The B.E. Journal of Theoretical Economics

Advances

Volume 11, *Issue* 1	2011	*Article* 20

Strategy-Proof Compromises

Peter Postl*

*University of Birmingham, p.postl@bham.ac.uk

Recommended Citation
Peter Postl (2011) "Strategy-Proof Compromises," *The B.E. Journal of Theoretical Economics*: Vol. 11: Iss. 1 (Advances), Article 20.

Strategy-Proof Compromises*

Peter Postl

Abstract

We study strategy-proof decision rules in the variant of the canonical public good model proposed by Borgers and Postl (2009). In this setup, we fully characterize the set of budget-balanced strategy-proof deterministic mechanisms, which are simple threshold rules. For smooth probabilistic mechanisms, we provide a necessary and sufficient condition for dominant strategy implementation. When allowing for discontinuities in the mechanism, our necessary condition remains valid, but additional conditions must hold for sufficiency. We also show that, among ex post efficient decision rules, only dictatorial ones are strategy-proof. While familiar in spirit, this result is not the consequence of any known result in the literature.

KEYWORDS: compromise, public good provision, dominant strategy implementation, strategy-proof, dictatorship

*I am indebted to Siddhartha Bandyopadhyay, Tilman Borgers, Martin Jensen, Indrajit Ray, Jaideep Roy, Arunava Sen and Rakesh Vohra.

1 Introduction

Since the "Wilson doctrine" (Wilson, 1987), much emphasis has been placed on the design of "detail free" mechanisms that do not rely excessively on common knowledge assumptions about the environment for which they are intended. In particular, the main objective is to avoid the assumption that agents' beliefs about each other are common knowledge. As one way of responding to Wilson's critique, the literature has revisited canonical mechanism design problems such as auctions, bilateral trade, and the provision of a public good (see, e.g., Börgers, 2006, Chung and Ely, 2006, Mookherjee and Reichelstein, 1992), offering characterizations of either ex post incentive compatible or dominant strategy incentive compatible mechanisms, for which assumptions about the agents' beliefs are superfluous.[1]

The present paper contributes to this literature by focusing on dominant strategy implementation in the compromise model of Börgers and Postl (2009). The question of what shape strategy-proof decision rules take in this model is of broader interest because of its close formal connection with canonical models of mechanism design for public good provision. While formally without transferable utility, the compromise setting is essentially a public good model with two agents who "pay" for a given probability of obtaining a compromise outcome (i.e. the public good) by surrendering probability of their respective favorite alternatives. The fact that the transferable resource is probability mass gives rise to individual liquidity constraints that impose additional restrictions on the agents' payments towards the public good over and above those needed to ensure "budget balance".[2] A second difference between the public good model and the compromise model is that in the latter, agents cannot opt out of the mechanism.

In the literature, existing characterizations of strategy-proof budget-balanced public good provision mechanisms with voluntary participation have so far been restricted to deterministic mechanisms that specify, on the basis of the participants' stated preferences, whether or not the public good will be provided (see Chung and Ely, 2006, who allow for the possibility of interdependent valuations, and Börgers, 2006, who provides a characterization for two agents with private valuations). In the compromise setting, where a mechanism is a cardinal probabilistic decision rule, a restriction to deterministic mechanisms seems less natural than in the public good setup. By studying strategy-proof decision rules in the compromise

[1]There are alternative responses to the Wilson doctrine that do not involve a strengthening of the implementation concept as a way of dispensing with the need for specifying agents' beliefs about each other. Examples include Bergemann and Morris (2005), Chung and Ely (2007), and Smith (2010).

[2]Budget balance in the compromise model refers to the requirement that the probabilities assigned by the mechanism to the various possible outcomes must sum up to 1.

model, we would like to make a first step towards a characterization of strategy-proof budget-balanced and ex post individually rational *probabilistic* public good provision mechanisms.

We make two main contributions in this paper. The first is a characterization of the set of strategy-proof decision rules in the compromise model. While this characterization builds on what are now standard methods in mechanism design theory (such as the integral form envelope theorem of Milgrom and Segal, 2002), we go beyond the established payoff equivalence result (see Holmström's Lemma in Milgrom, 2004) by explicitly accounting for the fact that decision rules are probability distributions over the set of social alternatives. In terms of the public good analogy, this means that we incorporate the ex post budget balance constraints directly into the characterization of strategy-proof public good provision mechanisms. What is particularly interesting is that any strategy-proof mechanism involves "payments" by the agents that automatically satisfy the individual liquidity constraints. In other words, with dominant strategy implementation, the key distinguishing feature of the compromise model (i.e. the agents' liquidity constraints) ceases to play a role, and consequently the set of strategy-proof decision rules is identical to the set of budget-balanced provision mechanisms in the corresponding public good model.

The key component of any strategy-proof decision rule is the function which, based on the agents' stated preferences, determines the probability of the compromise. The analogue of this function in the public good model is the so called public good provision rule, which determines the probability of public good provision. Our characterization of strategy-proof mechanisms in the compromise model implies restrictions on the shape of admissible public good provision rules in the public good model. These restrictions go further than the monotonicity requirement that usually characterizes strategy-proof mechanisms in the literature (see, e.g., Mookherjee and Reichelstein, 1992). However, providing a sharp characterization of all admissible provision rules remains a difficult problem in general. For the class of deterministic mechanisms, we are able to provide a complete description of the set of strategy-proof and budget-balanced provision rules. For probabilistic mechanisms, however, it is decidedly more difficult to obtain such a characterization. For the class of twice continuously differentiable provision rules, we show that a necessary and sufficient condition for strategy-proofness is additivity. I.e. the provision rule is the sum of two functions, each of which depends solely on the valuation of one agent.[3] In order to move beyond continuously differentiable provision rules, we consider a

[3]Our work in this part of the paper is related to the small literature on strategy-proof cardinal probabilistic decision rules due to Gibbard (1977), Freixas (1984), and Barberà et al. (1998). I am indebted to Arunava Sen for drawing my attention to this literature. Note that all these papers predate the development of envelope theorems for functions that are not everywhere continuously differentiable (Milgrom and Segal, 2002). The approach in Freixas (1984), when applied to our

DOI: 10.2202/1935-1704.1684

special class of piecewise differentiable rules. While (piecewise) additivity remains necessary, we show that sufficiency requires any admissible provision rule to satisfy additional conditions that restrict the rule's behavior at points of discontinuity.

Our second main contribution in this paper is to show that an ex post efficient decision rule (as defined by Holmström and Myerson, 1983) is strategy-proof if and only if it always selects the favorite alternative of the same agent. Against the backdrop of "dictatorship" results in the classical literature on dominant strategy implementation of efficient decision rules (see, e.g., Gibbard, 1973, Satterthwaite, 1975, or Aswal et al., 2003), this result may not sound surprising. However, it is important to emphasize that our result is neither a consequence of this literature (which considers only deterministic social choice rules), nor of the literature on strategy-proof implementation of *probabilistic* social choice rules.[4] Despite the impossibility result in Börgers and Postl, which shows that no ex ante incentive efficient decision rules exist under Bayesian implementation, it is not clear a priori that there are no strategy-proof ex post efficient decision rules here. The reason is that while strategy-proofness on one hand is a more restrictive implementation concept than Bayesian incentive compatibility, ex post efficiency on the other hand is a much weaker efficiency notion than ex ante efficiency.[5] It should therefore come as no surprise that any given strategy-proof deterministic mechanism can be rendered ex post *constrained* efficient, as we illustrate with examples at the end of the paper.

The remainder of this paper is structured as follows: In Section 2, we introduce the model and basic definitions. Section 3 contains our characterization of the set of strategy-proof decision rules. In Section 4, we characterize (in the terminology of the public good model) all deterministic strategy-proof and budget-balanced public good provision rules. Section 5 contains our results for dominant strategy implementation of probabilistic mechanisms. In Section 6, we study ex post efficient decision rules. Section 7 offers a brief conclusion. The appendix in Section 8 contains all longer proofs.

---

setup, rules out all but constant mechanisms. The focus in Barberà et al. (1998) is on continuous and on twice continuously differentiable rules.

[4]In Gibbard (1977), preference intensities are not considered. While studying cardinal decision rules that take account of preference intensities, Dutta et al. (2007) assume a universal preference domain. The compromise model, in contrast, features a very restrictive preference domain with only one ordinal ranking per agent.

[5]In fact, Wilson (1993) points out: "Ex post efficiency is rarely invoked because it is a very weak criterion [...]."

2 The Model

Two agents $i \in I := \{1,2\}$ must choose one alternative from the set $A := \{a_0, a_1, a_2\}$. Each agent i prefers alternative a_i over alternative a_0, and alternative a_0 over alternative a_{-i} (subscript $-i$ refers to the agent other than i). These ordinal preferences are common knowledge. We refer to alternative a_0 as the compromise because it is the middle-ranked alternative of both agents. Agent i's von Neumann Morgenstern utility function is $u_i : A \to \mathbb{R}$. Utilities are normalized so that $u_i(a_i) = 1$ and $u_i(a_{-i}) = 0$ for all $i \in I$. These aspects of the von Neumann Morgenstern utility functions are common knowledge. For each agent $i \in I$ denote by t_i the utility of the compromise $u_i(a_0)$. We refer to $t_i \in [0,1]$ as agent i's type. Each agent observes his own type, but not that of the other agent. Define by $t := (t_1, t_2)$ a generic type-pair in $T := [0,1]^2$.

Definition 1 *A **decision rule** is a function $f : T \to \Delta(A)$, where $\Delta(A)$ is the set of all probability distributions over A.*

Denote by $f_i(t)$ the probability that decision rule f assigns to agent i's favorite alternative when the type-pair is t, and let $f_0(t)$ denote the probability that f assigns to the compromise. For a decision rule f and type-pair t, agent i's expected utility is $u_i(t|t_i) := f_i(t) + f_0(t)t_i$. As the agents' types are privately observed, only incentive compatible decision rules can be implemented. We focus here on implementation of truth-telling in dominant strategies.

Definition 2 *A decision rule f is **strategy-proof** if for all $i \in I$, all $t_i, t_i' \in [0,1]$, and all $t_{-i} \in [0,1]$:*

$$f_i(t_i, t_{-i}) + f_0(t_i, t_{-i})t_i \geq f_i(t_i', t_{-i}) + f_0(t_i', t_{-i})t_i.$$

As shown in Börgers and Postl (2009), the above compromise model can be re-interpreted as a model of mechanism design for the provision of a public good. By introducing a *default outcome* in which agent i's favorite alternative is selected with probability $\delta_i \in [0,1]$ (for each $i \in I$, with $\delta_1 + \delta_2 = 1$), we can view the difference $\delta_i - f_i(t)$ as agent i's "payment" towards a public good (i.e. the compromise) when the type-pair is t. The definition of decision rules above implies individual liquidity constraints for the agents, which arise because the probability of each agent's favorite alternative is a number between 0 and 1. Agent i's "payment" towards the public good must therefore be a number in $[\delta_i - 1, \delta_i]$ for all $t \in T$.

It is customary in the public goods context to assume that agents are free to opt out of any proposed mechanism. If we allow for this in the modified compromise setup, then the following individual rationality constraints have to be taken into account:

Definition 3 *A decision rule is **ex post individually rational** if:*

$$f_i(t) + f_0(t)t_i \geq \delta_i \text{ for all } t \in T \text{ and all } i \in I.$$

3 Strategy-Proof Decision Rules

In this section, we investigate the structure of the set of strategy-proof decision rules. Lemma 1 below adapts to our setting the characterization of strategy-proofness that is familiar from the mechanism design literature on quasilinear environments with transferable utility.[6]

Lemma 1 *A decision rule f is **strategy-proof** if and only if:*

(i) For all $i \in I$ and all $t \in T$: $f_0(t)$ is nondecreasing in t_i.

(ii) For every $i \in I$ and every $t \in T$:

$$f_i(t) = f_i(1, t_{-i}) + f_0(1, t_{-i}) - f_0(t)t_i - \int_{t_i}^{1} f_0(s, t_{-i})ds.$$

Lemma 1 highlights the central role played by the probability of the compromise in the characterization of strategy-proof decision rules. In particular, item (ii) of Lemma 1 tells us that the probability of an agent's favorite alternative (and therefore his "payment") is determined by the probability of the compromise, up to an additive term $f_i(1, t_{-i}) + f_0(1, t_{-i})$ that is independent of the agent's own type and therefore does not affect his incentives.

Observe that Lemma 1 is derived without making use of the fact that decision rules, as introduced in Definition 1, are probability distributions whose components sum up to 1 for all type-pairs. Lemma 1 is therefore of limited use when it

[6]The proof of Lemma 1 is familiar from the literature and therefore omitted. Item (ii) of Lemma 1, for instance, follows directly from the integral form envelope theorem (see, e.g., Theorem 3.1 in Chapter 3.2 of Milgrom, 2004), which establishes that an agent's utility under a strategy-proof decision rule is differentiable almost everywhere with respect to the agent's own type, and provides furthermore an expression for this partial derivative. Taking as given the agent's utility at the highest type (rather than the lowest type, as is customary in the literature), we obtain the expression in item (ii) of Lemma 1.

comes to constructing strategy-proof decision rules because it does not sufficiently restrict the class of functions f_0 that can be part of a strategy-proof rule. The following example illustrates this point:

Example 1 *Consider the function f_0 with $f_0(t) = 1$ if $t_1 + t_2 > 1$, and $f_0(t) = 0$ if $t_1 + t_2 < 1$. This function is nondecreasing as required by item (i) of Lemma 1, but cannot be part of a strategy-proof decision rule. To see this, note that by item (ii) of Lemma 1, the probability of each agent's favorite alternative is:*[7]

$$f_i(t) = f_i(1, t_{-i}) + (1 - f_0(t))(1 - t_{-i}) \quad \forall i \in I. \tag{1}$$

For any type-pair t such that (s.t.) $t_1 + t_2 > 1$ we have $f_0(t) = 1$, and therefore $f_i(t) = 0$ for all $i \in I$. Thus, by equation (1): $f_i(1, t_{-i}) = 0$ for all $t_{-i} \in [0,1]$. This, however, leads to a contradiction: for any t s.t. $t_1 + t_2 < 1$ we obtain $f_0(t) + f_1(t) + f_2(t) = 2 - t_1 - t_2 > 1$.[8]

Example 1 highlights the need for explicit restrictions on functions f_0 that can be part of a strategy-proof decision rule. In order to derive such restrictions, we now account explicitly for the requirement that the functions f_0, f_1 and f_2 that together constitute a decision rule, must sum up to 1 for all type-pairs. This yields the following characterization of strategy-proof decision rules that goes beyond existing characterizations of strategy-proof mechanisms in the literature (such as the one for the canonical public good model in Börgers, 2006).

