Prevalence of Truthtelling and Implementation

Olivier Bochet* Norovsambuu Tumennasan †

April 10, 2017

Abstract

A decision rule exhibits the *(group) prevalence of truthtelling* if it is both (i) (group) strategy-proof and (ii) for any misreports of preferences leading to an outcome that differs from the one prescribed at the true preference profile, there always exists an agent (resp. a group of agents) that would benefit by reverting to truthtelling. When a decision rule satisfies (ii), we say that it is *(group-) resilient*. In general, (group-) resilience implies (group) strategy-proofness. Hence under resilience only item (ii) is required for the prevalence of truthtelling. We characterize both notions of resilience in terms of well-known conditions from the mechanism design literature. In particular the combination of *strategy-proofness* and *non-bossiness in welfare* is equivalent to group resilience. As such non-bossiness has a rather unexpected strategic implication. Individual resilience is more demanding and an extra condition *–outcome-rectangular property–* is needed for its characterization. We next provide some mechanism design foundations. Individual resilience is equivalent to secure implementation (Saijo et al., 2007), and a by-product is a new characterization of this concept based on and intuitive and easy-to-check condition. We tie group resilience to a new notion that we call *group secure implementation.* Importantly, our results show that most of the negative results on secure implementation uncovered in the literature vanish once coalitional moves are possible. We examine our findings for several models of interest.

1 Introduction

We introduce a new notion that we call the *prevalence of truthtelling.* We say that a decision rule exhibits the (group) prevalence of truthtelling if (i) it is (group) strategy-proof, and (ii) for any misreports of preferences leading to an outcome that differs from the one prescribed at the true preference profile, there always exists an agent (resp. a group

*Division of Social Science, New York University Abu Dhabi. *Email:* olivier.bochet@nyu.edu

†Department of Economics, Dalhousie University University. *Email:* norov@dal.ca

of agents) that would benefit by reverting to truthtelling.[1] Item (ii) is a new protective criterion that deals with situations where agents may end up at a preference misreport. When a (group) strategy-proof rule is responsive in the sense of (ii), truthtelling is focal since any misreport can be "signaled" by a simple reversion to truthtelling. If a rule satisfies the requirement in (ii), we say that it is *(group-) resilient*. Both notions of resilience are new and they both imply their counterpart non-manipulation notions, i.e. resilience implies strategy-proofness, and (when the domain is sufficientlly "large") group-resilience implies strategy-proofness. Therefore, item (i) in the prevalence of truthtelling is in general redundant.

From a practical point of view the prevalence of truthtelling is a desirable robustness requirement. However, from a theoretical perspective, if a decision rule is (group) strategy-proof, why should one be concerned that some preference misreports lead to outcomes which differ from the rule under the true preference profile? In the mechanism design literature, the (dominant strategy) revelation principle asserts that any decision rule that is implementable by some mechanism must be incentive compatible, i.e. truthtelling must be one of its direct revelation mechanism (dominant strategy) equilibrium. In that sense there is no loss of generality in restricting attention to direct mechanisms. A caveat with this approach is that direct revelation mechanisms typically admit many unwanted outcomes where some agents misreport their preferences. What happens when agents lie? We call an unwanted outcome, one that differs from the prescription of the rule at the true preference profile. There is a recent literature (Cason et al. (2006), Saijo et al. (2007)) underlining how agents may fail to identify and play their dominant strategies even if truthtelling is one of them, a feature which questions the salience of truthtelling. In particular both sets of authors observe that many strategy-proof rules of interest admit lot of unwanted Nash equilibrium. This features seems problematic as such rules typically perform poorly in practice, with agents failing to identify their dominant strategies. If the agents' behavior is best described as a mixture of dominant strategy and Nash equilibrium play (as the experimental results of see Cason et al. (2006) seem to suggest), the latter feature of strategy-proof rules may be of concern. In conjunction with the insights from Cason et al. (2006), Saijo et al. (2007) introduce the additional protective criterion of *secure implementation* which requires full double implementation in dominant strategies and Nash equilibrium. They identify a new condition, the *rectangular property* which together with strategy-proofness is both necessary and sufficient for secure implementation. Secure implementation puts some additional requirements on strategy-proof rules. More recently, Li (2016) introduces the new robustness notion of obvious strategy-proofness.[2] Li's notion is a robustness notion motivated by agents' cognitive limitations, arguing that some strategy-proof mechanisms are easier to understand than others.

[1]Often in the paper, we refer to decision rules simply as rules. This should cause no confusion.

