Theta$,*

$$\int_{s^T \in S^T} \mathbb{1}\big(\mu([\theta]|s^T) = 1\big) d\mathbb{P}_\pi(s^T|\theta) = 1.$$

4 Conditioning on Signals and Reports

Throughout this section, implementation refers to $\mathbb{M}$. The disclosure technology can respond to a mismatch between past signals and reports, while the allocation rule uses only reports. The following lemmas make precise the available menu and the incentives when further information is withheld after a false report.

Lemma 2 *Any mechanism in $\mathbb{M}$ that implements f can be replaced by one with the same distribution of truthful signals and allocations such that no further information is provided after the first false report. Moreover, after any reported history s^t that has* *positive probability under truth-telling, the outcomes obtainable through subsequent reports are exactly* $f(\operatorname{Supp}\mu(\cdot \mid s^t))$.

For any reported history $\hat{s}^t$, define the available menu by

$$A(\hat{s}^t) := \{a(\hat{s}^T) : \hat{s}^T \in S^T,\ \hat{s}^T \supset \hat{s}^t\}. \tag{3}$$

This definition includes every continuation of the reported history, including continuations that have probability zero under truth-telling. By Lemma 2, we may work with mechanisms for which, at every positive-probability truthful history s^t,

$$A(s^t) = \{f(\theta) : \mu(\theta \mid s^t) > 0\}.$$

After observing s^{t+1}, an agent who has previously reported s^t can obtain any outcome in $A(s^t)$ by committing to a suitable sequence of reports. Once he makes a false report, he receives no further information. The following inequalities therefore characterize incentive compatibility.

Lemma 3 *Suppose a mechanism produces f under truthful reporting and provides no further information after a false report. It is incentive compatible if and only if, for every $t < T$, every positive-probability truthful history s^{t+1}, and every $x \in A(s^t)$,*

$$\sum_{\theta\in\Theta} \mu(\theta \mid s^{t+1}) u(f(\theta);\theta) \geq \sum_{\theta\in\Theta} \mu(\theta \mid s^{t+1}) u(x;\theta).$$

The menu inequalities allow us to remove learning stages that do not reduce the available outcomes. The proof below also covers an infinite original horizon by using the terminal posterior on paths with no finite menu reduction.

Theorem 1 *Without loss of generality, at every nonterminal truthful history, each next signal strictly reduces the set of available outcomes. Moreover, the signal sets can be chosen so that*

$$|S_t| \leq n - t + 1$$

at every active round t.

Proposition 3 *Any mechanism in $\mathbb{M}$ that implements f can be replaced by one with $T \leq |X_f| - 1 \leq n - 1$.*

Corollary 1 *When $|X_f| = 2$, one round is without loss of generality.*

A constant target is also statically implementable. Thus, dynamic disclosure adds nothing to static implementation when $|\Theta| \leq 2$.

Corollary 2 *Suppose $\Theta = \{1, 2, 3\}$, f is injective, and μ has full support. Write $u_{ji} = u(f(j); i)$ and $\mu_i = \mu(i)$. Up to relabeling, exactly one of the following cases applies.*

1. *If $u_{ii} \geq u_{ji}$ for every i, j, then f is statically implementable.*
2. *Suppose the only strict envy relation is $3E2$, that is, $u_{23} > u_{33}$ and $u_{ii} \geq u_{ji}$ for every other ordered pair $i \neq j$. Then f is implementable if and only if*

$$\mu_1(u_{11} - u_{21}) \geq \mu_3(u_{23} - u_{33}).$$

3. *Suppose the only strict envy relations are $2E3$ and $3E2$. Then f is implementable if and only if there exists $\alpha \in [0, \mu_1]$ such that*

$$\alpha(u_{11} - u_{31}) \geq \mu_2(u_{32} - u_{22}),$$
$$(\mu_1 - \alpha)(u_{11} - u_{21}) \geq \mu_3(u_{23} - u_{33}).$$

4. *In all other cases, f is not implementable.*

Whenever implementation is possible in cases 2 or 3, it is achievable in two rounds with a mechanism in $\mathbb{M}_R$.

For case 2, let $S_1 = \{\{2\}, \{1, 3\}\}$ and $S_2 = \Theta$. In the first round, reveal whether $\theta = 2$:

$$s_1 = \begin{cases} \{2\}, & \theta = 2, \\ \{1, 3\}, & \theta \in \{1, 3\}. \end{cases}$$

After first-round report $\hat{s}_1 \in S_1$, send

$$s_2 = \begin{cases} 2, & \hat{s}_1 = \{2\}, \\ \theta, & \hat{s}_1 = \{1, 3\} \text{ and } \theta \in \{1, 3\}, \\ 1, & \hat{s}_1 = \{1, 3\} \text{ and } \theta = 2. \end{cases}$$

Define the allocation for every report pair by

$$a(\hat{s}_1, \hat{s}_2) = \begin{cases} f(2), & \hat{s}_1 = \{2\}, \\ f(3), & \hat{s}_1 = \{1,3\} \text{ and } \hat{s}_2 = 3, \\ f(1), & \hat{s}_1 = \{1,3\} \text{ and } \hat{s}_2 \neq 3. \end{cases}$$

Thus, reporting the branch $\{1,3\}$ leaves only $f(1)$ and $f(3)$ available. The inequality in case 2 deters reporting $\{2\}$ after receiving $\{1,3\}$; all remaining incentive constraints follow from the assumed preferences.

For case 3, choose any feasible α and let $p = \alpha/\mu_1$. Set $S_1 = \{\{1,2\},\{1,3\}\}$ and $S_2 = \Theta$. In state 1, send $\{1,2\}$ with probability p and $\{1,3\}$ with probability $1-p$. In state $i \in \{2,3\}$, send $\{1,i\}$ with probability one. After first-round report $\hat{s}_1 \in S_1$, send

$$s_2 = \begin{cases} \theta, & \theta \in \hat{s}_1, \\ 1, & \theta \notin \hat{s}_1, \end{cases} \qquad a(\hat{s}_1, \hat{s}_2) = \begin{cases} f(\hat{s}_2), & \hat{s}_2 \in \hat{s}_1, \\ f(1), & \hat{s}_2 \notin \hat{s}_1. \end{cases}$$

Within either truthful branch, the two states weakly prefer their prescribed outcomes to each other's. After a false first-round report, the second signal is always 1 and reveals no additional information. The two inequalities in case 3 deter choosing the other branch and its envied outcome.

Corollary 3 *Without loss of generality, each active signal strictly reduces posterior state support, and terminal posteriors are degenerate. For $n \geq 2$, this can be achieved in at most $n-1$ active rounds.*

To obtain this representation, take the mechanism of Theorem 1. Whenever a branch has a constant target but retains several states, reveal its remaining state and keep its allocation fixed regardless of that final report. This does not change incentives. Along every active edge at least one state is eliminated, so the number of active rounds is at most $n-1$. This representation is useful for the belief recursion and the linear characterization; it need not attain the sharper outcome-based bound at the same time.

When a prior q has a proper support $B \subsetneq \Theta$, implementation at q refers to the problem with state space B, utility functions restricted to B, and target $f|_B$.