Proposition 1 *Given a function $f_0 : T \to [0,1]$ and constants $f_1(1,1)$, $f_2(1,1) \in [0,1]$ s.t. $f_1(1,1) + f_2(1,1) = 1 - f_0(1,1)$, there exist functions $f_i : T \to [0,1]$ ($\forall i \in I$) s.t. (f_0, f_1, f_2) is a **strategy-proof** decision rule if and only if:*

(i) For all $i \in I$ and all $t \in T$: $f_0(t)$ is nondecreasing in t_i.

(ii) For all $i \in I$ and all $t \in T$:

$$f_i(t) = f_i(1,1) + f_0(1, t_{-i})t_{-i} + \int_{t_{-i}}^{1} f_0(1,s)ds - f_0(t)t_i - \int_{t_i}^{1} f_0(s, t_{-i})ds,$$

[7] The two expressions for f_i obtained from item (ii) in Lemma 1, one for type-pairs t such that $t_1 + t_2 < 1$, and the other for type-pairs t such that $t_1 + t_2 > 1$, can easily be gathered into the single expression in (1).

[8] The function f_0 in Example 1, coupled with functions f_i given by (1) with $f_i(1, t_{-i}) = 0$ for all $t_{-i} \in [0,1]$ and all $i \in I$, constitutes the Vickrey-Clarke-Groves mechanism used in the proof of the main impossibility result in Börgers and Postl (2009).

(iii) For all $t \in T$:

$$[f_0(t_1,1) - f_0(t)]t_1 + [f_0(1,t_2) - f_0(t)]t_2 - [f_0(1,1) - f_0(t)] + \int_{t_1}^{1} [f_0(s,1) - f_0(s,t_2)]ds + \int_{t_2}^{1} [f_0(1,s) - f_0(t_1,s)]ds = 0.$$

The proof of Proposition 1 is in the Appendix. If the compromise model is given the public good interpretation mentioned in Section 2, then Proposition 1 provides a characterization of strategy-proof public good provision mechanisms that are ex post budget-balanced and satisfy the agents' individual liquidity constraints.

In order to convey some insight into the derivation of Proposition 1, observe that the additive term $f_i(1,t_{-i}) + f_0(1,t_{-i})$ in item (ii) of Lemma 1 not only represents agent i's utility from the decision rule when he has the highest type 1, but also represents the probability that agent i's least preferred alternative is *not* chosen. We can therefore equivalently write the additive term $f_i(1,t_{-i}) + f_0(1,t_{-i})$ as $1 - f_{-i}(1,t_{-i})$. Noting that $f_{-i}(1,t_{-i})$ is itself determined by item (ii) of Lemma 1, we obtain the expressions for the probability of each agent's favorite alternative given in item (ii) of Proposition 1.[9] The requirement that the functions f_1 and f_2 in item (ii), together with the function f_0, must sum up to one for all type-pairs then yields item (iii) of Proposition 1. It is this item that furnishes the desired restriction on functions f_0 that can be part of a strategy-proof decision rule.

While item (iii) of Proposition 1 allows us to check if a given nondecreasing function f_0 can be part of a strategy-proof decision rule, it would be useful to know the full class of functions f_0 for which this is the case. More precisely, we would like to have a characterization of all nondecreasing functions f_0 that satisfy item (iii) of Proposition 1. Henceforth, we shall call such functions admissible:

Definition 4 *A nondecreasing function $f_0 : T \to [0,1]$ is **admissible** if it satisfies item (iii) of Proposition 1.*

While it is difficult in general to obtain a full characterization of all admissible functions f_0, it is easy to see that the following condition is *sufficient*:

Proposition 2 *If f_0 is an **additive function** of the form $f_0(t) = f_0^1(t_1) + f_0^2(t_2)$, where $\forall i \in I$, $f_0^i : [0,1] \to [0,1]$ is a nondecreasing function, then f_0 is admissible.*

[9] The proof of Proposition 1 also shows that the probabilities of the agents' favorite alternatives in item (ii) of Proposition 1 always take values in $[0,1]$.

Proof. The monotonicity of the functions f_0^i ($\forall i \in I$) implies that f_0 satisfies item (i) of Proposition 1. This, in turn, ensures that f_0 is Riemann integrable, and therefore the integrals in items (ii) and (iii) are well-defined. Additivity of f_0 implies for all $i \in I$ and all $t_i, t_i', t_{-i} \in [0,1]$: $f_0(t_i', t_{-i}) - f_0(t_i, t_{-i}) = f_0^i(t_i') - f_0^i(t_i)$. It is now straightforward to verify that item (iii) of Proposition 1 is satisfied. □

In order to explore which functions f_0 are admissible when we move beyond additivity, we study in the next section a specific class of functions that has received particular attention in the literature on public good provision mechanisms (see, e.g., Chung and Ely, 2006, and Börgers, 2006).

4 Binary Decision Rules

In this section we focus on binary decision rules where, conditional on the agents' types, the probability of the compromise is either 0 or 1. Binary decision rules in our setting correspond to deterministic provision mechanisms in the public good setting with quasilinear preferences and transferable utility.

Definition 5 *A decision rule f is a* ***binary decision rule*** *if $f_0 : T \to \{0,1\}$.*

In Proposition 3 below, we provide a full characterization of all admissible binary functions f_0. For each such function, we also state the probabilities f_1 and f_2 of the agents' favorite alternatives that render the binary decision rule (f_0, f_1, f_2) strategy-proof.

Proposition 3 *A* ***binary decision rule*** *is* ***strategy-proof*** *if and only if it belongs to one of the following three categories:*[10]

I. *Binary decision rules that depend on neither agent's type:*
   (i) $f_0(t) = 1$ *and* $f_1(t) = 0$ $\forall t \in T$.
   (ii) $f_0(t) = 0$ *and* $f_1(t) = a$ $\forall t \in T$ *and any* $a \in [0,1]$.

II. *Binary decision rules that depend on one agent's type:*
   There is an agent $i \in I$ and a threshold $\tau_i \in [0,1]$ s.t. $\forall t \in T$:

$$\begin{aligned} f_0(t) &= 0 \text{ and } f_i(t) = \tau_i \text{ if } t_i < \tau_i, \\ f_0(t) &= 1 \text{ and } f_i(t) = 0 \text{ if } t_i \geq \tau_i. \end{aligned}$$

[10] We list in each category, and for every $t \in T$, only the probabilities of two alternatives. The probability of the omitted alternative can be computed by subtracting from 1 the probabilities given in the proposition.

DOI: 10.2202/1935-1704.1684

III. Binary decision rules that depend on both agents' types:

(i) There are thresholds $\tau_1, \tau_2 \in [0,1]$ *with* $\tau_1 + \tau_2 = 1$ *s.t.* $\forall t \in T$*:*

$$\begin{aligned} f_0(t) &= 0 \text{ and } f_1(t) = \tau_1 \text{ if } t_1 < \tau_1 \wedge t_2 < \tau_2, \\ f_0(t) &= 1 \text{ and } f_1(t) = 0 \text{ otherwise.} \end{aligned}$$

(ii) There are thresholds $\tau_1, \tau_2 \in [0,1]$ *with* $\tau_1 + \tau_2 = 1$ *s.t.* $\forall t \in T$*:*

$$\begin{aligned} f_0(t) &= 1 \text{ and } f_1(t) = 0 \text{ if } t_1 \geq \tau_1 \wedge t_2 \geq \tau_2, \\ f_0(t) &= 0 \text{ and } f_1(t) = \tau_1 \text{ otherwise.} \end{aligned}$$

The proof of Proposition 3 is in the Appendix. It uses item (iii) of Proposition 1 extensively to characterize admissible binary functions f_0. The key step in the proof is to note that if there are thresholds $\tau_i', \tau_i'' \in (0,1)$ and types t_{-i}' and t_{-i}'' such that $f_0(t_i, t_{-i}') = 0$ if $t_i < \tau_i'$ and $f_0(t_i, t_{-i}') = 1$ if $t_i > \tau_i'$, and $f_0(t_i, t_{-i}'') = 0$ if $t_i < \tau_i''$ and $f_0(t_i, t_{-i}'') = 1$ if $t_i > \tau_i''$, then the thresholds must be the same: $\tau_i'' = \tau_i'$. To see this, observe that for t_{-i}' (t_{-i}'', resp.) the probability of i's favorite alternative must be equal to the threshold τ_i' (τ_i'', resp.) whenever the compromise is *not* chosen. This is to ensure that agent i with a type below the threshold has no incentive to pretend his type is above the threshold, and vice versa. Now consider a type $\hat{t}_i < \min\{\tau_i', \tau_i''\}$. We have $f_0(\hat{t}_i, t_{-i}') = 0$ and $f_{-i}(\hat{t}_i, t_{-i}') = 1 - \tau_i'$, as well as $f_0(\hat{t}_i, t_{-i}'') = 0$ and $f_{-i}(\hat{t}_i, t_{-i}'') = 1 - \tau_i''$. In order to ensure that both types t_{-i}' and t_{-i}'' of the other agent report truthfully, it must hold that $\tau_i' = \tau_i''$. Given this observation, it is intuitive that admissible binary functions f_0 can only display the shapes in Proposition 3.

To conclude this section, we re-interpret the compromise model as a public good model and impose individual rationality constraints. It is easy to obtain from Proposition 3 the following result:

Corollary 1 *A binary decision rule is* ***strategy-proof and ex post individually rational*** *if and only if it belongs to one of the following categories in Proposition 3: Category I.(ii) with* $f_i(t) = \delta_i$ *for all* $i \in I$*, and Category III.(ii) with* $\tau_i = \delta_i$ *for all* $i \in I$*.*

The proof is in the Appendix. The set of individually rational and strategy-proof binary decision rules characterized in Corollary 1 corresponds to the set of strategy-proof ex post individually rational and ex post budget balanced deterministic provision mechanisms characterized by Chung and Ely (2006), and by Börgers (2006), for the canonical public good setup. The fact that these mechanisms are a strict subset of those in Proposition 3 illustrates, as pointed out in the introduction, that individual liquidity constraints are not binding in the public good variant of our model, where participation constraints must be respected. Note that this is the case not just with deterministic mechanisms, but holds for all public good provision rules when dominant strategy implementation is used.

5 Some Results for Non-Binary Decision Rules

We have shown in Proposition 2 that additivity is a sufficient condition for a function f_0 to be admissible. However, it is obvious from the previous section on binary decision rules that additivity is not necessary. The reason is that binary rules feature piecewise-constant functions f_0 that cannot be described as the sum of two univariate functions.

The task of characterizing all functions $f_0 : T \to [0,1]$ that satisfy item (iii) of Proposition 1 is a difficult problem in general. As a first step towards such a characterization, we derive here a *necessary* condition for a function f_0 to be admissible by adopting a differential approach.[11] If this appears restrictive, note that monotonicity of f_0 (as required by item (i) of Proposition 1) ensures the existence almost everywhere (a.e.) of the first-order partial derivatives $\partial f_0(t)/\partial t_i$ of f_0. However, it would be remiss not to emphasize that our necessary condition below relies on an additional smoothness assumption. In particular, we assume that the partial derivatives $\partial f_0(t_i,\cdot)/\partial t_i$ of f_0 are absolutely continuous. This assumption guarantees the existence a.e. of the cross partial derivatives $\partial^2 f_0(t)/\partial t_i \partial t_j$ of f_0. Given this assumption, we can show the following result:

Proposition 4 *Suppose f_0 is nondecreasing. Suppose also that for all $i \in I$ the partial derivative $\partial f_0(t_i,\cdot)/\partial t_i$ is an absolutely continuous function for every t_i where it exists. If f_0 satisfies item (iii) of Proposition 1, then for all $i \in I$ and almost all t_i, t_{-i}, $t'_{-i} \in [0,1]$:*

$$\frac{\partial f_0(t_i,t_{-i})}{\partial t_i} = \frac{\partial f_0(t_i,t'_{-i})}{\partial t_i}. \tag{2}$$

It is easy to show that any continuously differentiable function f_0 that satisfies condition (2) in Proposition 4 must be additive. To see this, assume that the partial derivatives $\partial f_0(t)/\partial t_i$ exist everywhere and are continuous. Set $t'_{-i} = 1$ and

[11] See Laffont and Maskin (1980) for a differential approach to efficient public good provision rules.

DOI: 10.2202/1935-1704.1684

integrate both sides of (2) from t_i up to 1. This yields the following additive function:

$$f_0(t) = f_0(t_1,1) + f_0(1,t_2) - f_0(1,1). \tag{3}$$

It follows directly from Proposition 2 that the function f_0 in (3) is admissible.[12] Thus, for the class of continuously differentiable nondecreasing functions, condition (2) in Proposition 4 is both necessary *and* sufficient for f_0 to be admissible.[13]

Obviously, limiting ourselves to continuously differentiable functions f_0 is too restrictive. However, it is difficult to establish in general a sufficient condition for strategy-proofness once we move outside this class. To see why, note that while a nondecreasing function $f_0(\cdot,t_{-i})$ (for given $t_{-i} \in [0,1]$) can display only countably many jump-discontinuities, there may nevertheless be a large number of jumps in the value of f_0.[14] We therefore consider a very limited departure from the class of continuously differentiable functions in order to explore what conditions beyond the one in Proposition 4 are needed to make a *discontinuous* function f_0 admissible. The particular class of functions we study now has been chosen because it includes the piecewise constant step functions f_0 that are associated with the strategy-proof binary decision rules in Proposition 3.