[2]See also Pycia and Troyan (2016).

We are not concerned here with complexity or limitations on agents' part. However, our approach is yet another try to add some robustness requirement to (group) strategy-proof rules. Let us go back to the full implementation requirement. Does the latter requirement improve the salience of truthtelling? In a direct mechanism, the full implementation requirement imposes that whenever agents end up at a joint lie, there always exists a profitable deviation to an alternative preference report. Hence the type of deviations arising from a joint lie are typically unrestricted. In contrast, our robustness requirement is to impose one unique type of profitable deviation: whenever agents are stuck at a joint lie, there exists a simple deviation in the form of a reversion to truthtelling. This added robustness should make the truth salient in a direct mechanism. We believe that among all possible deviations, the reversion to truthtelling is the easiest to coordinate and least taxing mentally. In that sense, our motivation is also connected to Li's. Perhaps importantly at this stage, note that we are not imposing any type of behavioral departure from the standard pure payoff maximization, e.g. for instance we do not assume that some agents have a preference for honesty (see for instance Dutta and Sen (2012) or Saporiti (2014)). The prevalence of truthtelling comes out as a robustness requirement on rules, not one on agents' preferences.

We study in turn the two type of resilience conditions. Note that by definition individual resilience is stronger than group resilience. Indeed for the former requirement the size of the deviating group is imposed to be one. Our goal is manifold. First, we identify well-known conditions from the literature which are equivalent to our notions of resilience. Let us start with its group version. We find that group resilience is equivalent to the combination of strategy-proofness and non-bossiness in welfare, another well-know condition. Strategy-proofness only requires that truthtelling be a dominant strategy. In conjunction with non-bossiness in welfare, it imposes that truthtelling is prevalent and has therefore a somewhat unexpected strategic interpretation. In addition this equivalence shows that the seemingly very demanding group resilience condition is in fact no more demanding than requiring a pair of well-known properties which have appeared regularly in the literature. While strategy-proofness is central in the mechanism design literature, non-bossiness in welfare is typically used to achieve different objectives, in particular to provide structure and tractability to classes of rules in axiomatic characterization theorems.[3] There are interesting classes of group resilient rules. For instance, when preferences are strict the large class of trading cycles identified in Pycia and Ünver (forthcoming), all of which are efficient. In contrast, under efficiency the class of individual resilient rules contains only priority rules (Fujinaka and Wakayama (2011)). Note that under strict preferences group-resilience is equivalent to group strategy-proofness, hence the first requirement in the group prevalence of truthtelling is redundant. When a stock of a resource is to be divided among agents with single-peaked preferences (Sprumont

[3]See the excellent survey by Thomson (2016) for a discussion of non-bossiness.

(1991)), for any fixed path rules (Moulin (1999)), sequential allotment rules (Barbera et al. (1997)) or more broadly for any efficient rule, group-resilience and group strategy-proofness are equivalent. Once more the first requirement in the group prevalence of truthtelling is redundant. In contrast, under efficiency the class of individual resilient rules contains only variants of priority rules (Bochet and Sakai (2010)). Because the individual version of resilience imposes that group deviations can only be initiated by a single agent, strategy-proofness and non-bossiness in welfare are not sufficient for its characterization. The combination of the latter two with a condition called the *outcome rectangular property* (Saijo et al. (2007)) is equivalent to individual resilience. The latter turns out to be very demanding and the class of rules that satisfy resilience is for sure narrow as we exemplify in the paper.

Last but not least, our goal is also to provide some mechanism design foundations for the different resilience notions we have introduced. An important consequence of our earlier results is that individual resilience is equivalent to secure implementation introduced in Saijo et al. (2007). A direct by-product of that equivalence is a new characterization of secure implementation in the form of an intuitive and easy-to-check condition. Saijo et al. identifies the rectangular property as one of the necessary condition for secure implementation, along with strategy-proofness. The news delivered by secure implementation is mostly negative (Saijo et al. (2007), Bochet and Sakai (2010), Fujinaka and Wakayama (2011)). We tie back group resilience to a new implementation concept that we introduce, *group secure implementation*. The latter requires double implementation in dominant strategies and strong Nash equilibrium using direct revelation mechanisms. A decision rule is group secure implementable if it satisfies both group strategy-proofness and a condition called the group reversal property. The latter is a weaker requirement than group resilience as it does not pin down the type of deviations that must occur when agents misreport. Group strategy-proofness is equivalent to group resilience when preferences are strict, and it is implied by group resilience when the domain is rich, or when it contains only single-peaked preferences. Hence most of the negative results on secure implementation disappears once one allows for coalitional moves. In fact, if the domain of preferences is "large" enough, any group resilient rule is group secure implementable.