Fix u and f. For $k = 1, \dots, n$, let

$$\Gamma^k := \{q \in \Delta(\Theta) : |\operatorname{Supp} q| \leq k \text{ and } f \text{ is implementable at } q\}.$$

For every nonempty $B \subseteq \Theta$, let $IC(B)$ be the set of beliefs $\nu \in \Delta(\Theta)$ with $\operatorname{Supp} \nu \subseteq B$ such that

$$\sum_{\theta \in B} \nu(\theta) u(f(\theta);\theta) \geq \sum_{\theta \in B} \nu(\theta) u(x;\theta) \qquad \text{for every } x \in f(B).$$

These are the posterior beliefs satisfying the inequalities in Lemma 3 when the preceding truthful history has support B.

Proposition 4 *The implementable belief sets satisfy*

$$\Gamma^1 = \{\delta_\theta : \theta \in \Theta\}, \qquad \Gamma^k = \bigcup_{\substack{\varnothing \neq B \subseteq \Theta \\ |B| \leq k}} \operatorname{co}\big(IC(B) \cap \Gamma^{k-1}\big), \quad k = 2, \dots, n.$$

The recursion separates the current reporting constraint from the continuation problem. A current signal must induce a posterior in $IC(B)$, where B is the preceding support. Its continuation must implement the restricted target. Bayes plausibility requires the current prior to be a convex combination of these posteriors.

A finite linear characterization. The recursive characterization also yields a finite linear feasibility problem. Let $n = |\Theta|$, and let H_t be the set of ordered lists of t distinct elements of Θ, with $H_0 = \{\varnothing\}$. Write $H = \bigcup_{t=0}^{n-1} H_t$ and $B_h = \Theta \setminus \{\theta_1, \dots, \theta_t\}$ for $h = (\theta_1, \dots, \theta_t)$. For $j \in B_h$, let hj denote the list obtained by appending j to h.

Proposition 5 *The SCF f is implementable under (u, μ) in $\mathbb{M}$ if and only if there exist vectors $q_h \in \mathbb{R}_+^\Theta$, one for each $h \in H$, satisfying*

$$q_\varnothing = \mu, \tag{4}$$

$$q_h(\theta) = 0 \quad (h \in H,\ \theta \notin B_h), \tag{5}$$

$$q_h = \sum_{j \in B_h} q_{hj} \quad (h \in H_t,\ 0 \leq t \leq n-2), \tag{6}$$

$$\sum_{\theta \in \Theta} q_{hj}(\theta)\big[u(f(\theta);\theta) - u(x;\theta)\big] \geq 0. \tag{7}$$

The last inequality is imposed for every $h \in H_t$ with $0 \leq t \leq n-2$, every $j \in B_h$, and every $x \in f(B_h)$.

The vector q_h records the joint probabilities of reaching h and of the different states. Its total mass is the probability of that history; whenever this mass is positive, its normalization is the posterior. At a terminal history $h \in H_{n-1}$, the remaining set B_h is a singleton, so the allocation is fixed by f. Zero-mass histories are ignored. Thus all restrictions are linear in the unknown vectors q_h.

The following proposition characterizes when $IC(\Theta)$ contains a full-support belief.

Proposition 6 *The set $IC(\Theta)$ contains no full-support belief if and only if there exists a lottery $p \in \Delta(X_f)$ such that*

$$\sum_{x \in X_f} p(x) u(x;\theta) \geq u(f(\theta);\theta) \qquad \text{for every } \theta \in \Theta,$$

with strict inequality for at least one state.

Thus, the obstruction is a lottery over outcomes in the range of the SCF that is weakly preferred to the prescribed outcome in every state and strictly preferred in at least one state.

Observation 1 *Suppose there exists $\theta_0 \in \Theta$ such that*

$$u(f(\theta_0);\theta_0) > u(x;\theta_0) \quad \text{for every } x \in X_f \setminus \{f(\theta_0)\},$$

and

$$u(f(\theta);\theta) \geq u(f(\theta_0);\theta) \quad \text{for every } \theta \in \Theta.$$

Then f is implementable in at most two rounds for every full-support prior sufficiently concentrated on θ_0.

Proof: If $|X_f| = 1$, implementation is immediate. Otherwise, relabel $\theta_0 = 1$ and $\Theta = \{1, \ldots, n\}$. Set $S_1 = \{2, \ldots, n\}$ and $S_2 = \{1, \neg 1\}$. In the first round, send each signal with probability $1/(n-1)$ in state 1, and send signal i with probability one in state $i \neq 1$. After first-round report $\hat{s}_1 = i$, send $s_2 = \neg 1$ if $\theta = i$ and $s_2 = 1$ otherwise. Define

$$a(i, \hat{s}_2) = \begin{cases} f(1), & \hat{s}_2 = 1, \\ f(i), & \hat{s}_2 = \neg 1. \end{cases}$$

Truthful reporting implements f. Conditional on first-round signal i, the support is $\{1, i\}$; reporting truthfully in the second round is optimal by the assumed preferences between $f(1)$ and $f(i)$.

If the agent instead reports $j \neq i$ in the first round, the second signal is 1 in both possible states and reveals no additional information. The agent can then obtain either $f(1)$ or $f(j)$. Thus, it is sufficient that, for all $i \neq 1$ and $x \in X_f$,

$$\mu(1)\big[u(f(1);1) - u(x;1)\big] \geq (n-1)\mu(i)\big[u(x;i) - u(f(i);i)\big].$$

For $x = f(1)$, the inequality follows from the second assumption. For every other x, its left-hand coefficient is strictly positive. Since there are finitely many states and outcomes, all these inequalities hold when $\mu(1)$ is sufficiently close to one. □

5 Never-Implementable

Implementation in this section refers to $\mathbb{M}$.

The preceding results depend on cardinal utilities and prior beliefs. For example, the incentive constraint in Example 1 compares a probability-weighted loss in state 1 with a probability-weighted gain in state 3. Corollary 2 shows that reversing this inequality makes that SCF unimplementable. We now ask which ordinal preference patterns permit implementation for some compatible cardinal utilities and a full-support prior.

Because the mechanisms studied in this section can be restricted to allocations in X_f, let $\preceq = (\preceq_\theta)_{\theta\in\Theta}$ denote a family of preference orderings on X_f. Write $\preceq^u$ for the orderings induced by u on this finite set.

Definition 2 *The SCF f is* ***never-implementable for u****, if there exists no full-support belief μ, for which f is implementable. The SCF f is* ***never-implementable for $(\preceq, \mu)$****, if there exists no u with $\preceq^u = \preceq$ such that f is implementable for (u, μ). The SCF f is* ***never-implementable for $\preceq$****, if for any full-support belief μ, f is never implementable for $(\preceq, \mu)$.*

Proposition 7 *For any SCF f, any full-support prior μ, and any preference ordering $\preceq$ on X_f:*

1. *f is never-implementable for $(\preceq, \mu)$ if and only if it is never-implementable for $\preceq$.*

2. *If f is never-implementable for $\preceq^u$, it is never-implementable for u. The converse need not hold.*

Proof: The "if" part in 1 follows trivially by the definition of never-implementation. Let us prove the "only if" part. Suppose $\preceq$ is not never-implementable. So, there exists an implementable pair (u', μ') where $\preceq^{u'} = \preceq$. That is, a mechanism M exists that implements the SCF. It is easy to check that M is also IC for (u, μ) and implements it, where for any $(\theta, x) \in \Theta \times X_f$,

$$u(x;\theta) := \frac{\mu'(\theta)}{\mu(\theta)} u'(x;\theta).$$

The "if" part of 2 is trivial. Let us construct an example that shows that even if $\preceq^u$ is not never-implementable, it can be the case that u is never-implementable. Let $\Theta = \{1,2,3,4\}$, set $f(i) = x_i$, and choose the utility matrix

$$\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & -1 & -1 & -1 \\ x & 0 & -1 & -1 \\ -1 & -1 & 0 & -1 \\ -1 & x & x & 0 \end{pmatrix} \end{array},$$

where $x > 0$. These entries specify this example only. For $0 < x < 1/2$ and a uniform prior, the following successive partitions implement f: $\{\{1,2,3\},\{4\}\}$ in the first round, $\{\{1,3\},\{2\},\{4\}\}$ in the second round, and the singleton partition in the third round.