Definition 6 *Denote by $\mathscr{F}$ the class of nondecreasing functions $f_0 : T \to [0,1]$ where:*

(i) For every agent $i \in I$, there is a type $\tau_i \in [0,1]$ s.t. if, for any $t_{-i} \in [0,1]$, there exists some type $\hat{t}_i \in [0,1]$ for which $f_0^-(\hat{t}_i,t_{-i}) < f_0^+(\hat{t}_i,t_{-i})$, then $\hat{t}_i = \tau_i$. Furthermore, $f_0(\tau_i,t_{-i}) \in \{f_0^-(\tau_i,t_{-i}), f_0^+(\tau_i,t_{-i})\}$.[15]

(ii) For all $t_{-i} \in [0,1]$, the partial derivative $\partial f_0(\cdot,t_{-i})/\partial t_i$ of f_0 is continuous at every $t_i \in [0,1]$ where it exists.

Now consider a function f_0 in $\mathscr{F}$ and suppose that it satisfies the necessary condition (2) in Proposition 4. We can then show that f_0 must be piecewise additive:

[12] Barberà et al. (1998), who study the design of cardinal probabilistic decision rules, have a result which implies that twice continuously differentiable rules are additive. However, their results are limited to either continuous or twice continuously differentiable decision rules.

[13] While we omit for the sake of brevity a formal statement and proof, it is possible to show that in the presence of participation constraints, any additive public good provision rule f_0 is "degenerate": It provides the public good with probability zero for all type-pairs $t \in [0,1)^2$, and consists of the sum of functions f_0^i (for every $i \in I$) such that $f_0^i(t_i) = 0$ if $t_i < 1$, and $f_0^i(t_i) = \pi_i$ if $t_i = 1$, where $\pi_i \in [0, \delta_i - f_i(1,1)]$. This finding suggests that the focus on deterministic mechanisms in the literature may not be overly restrictive.

[14] For a definition of jump discontinuity, see e.g. Definition 4.49 in Apostol (1974).

[15] We denote by $f_0^-(\hat{t}_i,t_{-i})$ ($f_0^+(\hat{t}_i,t_{-i})$, resp.) the left-hand (right-hand, resp.) limit of f_0 at $\hat{t}_i$. I.e. $f_0^-(\hat{t}_i,t_{-i}) = \lim_{t_i \to \hat{t}_i^-} f_0(t_i,t_{-i})$ and $f_0^+(\hat{t}_i,t_{-i}) = \lim_{t_i \to \hat{t}_i^+} f_0(t_i,t_{-i})$.

Lemma 2 *If a function f_0 in $\mathscr{F}$ satisfies condition (2) in Proposition 4 then:*

$$f_0(t) = \begin{cases} f_0(t_1,1) + f_0(1,t_2) - f_0(1,1) + k & \text{if } t_i < \tau_i \ \forall i \in I, \\ f_0(t_1,1) + f_0(1,t_2) - f_0(1,1) & \text{if } \exists i \in I \text{ s.t. } t_i > \tau_i, \end{cases} \tag{4}$$

where $k \equiv [f_0^+(\tau_i,1) - f_0^-(\tau_i,1)] - [f_0^+(\tau_i,t_{-i}) - f_0^-(\tau_i,t_{-i})] = const.$ $\forall i \in I$ and $\forall t_{-i} < \tau_{-i}$.

The proof is in the Appendix. The additive term k in Lemma 2 represents the difference in the size of the jump in f_0 at the boundary point $(\tau_i,1)$, and at an interior point (τ_i,t_{-i}). The fact that this difference in jump-size must be constant for all $t_{-i} < \tau_{-i}$ restricts the types of functions in $\mathscr{F}$ that are admissible. Observe that, in contrast to the continuously differentiable case, condition (2) in Proposition 4 is *not* sufficient for functions in $\mathscr{F}$ to be admissible. The reason is that item (iii) of Proposition 1 implies further restrictions on admissible functions by limiting the types τ_i at which jumps in f_0 can occur.

Proposition 5 *A function f_0 in $\mathscr{F}$ is admissible if and only if it is a piecewise additive function as given in Lemma 2, with $k(1-\tau_1-\tau_2) = 0$.*

The proof is omitted as it is straightforward to verify that, for any type-pair t s.t. $t_i < \tau_i$ for all $i \in I$, the piecewise additive function f_0 in (4) satisfies item (iii) of Proposition 1 only if $k(1-\tau_1-\tau_2) = 0$. I.e. if the difference in jump size k is strictly positive, then the types τ_1 and τ_2 at which discontinuities in f_0 may arise must form a point on the cross-diagonal in the unit-square T.

We conclude this section by highlighting the usefulness of Proposition 5 for constructing admissible functions in $\mathscr{F}$ from given "boundary functions" $f_0(t_1,1)$ and $f_0(1,t_2)$. For example, we can generate in this way any binary decision rule in Category III.(ii) of Proposition 3.[16] We can also generate strategy-proof rules that are not piecewise constant, as the following example shows:

Example 2 *For all $i \in I$, let $\tau_i = 0.5$ and fix boundary functions $f_0(t_i,1)$ s.t. $f_0(t_i,1) = 0.375 + 0.25t_i$ if $t_i < 0.5$, and $f_0(t_i,1) = 0.5 + 0.5t_i$ if $t_i > 0.5$. This yields $k = 0.25 - [f_0^+(0.5,t_{-i}) - f_0^-(0.5,t_{-i})]$ for any $i \in I$ and all $t_{-i} < 0.5$. It is easy to verify*

[16] To see this, fix any pair (τ_1,τ_2) s.t. $\tau_1+\tau_2 = 1$. Define for each $i \in I$ a boundary function $f_0(t_i,1)$ s.t. $f_0(t_i,1) = 0$ if $t_i < \tau_i$, and $f_0(t_i,1) = 1$ if $t_i > \tau_i$. This implies $k = 1 - [f_0^+(\tau_i,t_{-i}) - f_0^-(\tau_i,t_{-i})]$ for any $i \in I$ and all $t_{-i} < \tau_{-i}$. The function $f_0(t)$ in Category III.(ii) of Proposition 3 then corresponds to the one we obtain from (4) for $k = 1$, which is required to ensure that $f_0(t) \in [0,1]$ $\forall t \in T$.

that the function $f_0(t)$ given by (4) takes values in $[0,1]$ for all $t \in T$, which is due to the fact that $f_0^+(0.5,t_{-i}) = f_0^-(0.5,t_{-i})$, and therefore $k = 0.25$:

$$f_0(t) = \begin{cases} 0.25(t_1+t_2) & \text{if } t_1,t_2 \le 0.5, \\ 0.375+0.25t_{-i}-0.5(1-t_i) & \text{if } t_i > 0.5, t_{-i} \le 0.5, \\ 0.5(t_1+t_2) & \text{if } t_1,t_2 > 0.5. \end{cases}$$

The associated probabilities $f_1(t)$ and $f_2(t)$ that render the decision rule (f_0, f_1, f_2) strategy-proof are then obtained from item (ii) of Proposition 1:

$$f_1(t) = \begin{cases} 0.5-0.125\left(t_1^2+t_2(2-t_2)\right) & \text{if } t_1,t_2 \le 0.5, \\ 0.53125-0.125(2t_1^2+t_2(2-t_2)) & \text{if } t_1 > 0.5, t_2 \le 0.5, \\ 0.59375-0.125(t_1^2+2t_2(2-t_2)) & \text{if } t_1 \le 0.5, t_2 > 0.5, \\ 0.5-0.25(t_1^2+t_2(2-t_2)) & \text{if } t_1,t_2 > 0.5. \end{cases}$$

6 Efficient Decision Rules

We have so far studied the characteristics of strategy-proof decision rules, leaving aside the question of which rule should be selected for the purpose of reaching a collective decision. To obtain a criterion for choosing between decision rules, we make recourse to the efficiency notions defined in Holmström and Myerson (1983).[17] In the spirit of dominant strategy implementation, we want to keep our model belief-free. Therefore, we focus here on ex post efficient decision rules. A decision rule f is said to be ex post efficient if, for given welfare weights $\lambda_i : T \to \mathbb{R}_+$ (for every $i \in I$) that depend arbitrarily on t, the decision rule attains the highest level of social welfare. Social welfare associated with a decision rule f is the aggregate over all $t \in T$ of the weighted sum of the agents' ex post utilities:

$$\int_T \left(\sum_{i\in I} \lambda_i(t) \left[f_i(t) + f_0(t)t_i \right] \right) dt. \tag{5}$$

As pointed out by Holmström and Myerson, ex post efficient decision rules are those that maximize, for *every* type-pair $t \in T$, the weighted sum of the agents' ex post utilities in the integrand of (5). Using the fact that decision rules are probability distributions over the set of alternatives A, we can write this sum as:

$$\lambda_1(t) f_1(t) + \left[\lambda_1(t)t_1 + \lambda_2(t)t_2\right] f_0(t) + \lambda_2(t) f_2(t).$$

This expression is additive in the probabilities f_0, f_1 and f_2, making it is easy to derive properties of decision rules that are welfare-maximizing among *all* decision

[17] Observe that in the compromise model, any decision (i.e. the alternative in A chosen by the decision rule) is Pareto efficient, because for every type-pair $t \in (0,1)^2$ it is impossible to make one agent better off by switching to a different alternative without making the other agent worse off.

rules. Adopting Holmström and Myerson's terminology, we call such decision rules ex post classically efficient.

Definition 7 *A decision rule f is **ex post classically efficient** if and only if for every $t \in T$:*

$$\lambda_1(t)t_1 + \lambda_2(t)t_2 > \max\{\lambda_1(t), \lambda_2(t)\} \quad \Rightarrow \quad f_i(t) = 0 \quad \forall i \in I,$$

$$\lambda_1(t)t_1 + \lambda_2(t)t_2 < \max\{\lambda_1(t), \lambda_2(t)\} \quad \Rightarrow \quad f_0(t) = 0,$$

$$\lambda_j(t) < \lambda_i(t) \quad (i, j \in I, j \neq i) \quad \Rightarrow \quad f_j(t) = 0.$$

Definition 7 states, for each $t \in T$, which alternatives in A must *not* be implemented by an ex post classically efficient decision rule. Observe that ex post classical efficiency need not prescribe a unique choice in A, in which case Definition 7 does not constrain the probabilities of any alternatives that have not been ruled out. Consequently, the set of ex post classically efficient decision rules contains both binary rules and non-binary rules.

As the agents' types are privately observed, any ex post classically efficient decision rule can be implemented only if it is strategy-proof. Given the considerable degrees of freedom in the choice of agents' welfare weights, it is interesting to ask if strategy-proof ex post classically efficient decision rules exist. A trivial example is the rule that always selects agent 1's favorite alternative: efficiency follows from the fact that this rule maximizes ex post welfare for every type-pair $t \in T$ if agent 2's weight $\lambda_2(t) = 0$ for all t; strategy-proofness of this rule follows from Category I.(ii) of Proposition 3. In order to exclude such cases, and to see if there exist strategy-proof ex post classically efficient decision rules that are responsive to the agents' types, we assume in the remainder of this section that agents' welfare weights are strictly positive: $\lambda_i(t) > 0$ for all $t \in T$ and all $i \in I$. Given this assumption, we can provide a sharper characterization of ex post classically efficient decision rules:

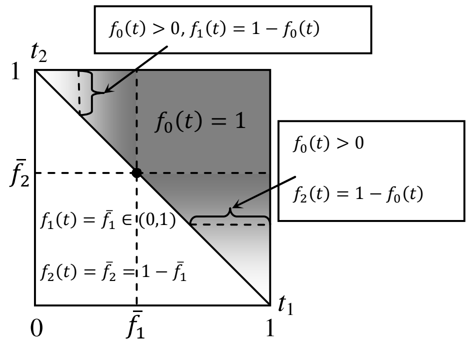

Figure 1: Illustration of proof of Proposition 6

Lemma 3 *Let f be an ex post classically efficient decision rule for some pair of welfare weights $\lambda_i : T \to \mathbb{R}_{++}$ ($\forall i \in I$). Then:*

$$\text{For } t_1 + t_2 < 1: \quad f_1(t) \in [0,1],\ f_2(t) = 1 - f_1(t).$$

$$\text{For } t_1 + t_2 = 1: \quad f_0(t), f_1(t) \in [0,1],\ f_2(t) = 1 - f_0(t) - f_1(t).$$

$$\text{For } t_1 + t_2 > 1:$$

$$\text{if } t_1 < 1, t_2 < 1: \quad f_0(t) \in [0,1],\ \exists i \in I : f_i(t) = 1 - f_0(t),$$

$$\text{if } t_i = 1, t_j < 1: \quad f_0(t) \in [0,1],\ f_j(t) = 1 - f_0(t),\ i, j \in I,\ j \neq i,$$

$$\text{if } t_1 = 1, t_2 = 1: \quad f_0(t) = 1.$$

The proof of Lemma 3 is in the Appendix. Lemma 3 states for each type-pair the alternatives that can be assigned positive probability by an ex post classically efficient decision rule. It is obvious from Lemma 3 that none of the strategy-proof binary decision rules in Proposition 3 are ex post classically efficient. We now show that even among all non-binary decision rules there is none that is both ex post classically efficient *and* strategy-proof:

Proposition 6 *If both agents' welfare weights $\lambda_i(\cdot)$ are strictly positive for all $t \in T$, then there exists no ex post classically efficient and strategy-proof decision rule.*

The proof, together with a series of lemmas that it builds on, is in the Appendix. However, the key idea of the proof is conveyed by Fig. 1. We show that *any* ex post classically efficient strategy-proof decision rule must have the shape displayed in Fig. 1. In particular, the compromise is never chosen for type-pairs t s.t. $t_1 + t_2 < 1$. For any such type-pair, each agent i's favorite alternative receives constant and strictly positive probability $\bar{f}_i$, with $\bar{f}_1 + \bar{f}_2 = 1$. For all type-pairs t s.t. $t_i > \bar{f}_i$ for all i, the compromise is selected with probability 1. Finally, for type-pairs t s.t. $t_i > \bar{f}_i$ and $1 - t_i < t_{-i} < \bar{f}_{-i}$, the decision rule assigns positive probability to alternatives a_0 and a_{-i} only. Now fix a type $\hat{t}_2 < \bar{f}_2$. In the proof of Proposition 6 we establish, for all $t_1 > 1 - \hat{t}_2$, that the probability of the compromise $f_0(t_1, \hat{t}_2)$ is a number in $(0,1)$. Next, fix some type $\hat{t}_1 > 1 - \hat{t}_2$ and consider agent 2 of type $\tilde{t}_2$, with $\bar{f}_2 < \tilde{t}_2 < 1$. Truthful revelation of his type gives agent 2 a utility of $\tilde{t}_2$ because the compromise is selected with probability 1. A report of $\hat{t}_2 < \bar{f}_2$ will instead give agent 2 a utility of $f_2(\hat{t}_1, \hat{t}_2) + f_0(\hat{t}_1, \hat{t}_2)\tilde{t}_2 = 1 - f_0(\hat{t}_1, \hat{t}_2)(1 - \tilde{t}_2)$. This follows from the fact that at $\hat{t}$, the decision rule assigns positive probability to a_0 and a_2 only. Incentive compatibility therefore requires that $f_0(\hat{t}_1, \hat{t}_2) = 1$. This, however, leads to a contradiction because there are types $t_1 \in (\bar{f}_1, 1 - \hat{t}_2)$ who would prefer to claim that their type is $\hat{t}_1 > 1 - \hat{t}_2$ in order to get the compromise for sure, rather than obtain their favorite alternative with probability $\bar{f}_1$ when reporting their true type.

It is important to reiterate that Proposition 6 is not a special case of the literature on the impossibility of implementing non-dictatorial cardinal and ex post efficient social choice rules. Furthermore, despite the formal similarity between the compromise model and canonical mechanism design models with transferable utility and quasi-linear preferences (such as Green and Laffont, 1979), Proposition 6 *cannot* be established by arguing that Vickrey-Clarke-Groves (VCG) mechanisms are the only ex post classically efficient and strategy-proof decision rules, but that they cannot serve as decision rules here because their components f_0, f_1 and f_2 do not sum up to 1 for all type-pairs. The reason we cannot construct the proof of Proposition 6 along these lines is that VCG mechanisms exist only in the special case where the agents have the *same* weights in the social welfare function.[18] In all other cases, we cannot maximize social welfare by choosing the function f_0 independently from the functions f_1 and f_2.

As there are no non-trivial ex post classically efficient decision rules that are strategy-proof, one may ask which, among all strategy-proof decision rules, maximize social welfare in (5) for *some* pair of weights $(\lambda_1(\cdot), \lambda_2(\cdot))$. We call such decision rules *ex post incentive efficient*.[19] In the absence of a full characterization

[18] See Example 1 in Section 3 above, esp. footnote 8.

[19] Clearly, the need to ensure incentive compatibility of the chosen rule will result in distortions relative to ex post classical efficiency. I.e. there will be type-pairs $t \in T$ for which the decision rule fails to maximize the weighted sum of the agents' utilities. An ex post incentive efficient decision rule therefore is one which yields the smallest aggregate distortion across all type-pairs t, in comparison with all other strategy-proof rules.

DOI: 10.2202/1935-1704.1684

of all functions f_0 that can be part of a strategy-proof decision rule, we restrict attention to binary rules. Given the vast degrees of freedom in the choice of welfare weights $\lambda_i(\cdot)$ it is not surprising that there are *many* ex post incentive efficient binary decision rules. In fact, as the following example shows, each category of non-trivial strategy-proof binary rules in Proposition 3 (i.e. Categories II and III) contains at least one rule that is ex post incentive efficient.

Example 3 *For weights $\lambda_1(t) = \lambda_2(t) = 4t_1t_2$, the welfare-maximizing strategy-proof binary decision rule is as shown in Category III.(i) of Proposition 3, with $\tau_1 = \tau_2 = 0.5$.*
For weights $\lambda_1(t) = \lambda_2(t) = 4(1-t_1)(1-t_2)$, the welfare-maximizing strategy-proof binary decision rule is as shown in Category III.(ii) of Proposition 3, with $\tau_1 = \tau_2 = 0.5$.
For weights $\lambda_1(t) = \lambda_2(t) = 1$ for all $t \in T$, the following strategy-proof binary decision rules in Proposition 3 maximize social welfare: Category II with $\tau_i = 0.5$ for any $i \in I$, III.(i) with $\tau_1 = \tau_2 = 0.5$, and Category III.(ii) with $\tau_1 = \tau_2 = 0.5$.

In order to obtain a stronger selection criterion, we could turn to the notion of ex ante incentive efficiency (see Holmström and Myerson, 1983). This, however, would require us to assume that the mechanism designer has a well-defined subjective probability distribution that represents his beliefs about the agents' types. As pointed out by Chung and Ely (2007), such an assumption would introduce an asymmetry into the model: By using strategy-proof (rather than Bayesian incentive compatible) decision rules, the mechanism designer avoids completely the need to make assumptions regarding the agents' beliefs about each others' types. However, a reluctance to formulate any view regarding the agents' beliefs about each other seems at odds with a precisely held subjective belief about the agents' types themselves.[20] We therefore do not pursue this issue further here.[21]

7 Conclusion

Adopting strategy-proofness as our implementation concept, we have taken a unified approach to the compromise model of Börgers and Postl (2009) and the closely related canonical public good model. We have shown that strategy-proof decision

[20]Chung and Ely (2007) provide a foundation for using strategy-proof mechanisms in the context of optimal auction design.

[21]For a prior-free approach to efficient mechanisms in the closely related public good model, see Smith (2010).

rules that are smooth, and those that fall into a special class of discontinuous mechanisms (which includes the deterministic mechanisms widely studied in the literature), must be (piecewise) additive. However, for strategy-proofness, any piecewise additive decision rule that displays jump-discontinuities must satisfy further conditions that limit the location and magnitude of these jumps. In future work, it would be interesting to see how many jump-discontinuities a strategy-proof decision rule can support, and to what extent such rules allow us to approximate more closely ex post efficient rules. Regarding the question of which ex post efficient rules are strategy-proof, we have proved a "dictatorship result" that, while familiar in spirit, is not the consequence of any known dictatorship result in the literature. An interesting open question is the extent to which universal preference domain assumptions in this literature can be weakened while still sustaining dictatorship as the only way of implementing an ex post efficient cardinal probabilistic decision rule.

8 Appendix

Proof of Proposition 1. Item (i) is the same as item (i) in Lemma 1. As the proof is familiar from the literature it is omitted here. In what follows, we explicitly derive the expressions in items (ii) and (iii) of Proposition 1.

Item (ii). We show that the additive term $f_i(1,t_{-i})+f_0(1,t_{-i})$ in item (ii) of Lemma 1 can be expressed solely in terms of the function f_0 and constants $f_i(1,1)$. To see this, suppose that f is an incentive compatible decision rule, so that the probabilities $f_1(t_1,t_2)$ and $f_2(t_1,t_2)$ of alternatives a_1 and a_2, resp., are given by item (ii) of Lemma 1. As the probabilities f_1, f_2 and f_0 together sum up to 1 for every type-pair we obtain:

$$f_1(1,t_2)+f_0(1,t_2)-\int_{t_1}^{1} f_0(s,t_2)ds+f_2(t_1,1)+f_0(t_1,1)-\int_{t_2}^{1} f_0(t_1,s)ds \\ = 1+f_0(t_1,t_2)(t_1+t_2-1). \quad (6)$$

For $t_1=1$ equation (6) reduces to:

$$f_1(1,t_2)+f_0(1,t_2)+f_2(1,1)+f_0(1,1)-\int_{t_2}^{1} f_0(1,s)ds = 1+f_0(1,t_2)t_2. \quad (7)$$

Solving equation (7) for $f_1(1,t_2)$ yields:

$$f_1(1,t_2) = 1-f_0(1,1)-f_2(1,1)+f_0(1,t_2)t_2-f_0(1,t_2)+\int_{t_2}^{1} f_0(1,s)ds. \quad (8)$$

DOI: 10.2202/1935-1704.1684

As the probabilities f_0, f_1 and f_2 sum up to 1 at every $(t_1,t_2) \in [0,1]^2$ we have $f_1(1,1) = 1 - f_0(1,1) - f_2(1,1)$. We can therefore write equation (8) as:

$$f_1(1,t_2) = f_1(1,1) + f_0(1,t_2)t_2 - f_0(1,t_2) + \int_{t_2}^{1} f_0(1,s)ds. \tag{9}$$

Substituting the expression for $f_1(1,t_2)$ in (9) into the probability of agent 1's favorite alternative in item (ii) of Lemma 1 we obtain:

$$f_1(t_1,t_2) = f_1(1,1) + f_0(1,t_2)t_2 + \int_{t_2}^{1} f_0(1,s)ds - f_0(t_1,t_2)t_1 - \int_{t_1}^{1} f_0(s,t_2)ds. \tag{10}$$

Equation (10) is the probability of agent 1's favorite alternative in item (ii) of Proposition 1. In the same way we can derive the probability of agent 2's favorite alternative.

We now show that the functions f_i in item (ii) of Proposition 1 only take values in $[0,1]$ for every $(t_1,t_2) \in [0,1]^2$. For this purpose, consider the probability of agent 1's favorite alternative in (10). Fix a value of t_2 and consider the behavior of $f_1(t_1,t_2)$ as a function of t_1. The partial derivative w.r.t. t_1 is:

$$\frac{\partial f_1(t_1,t_2)}{\partial t_1} = -\frac{\partial f_0(t_1,t_2)}{\partial t_1} t_1.$$

Monotonicity of f_0 implies that $f_1(t_1,t_2)$ is nonincreasing for all $t_1 \in [0,1]$. We therefore only have to show that for any monotone function f_0 and all $t_2 \in [0,1]$ it holds that $f_1(1,t_2)$ is nonnegative and that $f_1(0,t_2)$ is no larger than 1. To show that this is true, fix some value for $f_0(1,t_2)$ and consider the term:

$$f_1(1,1) + f_0(1,t_2)t_2 + \int_{t_2}^{1} f_0(1,s)ds, \tag{11}$$

which is part of the expression for the probability of agent 1's favorite alternative in (10). We now ask how the choice of function f_0 affects the magnitude of (11). First note that for a fixed value of t_2, the minimum of (11) is attained by setting $f_0(1,s) = f_0(1,t_2)$ for all $s > t_2$. This yields:

$$f_1(1,1) + f_0(1,t_2)t_2 + \int_{t_2}^{1} f_0(1,t_2)ds = f_1(1,1) + f_0(1,t_2). \tag{12}$$

Now obtain from (10) the probability of agent 1's favorite alternative at the point $(1,t_2)$, making use of the expression in (12). Denoting the resulting function by $f_1^{\min}(1,t_2)$, we obtain:

$$f_1^{\min}(1,t_2) = f_1(1,1) + f_0(1,t_2) - f_0(t_1,t_2) = f_1(1,1). \tag{13}$$

Equation (13) shows that, for all $t_2 \in [0,1]$, $f_1^{\min}(1,t_2)$ is equal to the given constant $f_1(1,1)$. This implies that $f_1(t_1,t_2) \geq f_1(1,1) \geq 0$ for all (t_1,t_2), as required.

Observe that for a fixed value of t_2, the maximum of (11) is attained by setting $f_0(1,s) = f_0(1,1)$ for all $s > t_2$. This yields:

$$f_1(1,1) + f_0(1,t_2)t_2 + \int_{t_2}^{1} f_0(1,1)ds = 1 - f_2(1,1) - t_2(f_0(1,1) - f_0(1,t_2)). \quad (14)$$

Now obtain from (10) the probability of agent 1's favorite alternative at the point $(0,t_2)$, making use of the expression in (14). Denoting the resulting function by $f_1^{\max}(0,t_2)$, we obtain:

$$f_1^{\max}(0,t_2) = 1 - f_2(1,1) - t_2(f_0(1,1) - f_0(1,t_2)) - \int_0^1 f_0(s,t_2)ds. \quad (15)$$

Equation (15) shows that $f_1^{\max}(0,t_2)$ takes a value smaller than 1 for all $t_2 \in [0,1]$. This implies $f_1(t_1,t_2) \leq 1$ for all (t_1,t_2), as required.