Our findings on the group prevalence of truthtelling are particularly interesting in environments in which agents communicate during or prior to submitting their reports. Indeed pre-play communication is usually not prohibited in many applications. For instance, many cities in the US allocate students to schools using some centralized allocation mechanism. Here, students or parents can discuss their preference reports before submitting to the relevant authorities. Patients can also freely discuss their organ preferences prior submitting them to donor-patient matching systems. In such environments, it is unreasonable to expect agents to be stuck at a Nash equilibrium that is vulnerable to coalitional deviations.[4] Our findings underline also the fragility and perhaps lack of credibility of the failure of the outcome rectangular property.

The paper proceeds as follows. We introduce the model and some of the necessary definitions in Section 2. In Section 3, we characterize both notions of resilience in the form of well-known conditions from the mechanism design literature. We provide some background on secure implementation in Section 3 and examples highlighting the intuition conveyed in Bochet and Sakai (2010). Section 4 deals with the strategic foundations of group resilience and its connection with group strategy-proofness. Section 5 provides a discussion and extension of our results. We provide some concluding remarks in Section 6. Some of our proofs are relegated to the Appendix.

2 Setup

Let $N = \{1, \ldots, n\}$ be a *set of agents.* Let $A = A_1 \times \ldots \times A_n$ be a *set of alternatives.* For $i \in N$, we call A_i *agent i's individual set of alternatives.* We assume that if $A_i \subseteq \mathbb{R}^m$ and $|A_i| = \infty$, then A_i is convex. Let $x = (x_1, \ldots, x_n) \in A$ be an alternative and $\mathbb{1} \equiv (1, \ldots, 1) \in \mathbb{R}^n$. If alternative x is such that for all $i, j \in N$, $x_i = x_j = \alpha$, then we denote $x = \alpha\mathbb{1}$. Next, let $F \subseteq A$ be the *set of feasible alternatives.* If for all $x \in F$ there exists α such that $x = \alpha\mathbb{1}$, then the set of feasible alternatives F *determines a public goods economy.* Otherwise, the set of feasible alternatives F *determines an economy with at least one private goods component.* Hence, our model encompasses public and private goods economies.

To fix ideas, let us give two examples. It will be clear from these examples that given the set A of alternatives, the set F of feasible alternatives fully determines whether we are working with a public or private goods model. Note that the Cartesian product notation we use for the set of alternatives is for notational convenience only; none of our results require it.

Example 2.1. *Let* $A = \{a_1, \ldots, a_n\} \times \ldots \times \{a_1, \ldots, a_n\}$.
Public Goods Model: Suppose that the agents have to choose one candidate out of the set $\{a_1, \ldots, a_n\}$ *of possible candidates. Then,* $F = \{x \in A : \text{for all } i, j \in N,\ x_i = x_j\}$.
Private Goods Model: On the other hand, if agents have to allocate the set of indivisible objects or tasks $\{a_1, \ldots, a_n\}$ *among themselves, then* $F = \{x \in A : \text{for all } i, j \in N,\ x_i \neq x_j\}$. ◇

Example 2.2. *Let* $A = [0, 1] \times \ldots \times [0, 1]$.
Public Goods Model: Suppose that the agents have to choose a single point in the interval [0,1] that everyone will consume without rivalry, e.g., a public facility on a street (see

[4]Indeed, this fact is established in experimental studies such as Moreno and Wooders (1998).

Moulin, 1980). Then, $F = \{x \in A : \textit{for all } i, j \in N,\ x_i = x_j\}$.
Private Goods Model: On the other hand, if agents have to choose a division of one unit of an infinitely divisible good among themselves (see Sprumont, 1991), then feasibility is determined by the size of the resource and $F = \{x \in A : \textit{for all } i \in N,\ x_i \geq 0 \textit{ and } \sum_{i \in N} x_i = 1\}$. ◇

For all $i \in N$, *preferences* are represented by a complete, reflexive, and transitive binary relation R_i over A_i. As usual, for all $a, b \in A_i$, $a\, R_i\, b$ is interpreted as "i weakly prefers a to b", $a\, P_i\, b$ as "i strictly prefers a to b", and $a\, I_i\, b$ as "i is indifferent between a and b". Preferences R_i over A_i are *strict* if for all $a, b \in A_i$, $a\, R_i\, a$ implies $a\, P_i\, b$ or $a = b$. Given $i \in N$, and preference relation R_i, $p(R_i) = \{a \in A_i : aR_ib \text{ for all } b \in A_i\}$. The set $p(R_i)$ is the set of elements of A_i that are top-ranked under R_i by agent i.