However, when $x > 2$, u is never-implementable. By martingale property of beliefs, there should exist posteriors that don't exclude $\theta = 2$. Consider one signal s_1 and $\mu = \mu(s_1)$ such that $\mu_2 > 0$. IC requires

$$\begin{cases} \mu_1 x \le \mu_3 + \mu_4 \le 1 \\ (\mu_2 + \mu_3) x \le \mu_1 \le 1 \end{cases} \implies \begin{cases} \mu_1 \le \frac{1}{x} < \frac{1}{2} \\ \mu_2 + \mu_3 \le \frac{1}{x} < \frac{1}{2} \end{cases} \implies \mu_4 > 0.$$

By Theorem 1, it is WLOG to decrease the support of the beliefs. This means that it is WLOG to assume $\mu_1 = 0$ or $\mu_3 = 0$, or both. If $\mu_1 = \mu_3 = 0$, μ is not implementable since 2 would mimic 4. If either $\mu_1 = 0$ or $\mu_3 = 0$, we have two envy edges among the remaining three states that do not form a reciprocal pair, which means by Corollary 2, μ is not implementable. $\square$

Throughout the remainder of this section, assume that f is injective and that preferences are strict on $X_f = f(\Theta)$. Define the envy relation E by

$$\theta E \theta' \quad \Longleftrightarrow \quad u(f(\theta');\theta) > u(f(\theta);\theta).$$

Thus, for distinct states θ, θ', the absence of envy means $u(f(\theta');\theta) < u(f(\theta);\theta)$. We call f *possibly implementable given* E if it is implementable for some full-support prior and some utility representation inducing E. Unless stated otherwise, implementation in this section allows signal structures to condition on both past signals and past reports.

Definition 3 *The SCF f is* ***never-implementable for an envy relation*** *E if it is never-implementable for every strict preference ordering $\preceq$ on X_f inducing E.*

Proposition 8 *Let $\preceq$ be a strict preference ordering on X_f and let $E^{\preceq}$ be its envy relation. Then f is never-implementable for $\preceq$ if and only if it is never-implementable for $E^{\preceq}$.*

Proposition 9 *If $|\Theta| = 1$, then f is possibly implementable given E. If $|\Theta| \geq 2$, then f is possibly implementable given E if and only if there exist nonempty proper subsets $\Theta_1, \Theta_2 \subsetneq \Theta$ such that*

1. *$\Theta_1 \cup \Theta_2 = \Theta$;*
2. *for each $i \in \{1,2\}$, the restriction $f|_{\Theta_i}$ is possibly implementable given $E|_{\Theta_i \times \Theta_i}$;*
3. *for each $i \in \{1,2\}$ and each $\theta' \in \Theta \setminus \Theta_i$, there is a state $\theta \in \Theta_i$ such that $\neg(\theta E \theta')$.*

The subsets need not be disjoint. Moreover, whenever these conditions hold, for every full-support prior μ and every strict preference ordering on X_f inducing E, there is a utility representation of that ordering for which f is implementable at μ.

Example 3 *Let $\Theta = \{1,2,3\}$, and suppose the only envy relations are $1E3$ and $3E1$. Take*

$$\Theta_1 = \{1,2\}, \qquad \Theta_2 = \{2,3\}.$$

Each restriction has no envy edges and is statically implementable. The shared state 2 envies neither of the excluded states. The covering condition therefore holds, and the SCF is possibly implementable in two rounds. The first signal pools either state 1 or state 3 with state 2; the second signal reveals the state within the reported branch.

Example 4 *Let* $\Theta = \{1, 2, 3, 4\}$ *and*

$$E = \{(1, 3), (3, 1), (2, 4), (4, 2), (1, 4)\}.$$

Consider the cover

$$\Theta_1 = \{1, 2, 3\}, \qquad \Theta_2 = \{3, 4\}.$$

The restriction to Θ_1 *is possibly implementable by Example 3, and the restriction to* Θ_2 *contains no envy. Within* Θ_1*, state* 3 *does not envy the excluded state* 4*. Within* Θ_2*, state* 4 *does not envy state* 1*, and state* 3 *does not envy state* 2*. Thus Proposition 9 applies even though one of the continuation problems itself requires gradual disclosure.*

Corollary 4 *Suppose* $|\Theta| \geq 2$*,* f *is injective, and preferences are strict on* X_f*. The following statements hold for the envy relation* E*.*

1. *If there is a state* j *such that* iEj *for every* $i \neq j$*, then* E *is never-implementable.*
2. *If there is a state* i *such that* iEj *for every* $j \neq i$*, then* E *is never-implementable.*
3. *If the underlying undirected envy graph is disconnected, then* E *is possibly implementable. In fact, for every full-support prior and every strict preference ordering inducing* E*, some utility representation of that ordering admits implementation in two rounds in* $\mathbb{M}$*.*

6 Conditioning on Reports

Here implementation refers to $\mathbb{M}_R$: disclosure can depend on past reports and the state, but cannot respond separately to a mismatch with past signals. Example 1 belongs to this class. The information-withholding construction in Lemma 2 need not be available, so the menu inequalities alone do not characterize incentives. The following results establish special cases and a sufficient ordinal condition.

Proposition 10 *If* $|X_f| \leq 2$*, implementation in* $\mathbb{M}_R$ *is equivalent to static implementation. For three states and three prescribed outcomes, Corollary 2 also characterizes implementation in* $\mathbb{M}_R$*.*

Proof: The first claim follows from $\mathbb{M}_R \subseteq \mathbb{M}$ and Corollary 1, since every static mechanism belongs to $\mathbb{M}_R$. For the second claim, necessity follows from the same inclusion, and the constructions in Corollary 2 depend only on reports. $\square$

Proposition 11 *Assume that f is injective and preferences are strict on X_f. Suppose there exists a finite rooted binary tree whose nodes are nonempty subsets of Θ, with root Θ and singleton leaves, such that at every nonterminal node A, its two children A_1, A_2 are proper subsets of A, satisfy $A_1 \cup A_2 = A$, and obey*

$$\text{for each } i \in \{1,2\} \text{ and each } \theta' \in A_{3-i}, \qquad \exists \theta \in A_i \setminus A_{3-i} \text{ such that } \neg(\theta E \theta').$$

Then, for every full-support prior μ and every strict preference ordering on X_f inducing E, there is a utility representation of that ordering for which a report-only mechanism implements f.