Item (iii). The result is obtained by substituting the expressions for $f_1(1,t_2)$ in (9) and the corresponding expression for $f_2(t_1,1)$ into equation (6). Noting that $f_1(1,1) + f_2(1,1) = 1 - f_0(1,1)$ we obtain:

$$f_0(t_1,1)t_1 + f_0(1,t_2)t_2 - f_0(1,1) + \int_{t_1}^{1} f_0(s_1,1)ds_1 + \int_{t_2}^{1} f_0(1,s_2)ds_2$$
$$= f_0(t_1,t_2)(t_1+t_2-1) + \int_{t_1}^{1} f_0(s_1,t_2)ds_1 + \int_{t_2}^{1} f_0(t_1,s_2)ds_2,$$

which can easily be rearranged to yield item (iii) in Proposition 1. □

We now prepare the ground for the proof of Proposition 3. It is easy to verify that the binary decision rules in Proposition 3 are strategy-proof. We therefore only prove necessity here by deriving restrictions that item (iii) of Proposition 1 imposes on admissible functions f_0. A key building block in the proof of Proposition 3 is presented in Lemma A.1 below:

Lemma A.1 *If there are two type-pairs (t_1',t_2') and (t_1'',t_2'') in the interior of the unit square T with $0 < t_1' < t_1'' < 1$ and $0 < t_2'' < t_2' < 1$ s.t. $f_0(t_1',t_2') = f_0(t_1'',t_2'') = 1$ then either (i) $f_0(t_1',t_2'') = 1$; or (ii) $t_1''+t_2' \geq 1 > t_1'+t_2''$ and $\exists t_1^* \in (t_1', 1-t_2'']$ s.t. $f_0(t_1,t_2) = 0$ if $t_1 < t_1^* \wedge t_2 < 1-t_1^*$ and $f_0(t_1,t_2) = 1$ if $(t_1 \geq t_1^* \wedge t_2 \geq t_2'') \vee (t_1 \geq t_1' \wedge t_2 \geq 1-t_1^*)$.*

The function f_0 in case (i) of Lemma A.1 is illustrated in the left-hand panel of Fig. 2, while the f_0 in case (ii) is shown in the right-hand panel of Fig. 2.

DOI: 10.2202/1935-1704.1684

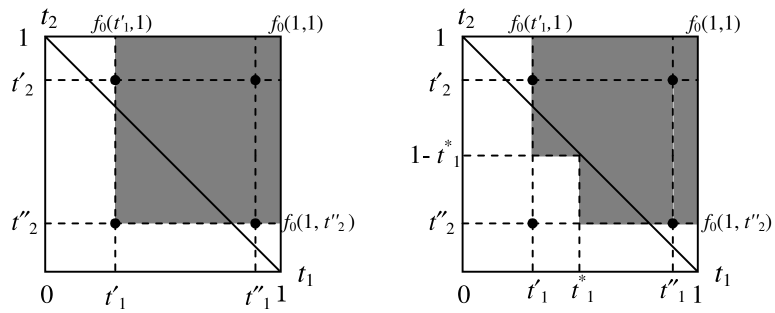

Figure 2: Illustration of Lemma A.1

Proof of Lemma A.1. Suppose there are two type-pairs (t_1', t_2') and (t_1'', t_2'') with $0 < t_1' < t_1'' < 1$ and $0 < t_2'' < t_2' < 1$ s.t. $f_0(t_1', t_2') = f_0(t_1'', t_2'') = 1$. Monotonicity of f_0 implies $f_0(t_1, t_2) = 1$ if $(t_1 \geq t_1'' \wedge t_2 \geq t_2'') \vee (t_1 \geq t_1' \wedge t_2 \geq t_2')$. Then, at (t_1', t_2'') item (iii) of Proposition 1 reduces to:

$$[1 - f_0(t_1', t_2'')](t_1' + t_2'' - 1) + \int_{t_1'}^{t_1''} [1 - f_0(s, t_2'')]ds + \int_{t_2''}^{t_2'} [1 - f_0(t_1', s)]ds = 0. \quad (16)$$

As f_0 takes values in $\{0, 1\}$ we have either $f_0(t_1', t_2'') = 0$ or $f_0(t_1', t_2'') = 1$. If $f_0(t_1', t_2'') = 1$, then equation (16) is satisfied, as claimed in Case (i) of Lemma A.1. If, instead, $f_0(t_1', t_2'') = 0$, then (16) reduces to:

$$t_1'' + t_2' - 1 - \int_{t_1'}^{t_1''} f_0(s, t_2'')ds - \int_{t_2''}^{t_2'} f_0(t_1', s)ds = 0. \quad (17)$$

A necessary condition for equation (17) to hold is that $t_1'' + t_2' \geq 1$. Furthermore, there exist types $t_1^* \in (t_1', t_1'']$ and $t_2^* \in (t_2'', t_2']$ s.t. $f_0(s, t_2'') = 0$ if $s < t_1^*$ and $f_0(s, t_2'') = 1$ if $s \geq t_1^*$ and $f_0(t_1', s) = 0$ if $s < t_2^*$ and $f_0(t_1', s) = 1$ if $s \geq t_2^*$. Thus, (17) reduces to:

$$t_1'' + t_2' - 1 - \int_{t_1^*}^{t_1''} ds - \int_{t_2^*}^{t_2'} ds = 0 \Leftrightarrow t_1^* + t_2^* = 1.$$

Note that types $t_1^* \in (t_1', t_1'']$ and $t_2^* \in (t_2'', t_2']$ s.t. $t_1^* + t_2^* = 1$ exist only if: $t_2'' < 1 - t_1^* \leq t_2' \Leftrightarrow t_1^* + t_2'' < 1 \leq t_1^* + t_2'$. The above necessary condition that $t_1'' + t_2' \geq 1$ ensures that there exists a value t_1^* s.t. $t_1^* + t_2'' < 1$. In order to ensure that there also exists a $t_2^* = 1 - t_1^*$ s.t. $1 \leq t_1^* + t_2'$ it must be that case that the lower bound on t_1^* (namely t_1') is s.t. $t_1' + t_2'' < 1$. This gives rise to Case (ii) in Lemma A.1. □

Proof of Proposition 3. The proof proceeds by checking separately what values an admissible function f_0 must take if it is of the form displayed in either the left-hand panel of Fig. 2, or the right-hand panel of Fig. 2.

Case 1. Consider types t_1', t_1^* and t_2'' with $0 < t_1' < t_1^* < 1$ and $0 < t_2'' < 1$. Now suppose that f_0 is of the form displayed in the right-hand panel of Fig. 2. Suppose $f_0(t_1', t_2'') = 0$. Now pick any $\hat{t}_1$ with $0 \leq \hat{t}_1 < t_1'$. Monotonicity of f_0 implies $f_0(\hat{t}_1, t_2'') = 0$. At $(\hat{t}_1, t_2'')$ item (iii) of Proposition 1 reduces to:

$$f_0(\hat{t}_1,1)\hat{t}_1 + t_2'' - 1 + \int_{\hat{t}_1}^{t_1'} f_0(s,1)ds + \int_{t_1'}^{t_1^*} ds + \int_{t_2''}^{1} [1 - f_0(\hat{t}_1,s)]ds = 0.$$

$\Leftrightarrow$

$$t_1^* - t_1' = \int_{t_2''}^{1} f_0(\hat{t}_1,s)ds - f_0(\hat{t}_1,1)\hat{t}_1 - \int_{\hat{t}_1}^{t_1'} f_0(s,1)ds. \qquad (18)$$

The left-hand side of (18) is strictly positive. Thus, there must be a threshold $s^* > t_2''$ s.t. $f_0(\hat{t}_1, s) = 0$ if $s < s^*$ and $f_0(\hat{t}_1, s) = 1$ if $s \geq s^*$. Therefore (18) implies:

$$t_1^* - t_1' = 1 - s^* - \hat{t}_1 - (t_1' - \hat{t}_1) \Leftrightarrow s^* = 1 - t_1^*.$$

Employing the same logic we can show for any type-pair $(t_1', \hat{t}_2)$ with $0 \leq \hat{t}_2 < t_2''$ that $f_0(s, \hat{t}_2) = 0$ for all $s < t_1^*$ and $f_0(s, \hat{t}_2) = 1$ for all $s \geq t_1^*$. Thus, if f_0 is as shown in the right-hand panel of Fig. 2 and $f_0(t_1', t_2'') = 0$, then f_0 is as described in Category III.(i) of Proposition 1 with $\tau_1 = t_1^*$.

Case 2. Consider a type-pair (t_1', t_2'') and suppose that $f_0(t_1', t_2'') = 1$ (see left-hand panel of Fig. 2 for an illustration, but note that we do not, at this point, make any assumptions about whether $t_1' + t_2'' < 1$, $t_1' + t_2'' = 1$, or $t_1' + t_2'' > 1$). Now suppose that $f_0(t_1, t_2) = 0$ for all $t_1 < t_1'$ (the case where $f_0(t_1, t_2) = 0$ for all $t_2 < t_2''$ is analogous to what follows and is therefore omitted). Pick any type-pair $(\hat{t}_1, \hat{t}_2)$ with $0 \leq \hat{t}_1 < t_1'$ and $0 \leq \hat{t}_2 < t_2''$. At $(\hat{t}_1, \hat{t}_2)$ it therefore holds that $f_0(\hat{t}_1, \hat{t}_2) = 0$. At $(\hat{t}_1, \hat{t}_2)$ item (iii) of Proposition 1 reduces to:

$$f_0(1,\hat{t}_2)\hat{t}_2 - 1 + \int_{t_1'}^{1} [1 - f_0(s,\hat{t}_2)]ds + \int_{\hat{t}_2}^{t_2''} f_0(1,s)ds + \int_{t_2''}^{1} ds = 0$$

$\Leftrightarrow$

$$1 - t_1' - t_2'' = \int_{t_1'}^{1} f_0(s,\hat{t}_2)ds - f_0(1,\hat{t}_2)\hat{t}_2 - \int_{\hat{t}_2}^{t_2''} f_0(1,s)ds. \qquad (19)$$

We now distinguish two subcases:

Case 2.i. First suppose that $f_0(s, \hat{t}_2) = 0$ for all $s \geq t_1'$, and that there exists a type $s^* \in (\hat{t}_2, t_2'']$ s.t. $f_0(1, s) = 0$ if $s < s^*$ and $f_0(1, s) = 1$ if $s \geq s^*$. Then (19) reduces to:

$$1 - t_1' - t_2'' = -\int_{s^*}^{t_2''} f_0(1,s)ds \Leftrightarrow s^* = 1 - t_1'.$$

Such a type s^* of agent 2 exists only if $\hat{t}_2 < 1 - t_1' \leq t_2''$. In this case it follows that $f_0(1, 1 - t_1') = 1$. As we also have $f_0(t_1', t_2'') = 1$ (by the opening assumption in Case

2.), we can appeal to Lemma A.1 and thereby obtain that $f_0(t_1', 1-t_1'') = 1$. This shows that f_0 is as described in Category III.(ii) of Proposition 1 with $\tau_1 = t_1'$.

Case 2.ii. Now suppose there exists a type $s^* \in [t_1', 1)$ s.t. $f_0(s, \hat{t}_2) = 0$ for all $s < s^*$ and $f_0(s, \hat{t}_2) = 1$ for all $s \geq s^*$. Then (19) reduces to:

$$1 - t_1' - t_2'' = \int_{s^*}^{1} ds - \hat{t}_2 - \int_{\hat{t}_2}^{t_2''} ds \Leftrightarrow s^* = t_1'.$$

This shows that for types t_1' and t_2'' s.t. $t_1' + t_2'' < 1$ it follows that f_0 must be as described in Category II. of Proposition 1 with $\tau_1 = t_1'$. □

Proof of Corollary 1. It is easy to see that decision rules in Categories I.(ii) and III.(ii) of Proposition 3 are individually rational for appropriately chosen thresholds/probabilities of the agents' favorite alternatives. However, decision rules in Category I.(i) violate ex post rationality as for each $i \in I$, any type $t_i < \delta_i$ has strictly lower utility from the decision rule than in the default outcome. Next note that decision rules in Category II violate ex post individual rationality: Consider a decision rule where $f_0(t_1, t_2) = 1$ for all (t_1, t_2) s.t. $t_1 > \tau_1$. Then for any (t_1, t_2) s.t. $t_1 > \tau_1$ and $t_2 < \delta_2$ it holds that agent 2's utility is t_2, which is strictly lower than in the default outcome. Finally, decision rules in Category III.(i) violate ex post individual rationality: Consider any (t_1, t_2) s.t. $t_1 < \min\{\tau_1, \delta_1\}$ and $t_2 > \tau_2$. In this case, $f_0(t_1, t_2) = 1$ and agent 1's utility is t_1, which is strictly lower than in the default outcome. □

Proof of Proposition 4. Suppose f_0 satisfies item (i) of Proposition 1. Suppose also that the cross partial derivatives $\partial^2 f_0(t_i, t_{-i})/\partial t_i \partial t_{-i}$ exist for almost all $t_{-i} \in [0,1]$ (for all $i \in I$). Now take any type-pair $t \in (0,1)^2$ at which the partial derivatives and cross partial derivatives of f_0 exist. Suppose also that the partial derivatives $\partial f_0(t_1, 1)/\partial t_1$ and $\partial f_0(1, t_2)/\partial t_2$ exist. Differentiating both sides of the equation in item (iii) of Proposition 1 with respect to t_1 yields:

$$\frac{\partial f_0(t_1, 1)}{\partial t_1} t_1 + \frac{\partial f_0(t_1, t_2)}{\partial t_1}(1 - t_1 - t_2) - \int_{t_2}^{1} \frac{\partial f_0(t_1, s)}{\partial t_1} ds = 0.$$

Differentiating both sides of this equation with respect to t_2 yields:

$$\frac{\partial^2 f_0(t_1, t_2)}{\partial t_1 \partial t_2}(1 - t_1 - t_2) - \frac{\partial f_0(t_1, t_2)}{\partial t_1} + \frac{\partial f_0(t_1, t_2)}{\partial t_1} = 0$$

$\Leftrightarrow$

$$\frac{\partial^2 f_0(t_1, t_2)}{\partial t_1 \partial t_2}(1 - t_1 - t_2) = 0.$$

This shows that for any given type t_1, we must have $\partial^2 f_0(t_1, s_2)/\partial t_1 \partial t_2 = 0$ for all s_2 s.t. $t_1 + s_2 \neq 1$. Now pick any two types $t_2' < t_2''$ s.t. $t_1 + t_2' \neq 1$ and $t_1 + t_2'' \neq 1$.