For public goods models, preferences R_i over the individual set of alternatives A_i can easily be extended to preferences over the set of alternatives A (since each agent consumes the same public alternative). Whenever our model captures a private goods component, we assume that agents only care about their own consumption. Then, for both public and private goods models, we can easily extend preferences R_i over the individual set of alternatives A_i to preferences over the set of alternatives A (both preference relations only depend on agent i's consumption in A_i). Therefore, from now on, we use R_i to describe agent i's preferences over A_i as well as over A, i.e., we use both notations $x\, R_i\, y$ and $x_i\, R_i\, y_i$. Note that for private goods models, strict preferences over A_i do not need to be strict over A.

For all $i \in N$, we call a set of preferences over A_i, denoted by $\mathcal{R}_i$, a *preference domain.* Let $\mathcal{R}^N \equiv \Pi_{i \in N} \mathcal{R}_i$ be the *domain of preference profiles.* A typical preference profile is $R = (R_i)_{i \in N}$ such that for all $i \in N$, $R_i \in \mathcal{R}_i$. Profile $R \in \mathcal{R}^N$, is often written as (R_i, R_{-i}), where $R_{-i} = (R_1, ..., R_{i-1}, R_{i+1}, ..., R_n)$. For a given preference profile $R \in \mathcal{R}^N$, we use the usual notations that $R_S \equiv (R_i)_{i \in S}$ and $R_{-S} \equiv (R_j)_{j \notin S}$. Likewise, for each $S \subset N$, we let $\mathcal{R}^S$ be the domain of preference profiles for S. Similar notations are used for $\mathcal{R}$.We say that alternative x *weakly dominates* y via group $S \subset N$ at profile R if $x\ R_i\ y$ for all $i \in S$ and $x\ P_j y$ for at least one $j \in N$. If everyone in S strictly prefers x to y then x *dominates* y. The notations wdom$[R, S]$ and dom$[R, S]$ denote the weak dominance and dominance at profile via group S.

We now define several preference domains that some of our results will cover.

Strict preference domain: Preferences R_i on $A_i \subseteq \mathbb{R}$ are *strict* if for all $x_i, y_i \in A_i$, $x_i\, R_i\, y_i$ implies $x_i\, P_i\, y_i$ or $x_i = y_i$. We say that a domain of preference profiles $\mathcal{R}^N$ is the domain of strict preferences if for each $i \in N$, each $\mathcal{R}_i$ consists of all the possible strict preferences over A_i.

Single-peaked preference domain: Preferences R_i on $A_i \subseteq \mathbb{R}$ are *single-peaked* if there exists a point $p(R_i) \in A_i$ such that for all $x_i, y_i \in A_i$ satisfying either $y_i < x_i \leq p(R_i)$

or $p(R_i) \leq x_i < y_i$, $x_i \; P_i \; y_i$. We say that a domain of preference profiles $\mathcal{R}^N$ is the domain of single-peaked preferences if for each $i \in N$, each $\mathcal{R}_i$ consists of all the possible single-peaked preferences over A_i. For each $R \in \mathcal{R}^N$, we let $p(R) = (p(R_1), ..., p(R_n))$ denote the profile of peaks at preference profile R.

Let the set of alternatives A, the set of feasible alternatives F and the domain of preference profile $\mathcal{R}^N$ be given. A *decision rule* $f : \mathcal{R}^N \to F$ is a function that assigns an alternative $f(R) \in F$ for each preference profile $R \in \mathcal{R}^N$. For each $R \in \mathcal{R}^N$, and each $i \in N$, we let $f_i(R)$ be what is alloted to i at $f(R)$. Obviously if the model is a public good model, $f_i(R) = f_j(R)$ for each $i, j \in N$.

A mechanism (or game form) is a pair $\Gamma = ((M_i)_{i \in N}, g)$ where M_i is agent i's message set, and $g : \prod_{i \in N} M_i \to F$ is the outcome function mapping each message profile to a feasible alternative. For each $R \in \mathcal{R}^N$, the pair (Γ, R) is a game in which the set of players is N, the set of strategy profiles is $M = \prod M_i$, and each player i's payoff is $g(m)$ where $m = (m_i)_{i \in N}$ is a message profile. In the paper we deal only with a special class of mechanisms, the *direct revelation mechanisms*. Given a decision rule f, the direct mechanism associated to f is $\Gamma^* = (\mathcal{R}^N, f)$. Whenever we make references to mechanisms, we confine our attention to pure strategies. We now introduce several important definitions which are used repeatedly throughout the paper.