The condition also applies to disjoint covers. When children overlap, it requires an exclusive state to deter each outcome in the other child, including outcomes common to the two children.

Conjecture 1 *Assume that f is injective and preferences are strict on X_f. Suppose that nonempty proper subsets $\Theta_1, \Theta_2 \subsetneq \Theta$ cover Θ, that each restriction $f|_{\Theta_i}$ is possibly implementable using report-only mechanisms, and that*

$$\text{for each } i \in \{1,2\} \text{ and each } \theta' \in \Theta_{3-i} \setminus \Theta_i, \qquad \exists \theta \in \Theta_i \setminus \Theta_{3-i} \text{ such that } \neg(\theta E \theta').$$

f is then possibly implementable using a report-only mechanism.

Unlike Proposition 11, this condition does not require an exclusive witness for an outcome common to the two subsets. A proof must choose the two continuation mechanisms jointly so that a false report cannot reveal information about their common states that makes a later deviation profitable.

The stronger report-only condition need not hold for a cover that satisfies Proposition 9. In Example 4, the exclusive parts of the displayed cover are $\Theta_1 \setminus \Theta_2 = \{1,2\}$ and $\Theta_2 \setminus \Theta_1 = \{4\}$. Both states 1 and 2 envy state 4. Hence this cover fails the report-only sufficient condition. Failure of this sufficient condition does not establish that the SCF is impossible to implement with report-only mechanisms.

Conjecture 2 *Every SCF implementable in $\mathbb{M}_R$ admits a report-only implementing mechanism with at most two possible signals at each active history and a strict reduction of the available outcome menu after every active signal until that menu is a singleton.*

The unrestricted argument does not establish this claim. Replacing a many-signal disclosure by successive binary disclosures changes when the agent learns and can report. Copying a continuation between overlapping branches also requires checking its incentives under the belief induced by a false report. A proof preserving these incentives is still needed. No result stated as proved here relies on this conjecture or on Conjecture 1.

7 Discussion

7.1 Multiple agents

With N agents, types are drawn independently from full-support priors μ_i on finite sets Θ_i. Agent i has utility $u_i(x;\theta_i)$, and the social choice function is $f : \prod_i \Theta_i \to X$. The disclosure technology takes the form

$$\pi_{it}(s_{it} \mid s_i^{t-1}, \hat{s}^{t-1}, \theta_i).$$

Conditional on the type profile and past histories, the agents' current signals are drawn independently according to these kernels. Thus, disclosure may respond to other agents' reports, although it does not directly condition on their types or private signals. Truthful reporting must maximize each agent's ex ante expected utility against every reporting strategy profile of the other agents. The revelation principle in the appendix applies to this setting.

Even static implementation introduces an additional restriction: a signal about one agent's type must work across all types of the other agents. Consider two agents with $\Theta_i = \{1, 2, 3\}$ and uniform priors, and let

$f(\theta_1, \theta_2)$	1	2	3
1	A	A	C
2	A	B	B
3	C	B	C

$u(x;\theta_i)$	A	B	C
1	2	0	10
2	10	2	0
3	0	10	2

.

Both agents have the displayed utility function. For each fixed θ_{-i}, the resulting single-agent problem is statically implementable. For example, when $\theta_{-i} = 1$, pooling types $\{1, 2\}$ induces them to choose A: their expected utilities from A and C are 6 and 5. Type 3 chooses C. The other two cases follow by cyclic permutation.

Nevertheless, the two-agent social choice function is not statically implementable. Every pair of rows of the allocation matrix differs in some column. If two types of agent 1 could receive the same signal, combining that signal with any signal of the distinguishing type of agent 2 would require two different allocations at the same report profile. Hence every signal reveals agent 1's type; the same argument applies to agent 2. Now let agent 2 always report a signal that can occur at type 1. When agent 1 learns that his type is 1, reporting a signal of type 3 changes his allocation from A to C, increasing his utility from 2 to 10. This contradicts dominance.

This example shows that static implementability of every single-agent restriction does not imply static implementability of the joint social choice function. Determining whether gradual disclosure can overcome this restriction requires a separate dynamic argument.

7.2 Stochastic social choice functions

The preceding analysis requires the prescribed outcome to be delivered on almost every truthful terminal history, conditional on each state. For a stochastic SCF $F : \Theta \to \Delta(Y)$ over primitive outcomes Y, a different requirement is natural: conditional on each state θ, the distribution of the realized allocation must equal $F(\theta)$. This permits the randomization of allocations to depend on the information disclosed. Treating each prescribed lottery as an indivisible outcome instead imposes the stronger requirement that the same lottery be assigned on almost every truthful terminal history in that state.

Example 5 (Implementation of a stochastic SCF) *Let $\Theta = \{1, 2, 3\}$, let the prior be uniform, and let $Y = \{\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_3\}$. Utilities over primitive outcomes are*

$$\begin{array}{c|ccc} & \theta = 1 & \theta = 2 & \theta = 3 \\ \hline \widetilde{x}_1 & -2 & -5 & 3 \\ \widetilde{x}_2 & 2 & -1 & -1 \\ \widetilde{x}_3 & 0 & 1 & -3 \end{array}.$$

The target lotteries are

$$F(1) = \tfrac{1}{2}\delta_{\widetilde{x}_1} + \tfrac{1}{2}\delta_{\widetilde{x}_2}, \qquad F(2) = \tfrac{1}{2}\delta_{\widetilde{x}_2} + \tfrac{1}{2}\delta_{\widetilde{x}_3}, \qquad F(3) = \tfrac{1}{2}\delta_{\widetilde{x}_3} + \tfrac{1}{2}\delta_{\widetilde{x}_1}.$$

Writing $x_i = F(i)$ and evaluating these lotteries by expected utility gives

$$\begin{array}{c|rrr} & \theta = 1 & \theta = 2 & \theta = 3 \\ \hline x_1 & 0 & -3 & 1 \\ x_2 & 1 & 0 & -2 \\ x_3 & -1 & -2 & 0 \end{array}.$$

If these three lotteries are treated as indivisible outcomes, the resulting deterministic SCF is not implementable: its only envy relations are $1E2$ *and* $3E1$*, a case ruled out by the three-state characterization.*

A one-round mechanism implements the stochastic SCF. It uses signals 12, 23*, and* 13*, with*

$$\pi(ij \mid \theta) = \begin{cases} \frac{1}{2}, & \theta \in \{i, j\}, \\ 0, & \theta \notin \{i, j\}. \end{cases} \qquad a(12) = \widetilde{x}_2, \quad a(23) = \widetilde{x}_3, \quad a(13) = \widetilde{x}_1.$$

Each signal induces the uniform posterior on its indicated pair. The expected utilities of the three possible reports are

$$\begin{array}{c|rrr} & \widehat{s} = 12 & \widehat{s} = 23 & \widehat{s} = 13 \\ \hline s = 12 & \frac{1}{2} & \frac{1}{2} & -\frac{7}{2} \\ s = 23 & -1 & -1 & -1 \\ s = 13 & \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \end{array}.$$

Truthful reporting is therefore optimal after every signal. In each state, the two prescribed primitive outcomes are delivered with equal probability, as required by F.