By integrating both sides of the equation $\partial^2 f_0(t_1,\cdot)/\partial t_1 \partial t_2 = 0$ from t_2' to t_2'' we obtain:[22]

$$\int_{t_2'}^{t_2''} \frac{\partial^2 f_0(t_1,s_2)}{\partial t_1 \partial t_2} ds_2 = 0$$

$\Leftrightarrow$

$$\left[\frac{\partial f_0(t_1,s_2)}{\partial t_1}\right]_{t_2'}^{t_2''} = 0$$

$\Leftrightarrow$

$$\frac{\partial f_0(t_1,t_2'')}{\partial t_1} - \frac{\partial f_0(t_1,t_2')}{\partial t_1} = 0,$$

which establishes the result in Proposition 4. □

Proof of Lemma 2. Consider a function f_0 in $\mathscr{F}$ and assume that it satisfies the necessary condition (2) in Proposition 4. Now fix a $t_i < \tau_i$ and integrate both sides of (2) from t_i up to 1, taking account of the singularity at τ_i. This yields:

$$f_0(t) = f_0(t_i,1) + f_0(1,t_{-i}) - f_0(1,1) + k_i(t_{-i}) \quad \forall t \text{ s.t. } t_i < \tau_i, \tag{20}$$

where

$$k_i(t_{-i}) = [f_0^+(\tau_i,1) - f_0^-(\tau_i,1)] - [f_0^+(\tau_i,t_{-i}) - f_0^-(\tau_i,t_{-i})].$$

Evaluating equation (20) for $i = 1$ and for $i = 2$ we obtain two expressions for $f_0(t)$. These two expressions must be identical for all t s.t. $t_1 < \tau_1$ and $t_2 < \tau_2$. This implies:

$$k_1(t_2) = k_2(t_1) \equiv k \quad \forall t \text{ s.t. } t_1 < \tau_1 \text{ and } t_2 < \tau_2.$$

Thus, $f_0(t) = f_0(t_1,1) + f_0(1,t_2) - f_0(1,1) + k$ for all t s.t. $t_1 < \tau_1$ and $t_2 < \tau_2$. Now fix a $t_i > \tau_i$. Integrating both sides of (2) from t_i up to 1 yields:

$$f_0(t) = f_0(t_i,1) + f_0(1,t_{-i}) - f_0(1,1) \quad \forall t \text{ s.t. } t_i > \tau_i. \tag{21}$$

It follows immediately that for all t s.t. $t_1 > \tau_1$ and $t_2 > \tau_2$ the function $f_0(t)$ is given by (21). For any t s.t. $t_1 < \tau_1$ and $t_2 > \tau_2$ the expression for $f_0(t)$ obtained from (20) by setting $i = 1$ must be identical to the expression for $f_0(t)$ obtained from (21) by setting $i = 2$. This implies that $k_1(t_2) = 0$ for all $t_2 > \tau_2$. In the same manner, we can establish that $k_2(t_1) = 0$ for all $t_1 > \tau_1$. Thus, $f_0(t)$ is given by (21) for all t s.t. $t_i < \tau_i$ and $t_{-i} > \tau_{-i}$. and $\forall t$ s.t. $t_1 > \tau_1$ and $t_2 > \tau_2$. □

[22] Observe that the results follows even in the case where $t_2' < 1 - t_1$ and $t_2'' > 1 - t_1$. While the function $\partial^2 f_0(t_1,s_2)/\partial t_1 \partial t_2$ may display a discontinuity at $s_2 = 1 - t_1$, the functional value $\partial^2 f_0(t_1, 1-t_1)/\partial t_1 \partial t_2$ does not affect the value of the integral from t_2' up to t_2''.

DOI: 10.2202/1935-1704.1684

Proof of Lemma 3. The proof proceeds by considering an exhaustive list of cases. In each case, we examine for a given a type-pair $t \in T$ and non-empty subset $S \subseteq A$, if a decision rule that assigns strictly positive probability to only the elements of S is compatible with ex post welfare-maximization.

Case 1. $S = A$: $f_0(t) \in (0,1)$ and $f_i(t) \in (0,1)$ for all $i \in I$. Efficiency requires that all three alternatives in A generate the same level of ex post welfare:

$$\lambda_1(t) = \lambda_1(t)t_1 + \lambda_2(t)t_2 = \lambda_2(t) \Leftrightarrow t_1 + t_2 = 1.$$

I.e. an ex post classically efficient decision rule f can assign strictly positive probability to all three alternatives only for type-pairs t on the cross-diagonal in the unit square T.

Case 2. $S = \{a_0, a_i\}$ for some $i \in I$: $f_0(t) \in (0,1)$ and $f_i(t) = 1 - f_0(t)$. Efficiency requires that the two alternatives in S generate the same level of ex post welfare:

$$\lambda_i(t) = \lambda_i(t)t_i + \lambda_j(t)t_j \geq \lambda_j(t) \quad \text{for } i, j \in I,\ j \neq i. \tag{22}$$

If $t_i < 1$, then we can obtain from (22):

$$\lambda_i(t) = \lambda_j(t)t_j/(1-t_i) \geq \lambda_j(t) \Leftrightarrow t_1 + t_2 \geq 1.$$

If, instead, $t_i = 1$, then (22) can be satisfied only if $t_j = 0$. In summary, if an ex post classically efficient decision rule f assigns positive probability to only the elements of $S = \{a_0, a_i\}$ then: either $t_i \in [0,1)$, $t_j \in [0,1]$ and $t_1 + t_2 \geq 1$; or $t_i = 1$ and $t_j = 0$.

Case 3. $S = \{a_1, a_2\}$: $f_1(t) \in (0,1)$ and $f_2(t) = 1 - f_1(t)$. Efficiency requires that the two alternatives in S generate the same level of ex post welfare:

$$\lambda_1(t) = \lambda_2(t) \geq \lambda_1(t)t_1 + \lambda_2(t)t_2 \Leftrightarrow t_1 + t_2 \leq 1.$$

I.e. if, for some type-pair t, an ex post classically efficient decision rule f assigns strictly positive probability to the agents' favorite alternatives only, then t must be on or below the cross-diagonal.

Case 4. $S = \{a_0\}$: $f_0(t) = 1$. Efficiency requires that ex post welfare under the compromise exceeds welfare under each of the other two alternatives:

$$\lambda_1(t)t_1 + \lambda_2(t)t_2 \geq \lambda_i(t) \quad \forall i \in I. \tag{23}$$

Suppose first that the type-pair t is below the cross-diagonal in T: $t_1 + t_2 < 1$, which implies $t_1, t_2 < 1$ and $t_j/(1-t_i) < 1$ for all $i, j \in I$, $j \neq i$. We therefore obtain from (23):

$$\lambda_i(t) \leq \lambda_j(t)t_j/(1-t_i) \quad \forall i, j \in I,\ j \neq i.$$

Combining these two inequalities, we can write:

$$\lambda_1(t) \leq \lambda_2(t)t_2/(1-t_1) < \lambda_2(t) \leq \lambda_1(t)t_1/(1-t_2).$$

This, however, constitutes a contradiction. It is therefore *not* efficient to implement the compromise for any type-pair below the cross-diagonal.
Now suppose that the type-pair t is on or above the cross-diagonal in T: $t_1+t_2 \geq 1$. For any agent i with $t_i < 1$ we obtain from (23) an upper bound on i's welfare weight $\lambda_i(t)$:

$$\lambda_i(t) \leq \lambda_j(t)t_j/(1-t_i).$$

For any agent i with $t_i = 1$, there is no upper bound on the value of his welfare weight $\lambda_i(t)$. However, regardless of whether $t_i < 1$ or $t_i = 1$ ($i \in I$) we can always find a pair of welfare weights $\lambda_1(t)$ and $\lambda_2(t)$ s.t. - where applicable - the upper bounds implied by (23) are respected. In summary, an ex post classically efficient decision rule f may implement the compromise with probability 1 for any type-pair t on or above the cross-diagonal.

Case 5. $S = \{a_i\}$ for some $i \in I$: $f_i(t) = 1$. Efficiency requires that ex post welfare under agent i's favorite alternative exceeds welfare under both the compromise and agent j's favorite alternative:

$$\lambda_i(t) \geq \lambda_j(t) \text{ and } \lambda_i(t) \geq \lambda_i(t)t_i + \lambda_j(t)t_j \quad \text{for } j \in I,\ j \neq i. \tag{24}$$

It is easy to verify that there exist welfare weights $\lambda_i(t)$ and $\lambda_j(t)$ that satisfy the inequalities in (24): If $t_i = 1$ and $t_j = 0$, any pair of weights with $\lambda_i(t) \geq \lambda_j(t)$ satisfies (24). If, instead, $t_i < 1$ then any pair of weights with $\lambda_i(t) \geq \max\{\lambda_j(t), \lambda_j(t)t_j/(1-t_i)\}$ satisfies (24). In summary, if an ex post classically efficient decision rule f implements agent i's favorite alternative with probability 1 then: either $t_i \in [0,1)$ and $t_j \in [0,1]$; or $t_i = 1$ and $t_j = 0$. $\square$

We now prepare the grounds for the proof of Proposition 6. The proof employs the results of three lemmas, each stating a property that any strategy-proof ex post classically efficient decision rule must display. The first lemma establishes that if the decision rule f assigns the compromise strictly positive probability for some interior type-pair *above* the cross-diagonal of the unit square T, then f must assign strictly positive and constant probabilities to alternatives a_1 and a_2, resp., for all type-pairs *below* the cross-diagonal.

Lemma A.2 *Let f be an ex post classically efficient and strategy-proof decision rule for some pair of strictly positive functions $\lambda_1(\cdot)$ and $\lambda_2(\cdot)$. If either $f_0(t') > 0$ for some type-pair $t' \in T$, with $t'_1 + t'_2 > 1$ and $t'_i < 1$ for all $i \in I$, or $f_0(t') \in (0,1)$ for some $t' \in T$, with $t'_1 + t'_2 > 1$ and $t'_i < 1$ for at least one $i \in I$, then there exist* *numbers* $\bar{f}_1 \in (0,1)$ *and* $\bar{f}_2 = 1-\overline{f}_1$ *s.t.* $f_1(t) = \bar{f}_1$ *and* $f_2(t) = \bar{f}_2 = 1-\bar{f}_1$ *for all* $t \in T$ *s.t.* $t_1+t_2 < 1$.

Proof of Lemma A.2. We distinguish the two cases listed in the lemma:

Case 1. Fix a type-pair (t_1',t_2') s.t. $t_1'+t_2' > 1$ and $t_i' < 1$ for all $i \in I$. Suppose that $f_0(t_1',t_2') > 0$. By Lemma 3 there exists an agent j s.t. $f_j(t_1',t_2') = 0$ and $f_i(t_1',t_2') = 1-f_0(t_1',t_2')$ $(i,j \in I,\ j \neq i)$. W.l.o.g. suppose that $f_2(t_1',t_2') = 0$ and $f_1(t_1',t_2') = 1-f_0(t_1',t_2')$. Now consider the type-pairs (t_1',t_2') and (t_1',t_2'') with $0 < t_2'' < 1-t_1'$. By Lemma 3 we know that f assigns positive (but possibly zero) probability to alternatives a_1 and a_2 only. As f is strategy-proof, agent 2 of type t_2'' cannot benefit from misrepresenting his type as t_2':

$$u_2(t_1',t_2''|t_2'') \geq u_2(t_1',t_2'|t_2'') \Leftrightarrow f_2(t_1',t_2'') \geq f_0(t_1',t_2')t_2''. \tag{25}$$

Similarly, agent 2 of type t_2' cannot benefit from misrepresenting his type as t_2'':

$$u_2(t_1',t_2'|t_2') \geq u_2(t_1',t_2''|t_2') \Leftrightarrow f_0(t_1',t_2')t_2' \geq f_2(t_1',t_2''). \tag{26}$$

By combining (25) and (26), it follows that, at type-pair (t_1',t_2''), f assigns alternative a_2 a probability $f_2(t_1',t_2'')$ that is strictly between 0 and 1: $0 < f_0(t_1',t_2')t_2'' \leq f_2(t_1',t_2'')$ $\leq f_0(t_1',t_2')t_2' < 1$, where $f_0(t_1',t_2') > 0$. By Lemma 3, the remaining probability must be assigned to alternative 1: $f_1(t_1',t_2'') = 1-f_2(t_1',t_2'')$.

Case 2. Fix a type-pair (t_1',t_2') s.t. $t_1'+t_2' > 1$ and $t_i' < 1$ for at least one agent $i \in I$. The only aspect of this case that is not already covered by Case 1 is where $t_i' < 1$ and $t_j' = 1 (i,j \in I,\ j \neq i)$. W.l.o.g. let $t_2' = 1$ and $t_1' < 1$. Suppose that $f_0(t_1',t_2') \in (0,1)$. By Lemma 3, we know that $f_2(t_1',t_2') = 0$ and $f_1(t_1',t_2') = 1-f_0(t_1',t_2')$. Now consider the type-pairs (t_1',t_2') and (t_1',t_2'') with $0 < t_2'' < 1-t_1'$. Following the same incentive compatibility argument as in (25) and (26) of Case 1, it follows that, at type-pair (t_1',t_2''), f assigns alternative a_2 a probability $f_2(t_1',t_2'')$ that is strictly between 0 and 1: $0 < f_0(t_1',t_2')t_2'' \leq f_2(t_1',t_2'') \leq f_0(t_1',t_2')t_2' < 1$, where $t_2' = 1$ and $f_0(t_1',t_2') \in (0,1)$. By Lemma 3, the remaining probability must be assigned to alternative 1: $f_1(t_1',t_2'') = 1-f_2(t_1',t_2'')$.