Nash Equilibrium: Fix a decision rule f. A message profile $\tilde{R}$ is a Nash equilibrium of $\Gamma^* = (\mathcal{R}^N, f)$ at profile $R \in \mathcal{R}^N$ if for each $i \in N$, $f_i(\tilde{R}) \; R_i \; f_i(R'_i, \tilde{R}_{-i})$ for each $R'_i \in \mathcal{R}_i$. For each $R \in \mathcal{R}^N$, let $NE(\Gamma^*, R)$ be the set of Nash equilibria of (Γ^*, R).

Dominant strategies: Fix a decision rule f. A message profile $\tilde{R}$ is a (weakly) dominant message of $\Gamma^* = (\mathcal{R}^N, f)$ at profile $R \in \mathcal{R}^N$, if for each $i \in N$, $R'_i \in \mathcal{R}_i$, and $\tilde{R}_{-i}$, $f_i(R_i, \tilde{R}_{-i}) \; R_i \; f_i(R'_i, \tilde{R}_{-i})$. For each $R \in \mathcal{R}^N$, let $DS(\Gamma^*, R)$ be the set of dominant strategies of (Γ^*, R).

As we mentioned in the introduction, preplay communication among participating agents is allowed in many practical mechanisms. Once pre-play communication is allowed, agents will not get stuck at any Nash equilibrium which is susceptible to coalitional deviations. In this paper, we assume that agents coordinate at a strong Nash equilibrium – a strategy that is immune to coalitional deviations (Aumann, 1959).

Strong Nash Equilibrium: Fix a decision rule f. A message profile $\tilde{R}$ is a strong Nash equilibrium of $\Gamma^* = (\mathcal{R}^N, f)$ at profile $R \in \mathcal{R}^N$, if there exists no R'_S such that $f(R'_S, \tilde{R}_{-S}) \, \mathrm{wdom}[R, S] f(\tilde{R})$. For each $R \in \mathcal{R}^N$, let $SNE(\Gamma^*, R)$ be the set of strong Nash equilibria of (Γ^*, R).

Strategy-Proofness: A decision rule f satisfies strategy-proofness if for each $R \in \mathcal{R}^N$, $i \in N$ and $R'_i \in \mathcal{R}_i$, $f_i(R) \; R_i \; f_i(R'_i, R_{-i})$.

We will introduce several additional definitions as we proceed. We do so to keep some of the less standard definitions close to where they are being discussed in the text. In particular the two new properties of resilience of a decision rules are deferred to the next section.

3 On the Prevalence of Truthtelling

In the mechanism design literature, the (dominant strategy) revelation principle asserts that any decision rule that is implementable by some mechanism must be incentive compatible, i.e. truthtelling must be one of its direct revelation mechanism (dominant strategy) equilibrium. In that sense there is no loss of generality to restrict attention to direct mechanisms. A caveat with this approach is that such a direct revelation mechanism typically admits many unwanted equilibria where some agents may coordinate on some joint lie –preference misreports. What happens when agents lie? We call an unwanted outcome, one that differs from the prescription of the rule at the true preference profile. If a rule admits unwanted outcomes, one may be concerned about the applicability of such a rule using its direct revelation mechanism. A way out is simply to impose full implementation in dominant strategies. This typically adds additional requirements on strategy-proof rules. In a direct mechanism, the full implementation requirement imposes that all dominant strategy equilibria deliver the right outcome under the true preference profile. There is a recent literature discussing how agents may fail to identify their dominant strategies, a feature which questions the salience of truthtelling. Several avenues have been emphasized. For instance, Saijo et al. (2007) argue that many strategy-proofs are plagued with unwanted Nash equilibria that delivers the wrong outcome. This features seems problematic as such rules typically perform poorly in practice, with agents failing to identify their dominant strategies –see also the extensive literature following Saijo et al. (2007). More recently, Li (2016) introduces the new robustness notion of obvious strategy-proofness. Li's notion is a robustness notion motivated by agents' cognitive limitations, arguing that some strategy-proof mechanisms are easier to understand than others.