Here, allocating while the agent remains uncertain about his type supports truthful reporting. The example also shows why this extension is not obtained simply by enlarging the deterministic outcome space to include the target lotteries.

8 Conclusion

Gradual disclosure changes implementation because an agent can commit to a continuation menu before learning which outcome he would ultimately prefer. When disclosure can condition on past signals and reports, withholding information after a deviation leads to a recursive characterization of implementable priors and an equivalent linear feasibility test. Under injectivity and strict preferences, a recursive condition on the envy relation identifies which social choice functions can be implemented for some compatible cardinal utilities.

The distinction between disclosure technologies matters for these conclusions. For disclosure based only on reports, we provide a sufficient condition, while a complete characterization remains open. Extending the analysis to strategic interactions among several learning agents is a further direction. The general revelation principle in the appendix provides a starting point, but the single-agent menu arguments do not by themselves establish the corresponding multi-agent results.

References

Baron, D. P. and Besanko, D. (1984). Regulation and information in a continuing relationship. *Information Economics and Policy*, 1(3):267–302.

de Oliveira, H. and Lamba, R. (2023). Rationalizing dynamic choices. Working paper, September. `https://ssrn.com/abstract=3332092`.

de Oliveira, H. and Lamba, R. (2025). Dynamic information design via deviation rules. Working paper, April. Author's manuscript.

Doval, L. and Ely, J. C. (2020). Sequential information design. *Econometrica*, 88(6):2575–2608.

Eső, P. and Szentes, B. (2007). Optimal information disclosure in auctions and the handicap auction. *The Review of Economic Studies*, 74(3):705–731.

Maskin, E. (1999). Nash equilibrium and welfare optimality. *The Review of Economic Studies*, 66(1):23–38.

Pavan, A., Segal, I., and Toikka, J. (2014). Dynamic mechanism design: A Myersonian approach. *Econometrica*, 82(2):601–653.

Appendix

For the revelation principle, consider N agents with finite type spaces Θ_i, independent full-support priors μ_i, and utilities $u_i(x;\theta_i)$. Signals to agent i may depend on his type, his own past signals, and all agents' past messages. Conditional on the type profile and past histories, the agents' current signals are drawn independently according to these kernels. Messages are private: agent i observes his own signals and messages. A pure strategy can therefore be written as a function of his signal history, since his own past messages are determined recursively by that strategy. Dominance means optimality in ex ante expected utility against every strategy profile of the other agents. The single-agent model is the case $N = 1$.

Lemma 4 (Revelation principle) *If an indirect mechanism implements f in pure ex ante dominant strategies, a direct mechanism implements f with truthful signal reporting as an ex ante dominant strategy. The construction preserves the restrictions defining $\mathbb{M}_R$, $\mathbb{M}_S$ and $\mathbb{M}_I$ whenever the indirect mechanism satisfies the corresponding restrictions.*

Proof: Write $\Theta = \prod_i \Theta_i$ and $\mu = \prod_i \mu_i$. Consider an indirect mechanism with horizon T, signal spaces S_i, message spaces $M_i = \prod_t M_{it}$, signal kernels

$$\widetilde{\pi}_{it}(s_{it} \mid s_i^{t-1}, m^{t-1}, \theta_i),$$

and terminal allocation $g(m^T)$. Let $\widetilde{\sigma}^* = (\widetilde{\sigma}_i^*)_i$ be an implementing profile of pure ex ante dominant strategies. For any reported signal history $\widehat{s}_i^t$, define the associated equilibrium message history by

$$\widetilde{\sigma}_i^{*t}(\widehat{s}_i^t) = \big(\widetilde{\sigma}_{i1}^*(\widehat{s}_i^1), \dots, \widetilde{\sigma}_{it}^*(\widehat{s}_i^t)\big).$$

The direct mechanism uses the same signal spaces and sets

$$a(\widehat{s}^T) = g\big(\widetilde{\sigma}^{*T}(\widehat{s}^T)\big),$$
$$\pi_{it}(s_{it} \mid s_i^{t-1}, \widehat{s}^{t-1}, \theta_i) = \widetilde{\pi}_{it}\big(s_{it} \mid s_i^{t-1}, \widetilde{\sigma}^{*(t-1)}(\widehat{s}^{t-1}), \theta_i\big).$$

For a direct reporting strategy σ_i, its induced reported history and corresponding indirect strategy are

$$\widetilde{s}_i^t(s_i^t) = \big(\sigma_{i1}(s_i^1), \dots, \sigma_{it}(s_i^t)\big),$$
$$\widetilde{\sigma}'_{it}(s_i^t) = \widetilde{\sigma}^*_{it}\big(\widetilde{s}_i^t(s_i^t)\big).$$

Thus the equilibrium strategy is evaluated at the entire reported history. For any direct strategy profile σ, the induced law of signals and outcomes equals the law in the indirect mechanism under $\widetilde{\sigma}'$. This follows recursively from equality of the signal kernels and the simulated messages at every finite history. For an infinite horizon, equality on cylinder events gives equality of the induced measures.

Let $U_i^{\mathrm{dir}}(\sigma)$ and $U_i^{\mathrm{ind}}(\widetilde{\sigma})$ denote ex ante expected utilities in the two mechanisms. Fix an arbitrary direct strategy profile σ_{-i} of the other agents. Truthful reporting I_i induces $\widetilde{\sigma}_i^*$ in the indirect mechanism, because $\widetilde{s}_i^t(s_i^t) = s_i^t$. Consequently, for every direct deviation σ_i,

$$\begin{aligned} U_i^{\mathrm{dir}}(I_i, \sigma_{-i}) &= U_i^{\mathrm{ind}}(\widetilde{\sigma}_i^*, \widetilde{\sigma}'_{-i}) \\ &\geq U_i^{\mathrm{ind}}(\widetilde{\sigma}'_i, \widetilde{\sigma}'_{-i}) = U_i^{\mathrm{dir}}(\sigma_i, \sigma_{-i}). \end{aligned}$$

The inequality is ex ante dominance in the indirect mechanism. Truthful reporting is therefore ex ante dominant in the direct mechanism. When all agents report truthfully, the simulated law is the original implementing law, so $a(s^T) = f(\theta)$ almost surely conditional on every type profile.

Finally, simulated past messages depend only on past reports. Substituting them in a report-only kernel therefore preserves that restriction. A kernel independent of messages remains independent of reports, and a history-independent kernel remains independent of histories. This proves the asserted preservation of classes. □

Proof of Lemma 1: This argument also applies when Θ is a finite space of type profiles and H is the joint terminal signal history. Let P_θ be the truthful law of H conditional on θ, and let $P = \sum_\theta \mu(\theta) P_\theta$ be its prior mixture. Implementation gives $a(H) = f(\theta)$ almost surely under the joint law of (θ, H). Conditioning on H yields

$$\sum_{\vartheta \in \Theta:\, f(\vartheta) = a(H)} \mu(\vartheta \mid H) = 1 \qquad P\text{-almost surely.}$$

Full support of the prior implies $P_\theta \ll P$ for every θ. Under P_θ, moreover, $a(H) = f(\theta)$ almost surely. Hence $\mu([\theta] \mid H) = 1$ P_θ-almost surely, as required. The same conditioning argument applies to the full infinite signal history; it does not require that a particular terminal history have positive probability. $\square$

Proof of Lemma 2. Let M implement f, and write P for its truthful distribution of signal histories. For each finite history h with $P(h) > 0$, let $B(h) = \operatorname{Supp} \mu(\cdot \mid h)$. First, every $x \in f(B(h))$ is obtainable by fixing a report sequence extending h. Indeed, choose $\theta \in B(h)$ with $f(\theta) = x$. Conditional on θ and h, truthful continuation implements x with probability one, so at least one such report sequence exists. Consequently, incentive compatibility of M implies

$$\sum_\theta \mu(\theta \mid hs) u(f(\theta); \theta) \geq \sum_\theta \mu(\theta \mid hs) u(x; \theta) \qquad \text{for every } x \in f(B(h)) \tag{8}$$

whenever $P(hs) > 0$: after observing hs, the agent can commit to a report sequence that obtains x.