The remainder of the proof applies to both Cases 1 and 2 above. Consider type-pairs (t_1',t_2'') and (t_1'',t_2''), with $t_1''+t_2' < 1$. At (t_1'',t_2'') f assigns positive probability only to alternatives a_1 and a_2. Strategy-proofness of f implies that agent 1 of type t_1' cannot benefit from misrepresenting his type as t_1'':

$$u_1(t_1',t_2''|t_1') \geq u_1(t_1'',t_2''|t_1') \Leftrightarrow f_1(t_1',t_2'') \geq f_1(t_1'',t_2''). \tag{27}$$

Similarly, agent 1 of type t_1'' cannot benefit from misrepresenting his type as t_1':

$$u_1(t_1'',t_2''|t_1'') \geq u_1(t_1',t_2''|t_1'') \Leftrightarrow f_1(t_1'',t_2'') \geq f_1(t_1',t_2''). \tag{28}$$

By combining (27) and (28), we obtain: $f_1(t_1'',t_2'') = f_1(t_1',t_2'') = 1-f_2(t_1',t_2'')$ and $f_2(t_1'',t_2'') = f_2(t_1',t_2'')$. Using the same logic, we can establish that $f_1(t_1'',t_2) =$

$f_1(t_1', t_2'') = 1 - f_2(t_1', t_2'')$ and $f_2(t_1'', t_2) = f_2(t_1', t_2'')$ for any $t_2 \in (t_2'', 1 - t_1'')$. Consequently f must assign the *same* probabilities $f_1(t_1, t_2) = f_1(t_1', t_2'') =: \bar{f}_1$ and $f_2(t_1, t_2) = f_2(t_1', t_2'') =: \bar{f}_2$ to alternatives a_1 and a_2, resp., for *every* type-pair (t_1, t_2) s.t. $t_1 + t_2 < 1$. □

The next lemma shows that if there is a type-pair (t_i', t_j') above the cross-diagonal and in the interior of the unit square T for which the decision rule f assigns positive probability to alternatives a_0 and a_i only, then f must assign the *same* probabilities to a_0 and a_i for *all* type-pairs (t_i', t_j) with $t_j > t_j'$.

Lemma A.3 *Let f be an ex post classically efficient and strategy-proof decision rule for some pair of strictly positive functions $\lambda_1(\cdot)$ and $\lambda_2(\cdot)$. If there exists a type-pair (t_i', t_j') with $t_i', t_j' < 1$ and $t_i' + t_j' > 1$ ($i, j \in I$ and $j \neq i$), s.t. $f_i(t_i', t_j') = 1 - f_0(t_i', t_j')$, then $f_0(t_i', t_j) = f_0(t_i', t_j')$ and $f_i(t_i', t_j) = 1 - f_0(t_i', t_j')$ for all $t_j > t_j'$.*

Proof of Lemma A.3. Suppose there is a type-pair (t_i', t_j') with $t_i', t_j' < 1$ and $t_i' + t_j' > 1$, for which decision rule f assigns positive probability to alternatives a_0 and a_i only: $f_0(t_i', t_j') \in [0, 1]$, $f_i(t_i', t_j') = 1 - f_0(t_i', t_j')$. Assume that for some type-pair (t_i', t_j) with $t_j > t_j'$, decision rule f assigns probability $f_0(t_i', t_j)$ to the compromise and the remaining probability $f_j(t_i', t_j) = 1 - f_0(t_i', t_j)$ to agent j's favorite alternative a_j. Now consider agent j of type t_j'. His utility from truthful revelation of his type is $u_j(t_i', t_j' | t_j') = f_0(t_i', t_j') t_j'$. Suppose agent j reports instead some type $t_j > t_j'$. This generates a utility of $u_j(t_i', t_j | t_j') = f_j(t_i', t_j) + f_0(t_i', t_j) t_j'$. The difference in agent j's utility between misrepresenting his type and truthful revelation of his type is:

$$\begin{aligned} & u_j(t_i', t_j | t_j') - u_j(t_i', t_j' | t_j') \\ = \;& f_j(t_i', t_j) + [f_0(t_i', t_j) - f_0(t_i', t_j')] t_j' \\ = \;& 1 - f_0(t_i', t_j) + [f_0(t_i', t_j) - f_0(t_i', t_j')] t_j'. \end{aligned} \tag{29}$$

Strategy-proofness of f requires that the utility difference in (29) be non-positive. Item (i) of Proposition 1 strategy-proofness implies that $f_0(t_i', \cdot)$ is non-decreasing: $f_0(t_i', t_j) \geq f_0(t_i', t_j')$ for all $t_j > t_j'$. We now distinguish three cases:

Case 1. $f_0(t_i', t_j') = 0$. In this case, the utility difference in (29) reduces to $u_j(t_i', t_j | t_j') - u_j(t_i', t_j' | t_j') = 1 - f_0(t_i', t_j)(1 - t_j')$. Due to the premise that $t_j' > 1 - t_i' > 0$, it is easy to verify that the utility difference is strictly positive for *all* values $f_0(t_i', t_j) \in [0, 1]$. This constitutes a contradiction to the premise that f is strategy-proof. Therefore, f must prescribe a probability distribution over $\{a_0, a_i\}$ for *all* type-pairs (t_i', t_j) with $t_j > t_j'$: $f_0(t_i', t_j) \in [f_0(t_i', t_j'), 1]$ and $f_i(t_i', t_j) = 1 - f_0(t_i', t_j)$.

Strategy-proofness requires furthermore that agent j of type t'_j cannot gain from pretending to be any type $t_j > t'_j$, and vice versa:

$$u_j(t'_i,t'_j|t'_j) \geq u_j(t'_i,t_j|t'_j) \Leftrightarrow f_0(t'_i,t'_j) \geq f_0(t'_i,t_j) \text{ for all } t_j > t'_j,$$

and

$$u_j(t'_i,t_j|t_j) \geq u_j(t'_i,t'_j|t_j) \Leftrightarrow f_0(t'_i,t_j) \geq f_0(t'_i,t'_j) \text{ for all } t_j > t'_j.$$

Together, these two incentive compatibility conditions imply: $f_0(t'_i,t_j) = f_0(t'_i,t'_j)$ for all $t_j > t'_j$. Therefore, f prescribes the *same* probability distribution over $\{a_0,a_i\}$ for *all* type-pairs (t'_i,t_j) with $t_j \geq t'_j$.

Case 2. $f_0(t'_i,t'_j) \in (0,1)$. Therefore, we have either $f_0(t'_i,t_j) = f_0(t'_i,t'_j)$ and $f_j(t'_i,t_j) = 1 - f_0(t'_i,t_j) > 0$, or $f_0(t'_i,t_j) > f_0(t'_i,t'_j)$ and $f_j(t'_i,t_j) = 1 - f_0(t'_i,t_j) \geq 0$. In both of these sub-cases, the utility difference in (29) is strictly positive. This constitutes a contradiction to the premise that f is strategy-proof. Employing the same argument as in Case 1, we can conclude that f must prescribe the same probability distribution over $\{a_0,a_i\}$ for *all* type-pairs (t'_i,t_j) with $t_j \geq t'_j$.

Case 3. $f_0(t'_i,t'_j) = 1$. In this case, by monotonicity in item (i) of Proposition 1, $f_0(t'_i,t_j) = f_0(t'_i,t'_j) = 1$ and $f_j(t'_i,t_j) = 1 - f_0(t'_i,t_j) = 0$. This implies immediately that f prescribes the *same* degenerate probability distribution over A for *all* type-pairs (t'_i,t_j) with $t_j \geq t'_j$. □

Lemma A.4 *Let f be an ex post classically efficient and strategy-proof decision rule for some pair of strictly positive functions $\lambda_1(\cdot)$ and $\lambda_2(\cdot)$. If there exists a type-pair (t'_i,t'_j), with $t'_i,t'_j < 1$ and $t'_i + t'_j > 1$ ($i,j \in I$, $j \neq i$), s.t. $f_i(t'_i,t'_j) = 1 - f_0(t'_i,t'_j)$, then $f_j(t_i,t'_j) = 0$ for all $t_i \in (1-t'_j,t'_i)$.*

Proof of Lemma A.4. By contradiction. Suppose there is a type-pair (t'_i,t'_j) with $t'_i,t'_j < 1$ and $t'_i + t'_j > 1$, for which decision rule f assigns positive probability to alternatives a_0 and a_i only: $f_0(t'_i,t'_j) \in [0,1]$, $f_i(t'_i,t'_j) = 1 - f_0(t'_i,t'_j)$. Assume now that at some type-pair (t_i,t'_j), with $t_i \in (1-t'_j,t'_i)$, the decision rule f assigns strictly positive probability to alternative a_j: $f_j(t_i,t'_j) > 0$. By Lemma 3 it must hold that $f_0(t_i,t'_j) + f_j(t_i,t'_j) = 1$. Now consider agent i of type t_i. His utility from truthful revelation of his type is $u_i(t_i,t'_j|t_i) = f_0(t_i,t'_j)t_i$. Suppose that agent i reports instead the type t'_i. This generates a utility of $u_i(t'_i,t'_j|t_i) = f_i(t'_i,t'_j) + f_0(t'_i,t'_j)t_i$.

The difference in agent i's utility between misrepresenting his type and truthful revelation of his type is:

$$\begin{aligned} & u_i(t_i', t_j'|t_i) - u_i(t_i, t_j'|t_i) \\ = \quad & f_i(t_i', t_j') + [f_0(t_i', t_j') - f_0(t_i, t_j')]t_i \\ = \quad & 1 - f_0(t_i', t_j') + [f_0(t_i', t_j') - f_0(t_i, t_j')]t_i. \end{aligned} \tag{30}$$

Strategy-proofness requires that the utility difference in (30) be non-positive. By item (i) of Proposition 1, monotonicity of f_0 implies that $f_0(t_i', t_j') \geq f_0(t_i, t_j')$. We now distinguish three cases:

Case 1. $f_0(t_i', t_j') = 0$. Monotonicity of f_0 implies that $f_0(t_i, t_j') = 0$, and therefore the utility difference in (30) is strictly positive. This is a contradiction to the premise that f is strategy-proof.

Case 2. $f_0(t_i', t_j') \in (0,1)$. In this case, $f_i(t_i', t_j) = 1 - f_0(t_i, t_j') > 0$ and therefore the utility difference in (30) is strictly positive. This is a contradiction to the premise that f is strategy-proof.

Case 3. $f_0(t_i', t_j') = 1$. In this case, the utility difference in (30) reduces to:

$$u_i(t_i', t_j'|t_i) - u_i(t_i, t_j'|t_i) = [1 - f_0(t_i, t_j')]t_i = f_j(t_i, t_j')t_i > 0,$$

where the last equality follows from the opening assumption that $f_0(t_i, t_j') + f_j(t_i, t_j') = 1$ and $f_j(t_i, t_j') > 0$. This, however, constitutes a contradiction to the premise that f is strategy-proof.

As each of the three cases above leads to a contradiction we can conclude that f must prescribe a probability distribution over $\{a_0, a_i\}$ for *all* type-pairs (t_i, t_j') with $t_i \in (1 - t_j', t_i')$. $\square$

Proof of Proposition 6. We start from the premise that f is an ex post classically efficient decision rule. Assume furthermore that f is strategy-proof. Lemma 3 in conjunction with Proposition 3 implies that f cannot be a binary decision rule. Therefore, there exists a type-pair $(\hat{t}_1, \hat{t}_2) \in T$ s.t. $f_0(\hat{t}_1, \hat{t}_2) \in (0,1)$. By Lemma 3 it must hold that $\hat{t}_1 + \hat{t}_2 \geq 1$ and $\hat{t}_i < 1$ for at least one agent $i \in I$ (as $f_0(\hat{t}_1, \hat{t}_2) < 1$ we cannot have $\hat{t}_1 = \hat{t}_2 = 1$). Due to the monotonicity of f_0 by item (i) of Proposition 1, we have $f_0(t_1, t_2) > 0$ for all type-pairs (t_1, t_2) with $t_1 \geq \hat{t}_1$ and $t_2 \geq \hat{t}_2$.

Step 1. Observe that regardless of whether the type-pair $(\hat{t}_1, \hat{t}_2)$ is a point on the cross-diagonal (where $\hat{t}_1 + \hat{t}_2 = 1$), a point in the interior of T (where $\hat{t}_1 + \hat{t}_2 > 1$ and $\hat{t}_i < 1$ for all $i \in I$), or a boundary point (where $\hat{t}_i < 1$ and $\hat{t}_j = 1$ for $i, j \in I$,

$j \neq i$), Lemma A.2 applies and guarantees that there are numbers $\bar{f}_1 \in (0,1)$ and $\bar{f}_2 = 1 - \bar{f}_1$ s.t. for all $(t_1,t_2) \in T$ with $t_1 + t_2 < 1$:

$$f_1(t_1,t_2) = \bar{f}_1 \text{ and } f_2(t_1,t_2) = \bar{f}_2. \tag{31}$$

This, in turn, implies that the probability of the compromise is strictly positive for all type-pairs above the cross-diagonal of the unit-square T:

$$f_0(t_1,t_2) > 0 \text{ for all } (t_1,t_2) \in T \text{ s.t. } t_1 + t_2 > 1. \tag{32}$$

To see this, consider any type-pair (t_1',t_2') with $t_1' + t_2' > 1$. Suppose that, contrary to (32), we have $f_0(t_1',t_2') = 0$. By Lemma 3 there must then be an agent $i \in I$ whose favorite alternative is chosen with probability 1. W.l.o.g. suppose that $f_1(t_1',t_2') = 1 - f_0(t_1',t_2') = 1$. Now consider agent 2 of type t_2'. His utility from truthful revelation of his type is $u_2(t_1',t_2',|t_2') = 0$. If agent 2 reports instead a type $t_2 < 1 - t_1'$, his utility is $u_2(t_1',t_2,|t_2') = \bar{f}_2 > 0$. This is a profitable deviation and constitutes a contradiction to the assumption that f is strategy-proof. Therefore, we must have $f_0(t_1',t_2') > 0$ as claimed in (32).