We are not concerned here with complexity or limitations on agents' part. However, our approach is yet another try to add some robustness requirement to strategy-proof rules. Let us go back to the full implementation requirement. Does the latter requirement improve the salience of truthtelling? In this paper, we want to focus on the prevalence of truthtelling and conditions which makes the truth salient in a direct mechanism. We say that a rule exhibits the prevalence of truthtelling if (i) it is strategy-proof, and (ii) for any misreports of preferences leading to a different outcome, there always exists a group of agents (resp. an agent) that would benefit by reverting to truthtelling. Hence not only should a rule be incentive compatible, but our robustness requirement is to impose one unique type of profitable deviation: whenever agents are stuck at a joint lie, there exists a simple deviation in the form of reversion to truthtelling.

If a rule satisfies the latter requirement, we say that it is (group-) resilient. The notion of resilience is new and it implies in general (group) strategy-proofness. As such, item (i) in the prevalence of truthtelling is redundant once (ii) is imposed. When a rule satisfies (group) resilience, truthtelling is prevalent. Not only agents have an incentive to tell the truth no matter what the others are doing, but in addition any joint lie can be "signaled" by a simple reversion to truthtelling. From a practical point of view the prevalence of truthtelling is paramount.

We now are ready to define the properties of individual and group-resilience –the main focus of this paper. As in any mechanism design problem we do not want any outcome not prescribed by the rule in question arise in the direct revelation game. This means that someone or some group depending on whether the communication is allowed among agents should deviate from a report that leads to a "bad" outcome. We believe that among all possible deviations, the reversion to truthtelling is the easiest to coordinate and least taxing mentally. In that sense, our motivation is also close to Li's. With this in mind, we state our notions of resilience below.

Definition 3.1. [**Resilience**] A decision rule f satisfies resilience if whenever $f(R) \neq f(\tilde{R})$ for some $R, \tilde{R} \in \mathcal{R}^N$, there must exist $i \in N$ such that $f_i(R_i, \tilde{R}_{-i})\ P_i\ f_i(\tilde{R})$.

Definition 3.2. [**Group-Resilience**] A decision rule f satisfies group-resilience if whenever $f(R) \neq f(\tilde{R})$ for some $R, \tilde{R} \in \mathcal{R}^N$, there must exist $S \subseteq N$ such that $f(R_S, \tilde{R}_{-S})\,\mathrm{dom}[R, S]\ f(\tilde{R})$.

Resilience says that whenever a preference report delivers an unwanted outcome in the direct revelation game, some agent must strictly benefit by reverting to truthtelling. For this notion, only unilateral reversion to truthtelling by agents are considered, which fits with the idea that agents are not allowed to communicate before they submit their preferences. However, if communication is allowed among agents, some may be able to coordinate away from an unwanted outcome in which case group-resilience is the appropriate notion. Note that group-resilience is the weaker of the two resilience notions: group resilience does not specify the size of deviating groups from a bad outcome while resilience does.

We now establish the connection between the resilience concepts and strategy-proofness. Specifically, each resilience concept implies strategy-proofness.

Lemma 3.3. *If a decision rule f satisfies group-resilience, then f satisfies strategy-proofness.*

Proof. In contrast to the lemma suppose that f is not strategy-proof. This means that there must exist $i \in N$, $R \in \mathcal{R}^N$ and $\tilde{R}_i \in \mathcal{R}_i$ such that

$$f(\tilde{R}_i, R_{-i})\ P_i\ f(R). \tag{1}$$

Consequently, $f(\tilde{R}_i, R_{-i}) \neq f(R)$. Then by group-resilience, there must exist S such that $f(R_S, R_{-S})\ P_j\ f(\tilde{R}_i, R_{-i})$ for all $j \in S$. Because the preferences for each $j \neq i$ are the same under both (R_S, R_{-S}) and $(\tilde{R}_i, R_{-i})$, it must be that $i \in S$. Hence, $f(R_S, R_{-S}) = f(R)\ P_i\ f(\tilde{R}_i, R_{-i})$, a contradiction with (1). □

Note that strategy-proofness implies neither resilience nor its group version. This fact will become clear next when we show that the resilience concepts are equivalent to some combinations of strategy-proofness and other widely used concepts in the literature. We start with the characterization of the individual version of resilience. We also discuss there the mechanism design foundations of resilience. Covering its group version takes more time. While the characterization of group resilience is straightforward, a lot needs to be discussed under the light of several prominent models used in the mechanism design literature.

3.1 Resilience

We first close the gap between resilience and strategy-proofness. For this we introduce below the *rectangular property*, a condition first introduced by Saijo et al. (2007).