We now specify allocations for all report sequences, including null sequences. For $r \in S^T$, set

$$C(r) := \bigcap_{\substack{0 \leq t \leq T,\ t < \infty \\ P(\hat{s}^t) > 0}} f(B(\hat{s}^t)),$$

where $\hat{s}^0$ is the empty history and $B(\hat{s}^0) = \Theta$. The sets in this intersection are nested, finite, and nonempty, so $C(r)$ is nonempty. Fix an ordering of X_f, and define

$$a^*(r) = \begin{cases} a(r), & a(r) \in C(r), \\ \text{the first element of } C(r), & a(r) \notin C(r). \end{cases}$$

Under truthful reporting, $a(r) = f(\theta)$ almost surely and the true state belongs to the support at every finite prefix almost surely. Thus $a^* = a$ almost surely under each state's truthful distribution. For any positive-probability prefix h, every report continuation of h is assigned an outcome in $f(B(h))$. Conversely, each outcome in $f(B(h))$ is assigned on some truthful continuation of h. The available menu after h is therefore exactly $f(B(h))$.

Keep the original signal kernels whenever previous signals and reports coincide. After their first disagreement, send only fixed, state-independent signals. Truthful signals and allocations are unchanged. Conditional on the history hs at the first false report, the agent learns nothing further and his eventual outcome belongs to $f(B(h))$. His continuation payoff is therefore bounded above by the truthful continuation payoff by (8). Decomposing any deviation according to its first false report proves incentive compatibility. This argument also covers $T = \infty$, because two distinct report and signal sequences have a first finite disagreement. On paths with no disagreement the original truthful allocation is preserved. □

Proof of Lemma 3. Necessity follows because, after observing s^{t+1}, the agent can obtain any $x \in A(s^t)$ by fixing a continuation of his reports. For sufficiency, consider any reporting strategy and condition on the truthful signal history at its first false report. All subsequent information is uninformative about the state, and every subsequent allocation belongs to the menu available just before that report. The stated inequalities bound the payoff from every such allocation, and hence from any lottery over them, by the truthful continuation payoff. Averaging over the first false report establishes the result. □

Proof of Theorem 1 and Proposition 3. For a belief q, write

$$V(q) := \sum_{\theta} q(\theta)u(f(\theta);\theta).$$

We prove the claim by induction on the number of distinct target outcomes. A constant target requires no information or report; if the definition requires $T \geq 1$, use one uninformative round instead.

Suppose the claim holds whenever the target has fewer than m distinct outcomes. Consider an implementing mechanism at a prior q^0 with support B and $A = f(B)$, where $|A| = m \geq 2$. By Lemma 2, take the mechanism to have canonical menus and to withhold information after the first false report. Let (q_t) be its truthful posterior process, with $q_0 = q^0$, and let q_T denote the terminal posterior, including the almost-sure martingale limit when $T = \infty$. Set

$$\tau := \inf\{t \geq 1 : f(\operatorname{Supp} q_t) \subsetneq A\}, \qquad Q := \begin{cases} q_\tau, & \tau < \infty, \\ q_T, & \tau = \infty. \end{cases}$$

Before τ, the canonical menu is A. Hence Lemma 3 gives

$$V(Q) \geq \sum_{\theta} Q(\theta)u(x;\theta) \qquad \text{for every } x \in A,$$

where on $\{\tau = \infty\}$ the inequality follows by taking limits along the bounded posterior martingale. Bounded optional sampling also gives $\mathbb{E}[Q] = q^0$.

If $\tau < \infty$, the original continuation implements f under Q, and its target has fewer than m outcomes. If $\tau = \infty$, Lemma 1 implies that Q is supported on one equivalence class of f, so the restricted target is constant. Thus Q almost surely belongs to

$$D := \left\{ q \in \Delta(B) : \begin{array}{l} |f(\operatorname{Supp} q)| < m, \quad f \text{ is implementable at } q, \\ V(q) \geq \sum_{\theta} q(\theta)u(x;\theta) \text{ for every } x \in A \end{array} \right\}.$$

The finite-dimensional barycenter lemma below implies $q^0 \in \operatorname{co} D$. Carathéodory's theorem therefore gives $q^1, \ldots, q^K \in D$ and positive weights $\lambda_1, \ldots, \lambda_K$, with $K \leq |B|$, such that

$$q^0 = \sum_{j=1}^{K} \lambda_j q^j.$$

Send first-round signal j with probability

$$\pi_1(j \mid \theta) = \frac{\lambda_j q^j(\theta)}{q^0(\theta)}, \qquad \theta \in B.$$

After a truthful report j, use the implementing continuation provided by the induction hypothesis at q^j. After any false report, provide no further information. Each continuation can be taken to have canonical menus, and unused report labels can be assigned to existing branches. The first-round menu is A, while after report j it is $f(\operatorname{Supp} q^j)$. Membership in D gives the first-round IC constraints; the continuation constraints hold by induction. Lemma 3 therefore establishes incentive compatibility.

Every first-round posterior has a strictly smaller outcome menu. The same is true at every nonterminal continuation node by induction. A path has at most $m - 1$ active rounds. At a node with state support C, the construction uses at most $|C|$ signals. Since each nonterminal transition also strictly reduces state support, the number of signals at round t is at most $n - t + 1$. Labels may be reused across histories, and terminated branches may be padded with uninformative signals. This proves Theorem 1 and gives $T \leq |X_f| - 1$ for $|X_f| \geq 2$. □

Finite-dimensional barycenter lemma. If a bounded random vector Z takes values almost surely in $D \subseteq \hat{s}^d$, then $\mathbb{E}[Z] \in \operatorname{co} D$. Indeed, the mean belongs to the closure of $\operatorname{co} D$. If it does not belong to $\operatorname{co} D$, it lies on the relative boundary of that closure. A supporting hyperplane through the mean then contains Z almost surely: the supporting linear functional is weakly on one side of its value at the mean and has that value in expectation. Restricting D to this proper affine hyperplane reduces its affine dimension. Induction on affine dimension yields the claim. □

Proof of Corollary 2: The displayed mechanisms establish sufficiency. For necessity, Theorem 1 permits two rounds. Every first-round posterior therefore has support of size at most two. A two-state continuation must be statically implementable, so a pair can be used only if neither state envies the other. A singleton first-round posterior can be used only if its state envies no outcome in the initial menu.