Step 2. Fix a type-pair (t_1',t_2') s.t. $\bar{f}_i < t_i' < 1$ for all $i \in I$. Assume w.l.o.g. that f assigns positive probability to alternatives a_0 and a_1: $f_0(t_1',t_2') \in [0,1]$ and $f_1(t_1',t_2') = 1 - f_0(t_1',t_2')$. Lemma A.3 implies for all (t_1',t_2) with $t_2 > t_2'$: $f_0(t_1',t_2) = f_0(t_1',t_2')$ and $f_1(t_1',t_2) = f_1(t_1',t_2') = 1 - f_0(t_1',t_2')$. Furthermore, Lemma A.4 implies for all (t_1,t_2) with $t_2 \geq t_2'$ and $t_1 \in (1-t_2,t_1')$: $f_1(t_1,t_2) = 1 - f_0(t_1,t_2)$, where $f_0(\cdot,t_2)$ is a nondecreasing function that takes values in $[0, f_0(t_1',t_2)]$ for every $t_2 \in [t_2',1]$.

Step 3. Fix agent 1's type at some $t_1'' = 1 - t_2'$. Having assumed that $t_2' > \bar{f}_2$, it follows that $t_1'' < \bar{f}_1$. We know from (31) that $f_2(t_1'',t_2) = \bar{f}_2$ for all $t_2 < t_2' = 1 - t_1''$. Also (by Lemma A.3), we have $f_0(t_1'',t_2) = f_0(t_1'',1)$ and $f_1(t_1'',t_2) = 1 - f_0(t_1'',1)$ for all $t_2 > t_2' = 1 - t_1''$. Strategy-proofness of f requires that agent 2 with a type above $1 - t_1''$ cannot gain by pretending to be any type $t_2 < 1 - t_1''$. In particular, for all $\varepsilon \in (0,t_1'']$:

$$\begin{aligned} & u_2(t_1'',1-t_1''+\varepsilon|1-t_1''+\varepsilon) \geq u_2(t_1'',t_2\,|1-t_1''+\varepsilon) \\ \Leftrightarrow \quad & f_0(t_1'',1)(1-t_1''+\varepsilon) \geq \bar{f}_2. \end{aligned} \tag{33}$$

Similarly, agent 2 with a type below $1 - t_1''$ cannot gain from pretending to be any type $t_2 > 1 - t_1''$. In particular, for all $\delta \in (0, 1 - t_1'']$:

$$\begin{aligned} & u_2(t_1'',1-t_1''-\delta|1-t_1''-\delta) \geq u_2(t_1'',t_2|1-t_1''-\delta) \\ \Leftrightarrow \quad & \bar{f}_2 \geq f_0(t_1'',1)(1-t_1''-\delta). \end{aligned} \tag{34}$$

In the limit as both $\varepsilon \downarrow 0$ and $\delta \downarrow 0$ we obtain from (33) and (34):

$$\bar{f}_2 = f_0(t_1'', 1)(1 - t_1'') \Leftrightarrow f_0(t_1'', 1) = \frac{\bar{f}_2}{1 - t_1''}. \tag{35}$$

As $t_1'' < \bar{f}_1$ the probability of the compromise $f_0(t_1'', 1)$ in (35) is well-defined. To see this, note that:

$$\begin{aligned} f_0(t_1'', 1) \leq 1 \quad &\Leftrightarrow \quad \bar{f}_2 \leq 1 - t_1'' \\ &\Leftrightarrow \quad t_1'' \leq 1 - \bar{f}_2 \\ &\Leftrightarrow \quad t_1'' \leq \bar{f}_1. \end{aligned}$$

Employing the same logic as above, we can conclude that $f_0(t_1, t_2) = \bar{f}_2/(1 - t_1)$ and $f_1(t_1, t_2) = 1 - f_0(t_1, t_2)$ for all (t_1, t_2) with $t_1 < t_1''$ and $t_2 > 1 - t_1''$. Note also that the incentive compatibility argument used here can be replicated to establish that decision rule f *must* assign to all type-pairs (t_1, t_2) with $t_1 \in (\bar{f}_1, t_1']$ and $t_2 \in (1 - t_1, \bar{f}_2)$ a probability distribution over alternatives a_0 and a_2.[23] Given this observation, we can employ Lemmas 5 and 6 to show furthermore that f must also prescribe a probability distribution over alternatives a_0 and a_2 for all type-pairs (t_1, t_2) with $t_1 > t_1'$ and $t_2 \in (1 - t_1, \bar{f}_2)$.

Step 4. Fix agent 2's type at $\tilde{t}_2$, with $\tilde{t}_2 < \bar{f}_2$. We know from (31) that $f_1(t_1, \tilde{t}_2) = \bar{f}_1$ for all $t_1 < 1 - \tilde{t}_2$. Also (by Lemma A.3), we have $f_0(t_1, \tilde{t}_2) = f_0(1, \tilde{t}_2)$ and $f_2(t_1, \tilde{t}_2) = 1 - f_0(1, \tilde{t}_2)$ for all $t_1 > 1 - \tilde{t}_2$. Strategy-proofness of f requires that agent 1 with a type above $1 - \tilde{t}_2$ cannot gain by pretending to be any type $t_1 < 1 - \tilde{t}_2$. In particular, for all $\varepsilon \in (0, \tilde{t}_2]$:

$$\begin{aligned} & u_1(1 - \tilde{t}_2 + \varepsilon, \tilde{t}_2 | 1 - \tilde{t}_2 + \varepsilon) \geq u_1(t_1, \tilde{t}_2 | 1 - \tilde{t}_2 + \varepsilon) \\ \Leftrightarrow \quad & f_0(1, \tilde{t}_2)(1 - \tilde{t}_2 + \varepsilon) \geq \bar{f}_1. \end{aligned} \tag{36}$$

Similarly, agent 1 with a type below $1 - \tilde{t}_2$ cannot gain from pretending to be any type $t_1 > 1 - \tilde{t}_2$. In particular, for all $\delta \in (0, 1 - \tilde{t}_2]$:

$$\begin{aligned} & u_1(1 - \tilde{t}_2 - \delta, \tilde{t}_2 | 1 - \tilde{t}_2 - \delta) \geq u_1(t_1, \tilde{t}_2 | 1 - \tilde{t}_2 - \delta) \\ \Leftrightarrow \quad & \bar{f}_1 \geq f_0(1, \tilde{t}_2)(1 - \tilde{t}_2 - \delta). \end{aligned} \tag{37}$$

[23] To see this, consider some type-pair $(\tilde{t}_1, \tilde{t}_2)$ with $\bar{f}_1 < \tilde{t}_1 < t_1'$ and $1 - \tilde{t}_1 < \tilde{t}_2 < \bar{f}_2$. Suppose at $(\tilde{t}_1, \tilde{t}_2)$ decision rule f assigns positive probability to a_0 and a_1. Then there exists a type $t_1 \in (\bar{f}_1, \tilde{t}_1)$ such that f prescribes the same probability distribution over alternatives a_0 and a_1 for all $t_2 > 1 - t_1$: $f_0(t_1, t_2) = f_0(t_1, 1)$ and $f_1(t_1, t_2) = 1 - f_0(t_1, 1)$. An analogous incentive compatibility argument as in Step 1 yields the requirement that $f_0(t_1, 1) = \bar{f}_2/(1 - t_1)$. This, however, cannot hold as the probability $f_0(t_1, 1)$ is at most 1, while the ratio $\bar{f}_2/(1 - t_1)$ is a number strictly greater than 1 for any $t_1 > \bar{f}_1$.

DOI: 10.2202/1935-1704.1684

In the limit, as both $\varepsilon \downarrow 0$ and $\delta \downarrow 0$, we obtain from (36) and (37):

$$\bar{f}_1 = f_0(1,\tilde{t}_2)(1-\tilde{t}_2) \Leftrightarrow f_0(1,\tilde{t}_2) = \frac{\bar{f}_1}{1-\tilde{t}_2}. \quad (38)$$

As $\tilde{t}_2 < \bar{f}_2$, the expression in (38) is well-defined. By the same logic, we obtain that for all (t_1,t_2) s.t. $t_2 < \bar{f}_2$ and $t_1 > 1-t_2$:

$$f_0(t_1,t_2) = \frac{\bar{f}_1}{1-t_2} \text{ and } f_2(t_1,t_2) = 1 - f_0(t_1,t_2). \quad (39)$$

Taking the limit of the probability of the compromise $f_0(t_1,t_2)$ in (39) as t_2 approaches $\bar{f}_2$ from below, we obtain for any given $t_1 > \bar{f}_1$:

$$\lim_{t_2 \uparrow \bar{f}_2} f_0(t_1,t_2) = \frac{\bar{f}_1}{1-\bar{f}_2} = 1.$$

Monotonicity of the function $f_0(t_1,\cdot)$, together with the fact that $f_0(t_1,t_2) \leq 1$ for all $(t_1,t_2) \in T$, implies that $f_0(t_1,\bar{f}_2) = 1$ for all $t_1 > \bar{f}_1$. Consequently, we have $f_0(t_1,t_2) = 1$ for all type-pairs with $t_1 > \bar{f}_1$ and $t_2 \geq \bar{f}_2$ (see Fig. 1 in Section 6).

Step 5. Now fix agent 1's type at $t_1' > \bar{f}_1$. Strategy-proofness of f requires that agent 2 with type $t_2' > \bar{f}_2$ cannot gain by pretending to be some type $\tilde{t}_2$, with $1-t_1' < \tilde{t}_2 < \bar{f}_2$, for which decision rule f assigns positive probability to alternatives a_0 and a_2 only:

$$\begin{aligned} & u_2(t_1',t_2'|t_2') \geq u_2(t_1',\tilde{t}_2|t_2') \\ \Leftrightarrow \quad & t_2' \geq f_2(t_1',\tilde{t}_2) + f_0(t_1',\tilde{t}_2)t_2' \\ \Leftrightarrow \quad & t_2' \geq 1 - f_0(t_1',\tilde{t}_2)(1-t_2') \\ \Leftrightarrow \quad & f_0(t_1',\tilde{t}_2) \geq 1. \end{aligned}$$

As the probability of the compromise $f_0(t_1',\tilde{t}_2)$ cannot exceed 1 it must be the case that $f_0(t_1',\tilde{t}_2) = 1$. This, however, contradicts (39) whereby $f_0(t_1',\tilde{t}_2) = \bar{f}_1/(1-\tilde{t}_2) < 1$. From this contradiction we can conclude that there exists *no* ex post classically efficient and strategy-proof non-binary decision rule. □

References

Apostol, T. M. (1974): *Mathematical Analysis*, Reading, MA, USA: Addison-Wesley.

Aswal, N., S. Chatterji, and A. Sen (2003): "Dictatorial domains," *Economic Theory*, 22, 45–62.

Barberà, S., A. Bogomolnaia, and H. van der Stel (1998): "Strategy-proof probabilistic rules for expected utility maximizers," *Mathematical Social Sciences*, 35, 89–103.

Bergemann, D. and S. Morris (2005): "Robust mechanism design," *Econometrica*, 73, 1521–1534.

Börgers, T. (2006): *Mechanism Design*, Ann Arbor, MI, USA: University of Michigan mimeo.

Börgers, T. and P. Postl (2009): "Efficient compromising," *Journal of Economic Theory*, 144, 2057–2076.

Chung, K.-S. and J. C. Ely (2006): "Ex-post incentive compatible mechanism-design," Northwestern University mimeo.

Chung, K.-S. and J. C. Ely (2007): "Foundations of dominant-strategy mechanisms," *Review of Economic Studies*, 74, 447–476.

Dutta, B., H. Peters, and A. Sen (2007): "Strategy-proof cardinal decision schemes," *Social Choice and Welfare*, 28, 163–179.

Freixas, X. (1984): "A Cardinal Approach to Straightforward Probabilistic Mechanisms," *Journal of Economic Theory*, 34, 227–251.

Gibbard, A. (1973): "Manipulation of voting schemes: A general result," *Econometrica*, 41, 587–601.

Gibbard, A. (1977): "Manipulation of voting schemes that mix voting with chance," *Econometrica*, 45, 665–681.

Green, J. R. and J.-J. Laffont (1979): *Incentives in Public Decision-Making*, Amsterdam, Netherlands: North Holland Publishing Co.

Holmström, B. and R. B. Myerson (1983): "Efficient and durable decision rules with incomplete information," *Econometrica*, 51, 1799–1819.

Laffont, J.-J. and E. Maskin (1980): "A differential approach to dominant strategy mechanisms," *Econometrica*, 48, 1507–1520.

Milgrom, P. R. (2004): *Putting Auction Theory to Work*, Cambridge, MA, USA: Cambridge University Press.

Milgrom, P. R. and I. R. Segal (2002): "Envelope theorems for arbitrary choice sets," *Econometrica*, 70, 583–601.

Mookherjee, D. and S. Reichelstein (1992): "Dominant strategy implementation of bayesian incentive compatible allocation rules," *Journal of Economic Theory*, 56, 378–399.

Satterthwaite, M. A. (1975): "Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions," *Journal of Economic Theory*, 10, 187–217.

Smith, D. (2010): "A prior free efficiency comparison of mechanisms for the public good problem," University of Michigan mimeo.

Wilson, R. (1987): "Game-theoretic analyses of trading processes," in T. F. Bewley, ed., *Advances in Economic Theory: Fifth World Congress*, Cambridge, UK: Cambridge University Press, 33–70.

Wilson, R. (1993): "Design of efficient trading procedures," in D. Friedman and J. Rust, eds., *The Double Auction Market: Institutions, Theories, and Evidence*, Reading, MA, USA: Addison-Wesley, 125–152.