Rectangular Property: A decision rule f satisfies the rectangular property if for each $R, \tilde{R} \in \mathcal{R}^N$ with $f(R_i, \tilde{R}_{-i})\ I_i\ f(\tilde{R})$ for all $i \in N$, we have $f(R) = f(\tilde{R})$.

Theorem 3.4. *A decision rule f satisfies resilience if and only if f satisfies both strategy-proofness and the rectangular property.*

Proof. We already know that resilience implies strategy-proofness. It is clear that resilience implies the rectangular property. For this, consider R and $\tilde{R}$ such that for each $i \in N$, $f_i(R_i, \tilde{R}_{-i})\ I_i\ f_i(\tilde{R})$. If $f(R) \neq f(\tilde{R})$, then resilience contradicts the previous indifferences. Hence $f(R) = f(\tilde{R})$.

Next we show that the combination of strategy-proofness and the rectangular property implies resilience. Pick R and $\tilde{R}$. By the rectangular property if $f_i(R_i, \tilde{R}_{-i})\ I_i\ f_i(\tilde{R})$ for each $i \in N$, then $f(R) = f(\tilde{R})$ so resilience holds. Suppose there exists $i \in N$ for whom the previous indifference does not hold. By strategy-proofness, $f_i(R_i, \tilde{R}_{-i})\ R_i\ f_i(\tilde{R})$. Hence $f_i(R_i, \tilde{R}_{-i})\ P_i\ f_i(\tilde{R})$ and resilience holds. □

As emphasized in Saijo et al. (2007), many strategy-proof rules of interest violate the rectangular property. Consequently, resilience is a rather demanding condition. We present here cases where the failure of the rectangular property is salient. We focus on three different examples. The first two show how the failure of resilience are not robust once coalitional moves are possible. The last example nevertheless shows that for some models even group resilience is out of reach.

Example 3.5. *Structure and coalitional instability of bad Nash equilibria I*
Let $F = \{x \in [0,\Omega]^n$: for all $i \in N$, $x_i \geq 0$ and $\sum_{i \in N} x_i = \Omega\}$ where $\Omega \in \mathbb{R}_{++}$. The preference domain is the one of single-peaked preferences, $\mathcal{R}^{sp}$, over $[0,\Omega]$. Then $(\mathcal{R}^{sp}, F)$ determine the Sprumont Model (Sprumont, 1991) of division under single-peaked preferences. A rule that is central in this model is the so-called *Uniform rule* f^U, defined for each $R \in \mathcal{R}^N$ and each $i \in N$ as,

$$f^U(R) = \begin{cases} min\{p(R_i), \lambda\} \text{ if } \sum\limits_{i \in N} P(R_i) \geq \Omega \\ max\{p(R_i), \lambda\} \text{ if } \sum\limits_{i \in N} P(R_i) \leq \Omega \end{cases}$$

where λ solves $\sum\limits_{i \in N} f^U(R) = \Omega$.

It is well-known that the uniform rule is strategy-proof. However, it fails to be resilient as shown below.

Let $n = 3$, $\Omega = 6$ and pick $R \in \mathcal{R}^N$ with peak profile $p(R) = (1,2,4)$. Consider f^U on the domain $\mathcal{R}^{sp}$ and let us show that f is not resilient. For instance, consider $\tilde{R}$ with $p(\tilde{R}) = (2,2,2)$. By reverting to truthtelling unilaterally, neither agent 1 nor 3 can change the outcome and f^U is therefore not resilient. Notice that $f^U(R) = (1,2,3)$ Pareto dominates $f^U(\tilde{R})$ at R. Agents 1 and 3 have the joint profitable deviation of simply reporting their true preferences so that the true uniform allocation is obtained. One can verify that there are an infinity of joint misreports at which resilience fails –all of these are in fact Nash equilibrium reports– which are described by the following sets,

$$\{\tilde{R} \in \mathcal{R}^N : 1 < p(\tilde{R}_1) \leq 2 = p(\tilde{R}_2) \leq p(\tilde{R}_3) < 4, \sum p(\tilde{R}_i) = 6\}$$
$$\{\tilde{R} \in \mathcal{R}^N : p(\tilde{R}_1) = 2, p(\tilde{R}_2) = 2, p(\tilde{R}_3) \leq 2\}$$
$$\{\tilde{R} \in \mathcal{R}^N : p(\tilde{R}_1) \geq 2, p(\tilde{R}_2) = 2, p(\tilde{R}_3) = 2\}.$$