In case 2, every first-round posterior assigning positive mass to state 3 must have support $\{1, 3\}$. Sum its IC constraint against $f(2)$, weighted by the probability of the signal. All the mass μ_3 of state 3 appears in these posteriors, whereas they use at most μ_1 mass of state 1. Since $u_{11} - u_{21} \geq 0$, the stated inequality follows.

In case 3, all mass of state 2 must appear in posteriors supported on $\{1, 2\}$, and all mass of state 3 in posteriors supported on $\{1, 3\}$. Let β_2 and β_3 be the respective masses of state 1 used by those posteriors. Then $\beta_2 + \beta_3 \leq \mu_1$. Summing their IC constraints against $f(3)$ and $f(2)$ gives the two inequalities in the statement with β_2 and β_3 in place of α and $\mu_1 - \alpha$. Any unused mass of state 1 can be added to either branch, since the loss coefficients are nonnegative. This yields a feasible α.

It remains to exclude the other envy patterns. A state that envies both other states cannot belong to an admissible pair or singleton. If both other states envy one state j, the initial IC constraint against $f(j)$ permits only the posterior concentrated on j, which cannot average to a full-support prior. After excluding these patterns, any remaining graph with more than one edge that is not a single reciprocal pair contains a state for which both possible pairs contain an envy edge, and that state itself has an outgoing edge. It can belong to neither an admissible pair nor a singleton. This exhausts the directed graphs on three vertices and proves the final case. □

Proof of Proposition 4. A singleton state space is implemented by assigning its target outcome, which gives the stated formula for Γ^1. Suppose the characterization is known through $k-1$.

For necessity, let $\mu \in \Gamma^k$ and $B = \operatorname{Supp}\mu$. If f is constant on B, every δ_θ, $\theta \in B$, belongs to $IC(B) \cap \Gamma^{k-1}$, and $\mu = \sum_{\theta \in B} \mu(\theta)\delta_\theta$. Otherwise, apply Theorem 1. Its first-round posteriors have strictly smaller outcome menus, hence proper state supports contained in B. Their continuations implement f, so these posteriors belong to Γ^{k-1}. Lemma 3 puts them in $IC(B)$, and Bayes plausibility expresses μ as their convex combination.

For sufficiency, suppose

$$\mu = \sum_{j=1}^{K} \lambda_j q^j, \qquad q^j \in IC(B) \cap \Gamma^{k-1}, \quad \lambda_j > 0, \quad \sum_j \lambda_j = 1,$$

for some $|B| \leq k$. All q^j are supported on $\operatorname{Supp}\mu$. For each θ in that support, send signal j with probability

$$\pi_1(j \mid \theta) = \frac{\lambda_j q^j(\theta)}{\mu(\theta)}.$$

The probabilities sum to one and induce posterior q^j after signal j. For each report j, append a canonical implementing mechanism for q^j, which exists by the induction hypothesis and Lemma 2. Withhold all further information after a false report. The menu before the first report is $f(\operatorname{Supp}\mu) \subseteq f(B)$. Membership of q^j in $IC(B)$ therefore gives all first-round IC constraints. The continuation IC constraints hold by construction. Lemma 3 shows that the combined mechanism is incentive compatible, and truthful reporting produces f in every state in $\operatorname{Supp}\mu$. Thus $\mu \in \Gamma^k$. □

Proof of Proposition 5: For sufficiency, discard zero-mass histories. At an on-path history h, choose signal j with state-dependent probability

$$\pi(j \mid \theta, h) = \frac{q_{hj}(\theta)}{q_h(\theta)} \qquad \text{whenever } q_h(\theta) > 0.$$

Flow conservation makes these probabilities sum to one. For state-history pairs with $q_h(\theta) = 0$, specify the signal law arbitrarily. Following a first false report, provide no further information. A terminal reported history selects the outcome of its unique remaining state; reports outside the positive-mass continuation tree are assigned a fixed outcome from the current continuation menu. On the truthful path, the joint state-history probabilities are q_h, and terminal allocations equal $f(\theta)$. A deviation after receiving j can yield only outcomes in $f(B_h)$. By (7), every such outcome is weakly worse, at that posterior, than truthful continuation. The simplified IC lemma therefore implies incentive compatibility.

For necessity, take a finite implementing tree with strictly decreasing posterior supports, as provided by the support reduction result. If a terminal posterior contains several states, the SCF must be constant on its support; append full revelation after fixing its allocation. If an edge eliminates several states, label that edge by one eliminated state and insert deterministic edges that eliminate the remaining states one at a time. Each inserted node has the same posterior as the child of the original edge. Its menu is contained in the original parent menu, so the IC inequalities for the original edge imply those for the inserted edges.

Every path now has an ordered list of eliminated states and ends after $n-1$ eliminations. Different nodes can have the same list. For each h, sum their joint state-history probability vectors and call the result q_h. These sums satisfy (4)–(6). All nodes with the same list have the same designated remaining set B_h. Because the edge IC inequalities are homogeneous and linear in the joint probabilities, their sums satisfy (7). □

Lemma 5 *Let D be a real $m \times n$ matrix. There is no vector $z \gg 0$ satisfying $Dz \leq 0$ if and only if there exists $q \geq 0$ such that $D^{\mathsf{T}} q \geq 0$ and $D^{\mathsf{T}} q \neq 0$.*

Proof: Let $\mathbf{1}$ denote the vector of ones in $\hat{s}^n$. By positive rescaling, $Dz \leq 0$ has a solution $z \gg 0$ if and only if $D(y+\mathbf{1}) \leq 0$ has a solution $y \geq 0$. By Farkas' lemma, the latter system is infeasible if and only if there exists $q \geq 0$ with

$$D^{\mathsf{T}} q \geq 0, \qquad q^{\mathsf{T}}(-D\mathbf{1}) < 0.$$

Given the first inequality, the second is equivalent to $D^{\mathsf{T}} q \neq 0$. □

Proof of Proposition 6: Index the rows by $x \in X_f$ and the columns by $\theta \in \Theta$, and define the matrix

$$D_{x\theta} = u(x;\theta) - u(f(\theta);\theta).$$

Then $IC(\Theta)$ contains a full-support belief if and only if $Dz \le 0$ has a solution $z \gg 0$: normalize such a vector z to have components summing to one. By Lemma 5, no such belief exists if and only if there is $q \ge 0$ with $D^\mathsf{T} q \ge 0$ and $D^\mathsf{T} q \ne 0$. Necessarily $q \ne 0$. Set $p(x) = q(x)/\sum_{y \in X_f} q(y)$ to obtain

$$\sum_{x \in X_f} p(x) u(x;\theta) - u(f(\theta);\theta) \ge 0$$

in every state, with strict inequality in at least one state. The converse follows by reversing the argument. □

Proof of Proposition 8: One direction follows from the definition. For the converse, suppose some utility representation u inducing $E^{\preceq}$ and some full-support prior admit an implementing mechanism. By Lemma 2, its allocations can be restricted to X_f.

For each state θ, leave the own-outcome value $b_\theta = u(f(\theta);\theta)$ unchanged. Assign the outcomes ranked above $f(\theta)$ under $\preceq$ distinct values in $(b_\theta, b_\theta + \varepsilon_\theta)$, in the required order, where ε_θ is smaller than every original positive gain $u(x;\theta) - b_\theta$. Assign the outcomes below $f(\theta)$ values smaller than all their original values, again in the required order. Empty groups need no assignment. The resulting utility $\widetilde{u}$ represents $\preceq$ and satisfies

$$\widetilde{u}(f(\theta);\theta) = u(f(\theta);\theta), \qquad \widetilde{u}(x;\theta) \le u(x;\theta) \quad (x \in X_f).$$

Thus truthful utility is unchanged in every state, while the payoff of every deviation weakly decreases. The same mechanism therefore implements f under $\widetilde{u}$. □

Proof of Proposition 9: The singleton case is immediate. We first prove necessity. By Theorem 1, an implementing mechanism can be chosen so that its first round has finitely many positive-probability signals, with posteriors μ_j supported on nonempty proper subsets $B_j \subsetneq \Theta$. Their supports cover Θ, and each continuation implements $f|_{B_j}$. For every $\theta' \notin B_j$, incentive compatibility gives

$$\sum_{\theta \in B_j} \mu_j(\theta)\big[u(f(\theta');\theta) - u(f(\theta);\theta)\big] \le 0.$$

Since μ_j has full support on B_j and comparisons are strict, some $\theta \in B_j$ must satisfy $\neg(\theta E \theta')$.

Take an inclusion-minimal subcollection of the sets B_j that covers Θ and partition this subcollection into two nonempty groups. Let Θ_1 and Θ_2 be the unions in the respective groups. Minimality implies that both unions are proper: each member of a minimal cover contains a state absent from all other members. Each union is possibly implementable. Indeed, mix the original posteriors in its group with positive weights and use their original continuations. The resulting prior has full support on the union, and the original incentive constraints remain valid when the menu is restricted to that union. Finally, if $\theta' \notin \Theta_i$, every original member $B_j \subseteq \Theta_i$ supplies a state that does not envy θ'. This proves the three conditions.

For sufficiency, recursively choose the two covers in the statement, until every terminal set is a singleton. Such a finite tree exists because every child is a proper subset of its parent. Write A_h for the set at node h. Fix an arbitrary full-support prior μ. At every nonterminal node, conditional on state θ, send each child signal whose set contains θ with equal probability. The posterior μ_h at every node then has full support on A_h.

We claim that for every child h of a node with set A and every $\theta' \in A$, either $A_h = \{\theta'\}$ or there is a $\theta \in A_h \setminus \{\theta'\}$ with $\neg(\theta E \theta')$. If $\theta' \notin A_h$, this is condition 3. Otherwise choose any different state in A_h and follow a descendant path ending at its singleton. At the first edge along this path that excludes θ', condition 3 supplies the required state.

Now fix any strict preference ordering on X_f inducing E. Choose a utility representation with $u(f(\theta); \theta) = 1$, placing every outcome ranked below $f(\theta)$ in $[0, 1/2]$ and every outcome ranked above it in $(1, 1 + \varepsilon]$, while respecting all rankings within each group. These are choices of a representation for the existence proof, not a normalization imposed on the model. Because the tree is finite, the number

$$m := \min\{\mu_h(\theta) : \theta \in A_h\}$$

is positive. Choose $0 < \varepsilon < m/2$. For every child h and every θ' in its parent's set, the claim implies

$$\sum_{\theta \in A_h} \mu_h(\theta) \big[u(f(\theta'); \theta) - u(f(\theta); \theta) \big] \leq \varepsilon - \frac{m}{2} < 0,$$

unless $A_h = \{\theta'\}$, in which case the expression is zero. Thus all fixed-outcome incentive constraints hold simultaneously for one utility representation, including in overlapping branches. Allocate $f(\theta)$ at singleton leaves and withhold further information after the first false report. Lemmas 2 and 3 then establish incentive compatibility of the constructed mechanism. This also proves the final claim for every prescribed full-support prior and strict ordering. □

Proof of Corollary 4: For the first claim, the initial IC inequality for outcome $f(j)$ can hold at a posterior only if that posterior assigns probability one to j: every other state strictly prefers $f(j)$ to its prescribed allocation. Such posteriors cannot average to a full-support prior.

For the second claim, use a finite tree with strictly decreasing supports. Consider a positive-probability path that ends in state i, and the edge on which i is first fully revealed. Its parent support contains another state j. At the resulting degenerate posterior, outcome $f(j)$ is strictly preferred to $f(i)$, violating the edge IC inequality.

For the last claim, let the vertices be divided into the connected components of the underlying undirected graph. Consider all pairs $\{i, j\}$ whose vertices lie in different components. Choose a first-round experiment whose positive-probability posteriors have exactly these pairs as supports and whose average is the given prior. Such an experiment exists: in each state, assign positive probability to every pair containing that state. After each first-round report, reveal which of the two states occurred and restrict allocations to their two prescribed outcomes. No state envies a state in another component, so this last-round allocation is incentive compatible.

Choose a utility representation of the prescribed strict ordering in which every gain from an envied outcome lies in $(0, \varepsilon)$ and every loss from an outcome below the prescribed outcome has magnitude at least one. At every first-round pair posterior and for every alternative desired outcome, one of the two states strictly loses from that outcome; the other either weakly loses or gains at most ε. There are finitely many such posteriors and inequalities, and each posterior gives both its states positive probability. Hence a sufficiently small $\varepsilon > 0$ makes all first-round IC inequalities hold. Withholding information after deviations completes the implementing mechanism. □

Proof of Proposition 11: Construct a report-only mechanism on the specified tree. At a node A with children A_1, A_2, conditional on a state $\theta \in A$, send each child signal whose set contains θ with equal probability. Conditional on any state $\theta \notin A$, send each child signal with probability $1/2$. Continue at the child that the agent reports, and allocate $f(\theta)$ at a singleton leaf $\{\theta\}$. This signal structure depends only on the state and past reports. Truthful reporting reaches the correct singleton in every state.

Consider a pure strategy whose first false report occurs with positive probability. At that history, let A_i be the received child signal and A_{3-i} the reported child. The agent's posterior has full support on A_i. Every state in $D = A_i \setminus A_{3-i}$ lies outside every subsequent reported node. Consequently, all states in D induce the same law of subsequent signals and outcomes under the deviating continuation. Choose any terminal outcome $f(\theta')$ that this continuation reaches with positive probability under that law. Since $\theta' \in A_{3-i}$, the displayed condition supplies $\theta \in D$ that does not envy θ'. The deviation therefore assigns a strictly worse outcome with positive probability in state θ.

The tree, signal spaces and report spaces are finite. Hence the minimum r of the strictly positive ex-ante probabilities of receiving a worse outcome over all pure strategies with a positive-probability first false report is positive. If there are no such strategies the result is immediate. Choose a representation of the prescribed ordering with own-outcome utility 1, below-own utilities in $[0, 1/2]$, and above-own utilities in $(1, 1+\varepsilon]$, where $0 < \varepsilon < r/2$. Every such deviation has expected utility at most $1 - r/2 + \varepsilon < 1$, whereas truthful reporting yields 1. Strategies that never falsely report on a positive-probability history produce the truthful outcome, and randomization cannot improve on pure strategies. Thus the mechanism implements f. □