Observe here that for each of these reports, agents 1 and 3 improve by jointly reverting to truthtelling. We will later show that the uniform rule is group-resilient. ◇

Example 3.6. *Structure and coalitional instability of bad Nash equilibria II*
Let $A_i = \{h_1, ..., h_n\}$ for each $i \in N$ and $F = \{x = (x_1, ..., x_n) \in A : x_i \neq x_j$ for each $i \neq j\}$, i.e., F determines a private good economy with indivisible goods, as in example 2.1. Let the preference domain $\mathcal{R}^N$ be the set of strict preferences over $\{h_1, ..., h_n\}$. Pick f^{TTC} to be the top-trading cycle rule in which each agent i is endowed with object h_i.[5]

[5] $f^{TTC}(R)$ is determined according to the following process:

Let $N = \{1, 2, 3\}$ and $A = \{h_1, h_2, h_3\}^3$. Let the set of preferences be as follows:

R_1	R_2	R_3	$\tilde{R}_1$	$\tilde{R}_2$
h_2	h_1	h_2	h_1	h_3
h_1	h_3	$\vdots$	$\vdots$	$\vdots$
h_3	h_2			

Observe here that the TTC rule f^{TTC} gives the following allocations:

$$f^{TTC}(R_1, R_2, R_3) = (h_2, h_1, h_3)$$
$$f^{TTC}(\tilde{R}_1, R_2, R_3) = f^{TTC}(R_1, \tilde{R}_2, R_3) = f^{TTC}(\tilde{R}_1, \tilde{R}_2, R_3) = (h_1, h_3, h_2)$$

When the state is $R = (R_1, R_2, R_3)$, the dominant strategy equilibrium is (R_1, R_2, R_3), i.e., truthtelling. Consider now the joint misreport $(\tilde{R}_1, \tilde{R}_2, R_3)$ which results in (h_1, h_3, h_2). Clearly, agent 3 has no incentive to unilaterally deviate as she obtains her top choice under this preference report. If either agent 1 or 2 unilaterally deviates from $(\tilde{R}_1, \tilde{R}_2, R_3)$ to truthtelling, then the outcome remains (h_1, h_3, h_2). Combining this with the result that truthtelling is a weakly dominant strategy for each agent, none of agents 1 and 2 have an incentive to deviate unilaterally from $(\tilde{R}_1, \tilde{R}_2, R_3)$. Thus, resilience is violated at $(\tilde{R}_1, \tilde{R}_2, R_3)$ –the latter is in fact a Nash equilibrium in the direct revelation mechanism of f^{TTC}. However, agents 1 and 2 jointly reverting from $(\tilde{R}_1, \tilde{R}_2)$ to (R_1, R_2) (while agent 3 reports R_3) leads to allocation (h_2, h_1, h_3). This is a profitable deviation for both agents. An important difference compared to the previous example is that the allocation under $(\tilde{R}_1, \tilde{R}_2, R_3)$ is not Pareto comparable with the one obtained under the report (R_1, R_2, R_3). Hence following the coalitional deviation by agents 1 and 2, agent 3 is worse-off since he was getting his top choice h_2 under the report $(\tilde{R}_1, \tilde{R}_2, R_3)$. ◇

Our last example shows that some decision rules of interest are neither resilient nor group-resilient. In particular our example shows that auction models will remain out of reach even if one considers group resilience.

Example 3.7. *Survival of bad Nash equilibria*

Step m: Each agent who have not been allocated an object in the previous steps points to the agent who owns her/his most preferred object among those which are not assigned to any agent yet. There exist at least one cycle of agents $\{i_1, \cdot, i_k\}$ such that each i_l where $l < k$ points to i_{l+1} while i_k points to i_1. Under the TTC rule, each agent in a cycle is allocated the object of the agent to whom she points.

The above process is terminated once every agent is allocated an object.

Let $F = \{(x,t) \in \{0,1\}^n \times \mathbb{R}^n : \sum_{i\in N} x_i = 1\}$. The feasible set F stands for a model in which there is one object to be given to one out of the n agents, and monetary transfers are possible. Let rule f^V be the Vickrey rule, i.e., the second price auction. Let $\mathcal{R}^N = \mathbb{R}_+$. For each $i \in N$, each preference relation R_i is indexed by a real number that stands for the valuation agent i attaches to the object. With a slight abuse of notation, each R_i is such that for (x,t), (x,t') with $t > t'$, then $(x_i, t_i)P_i(x_i, t'_i)$. Hence, preferences are said to be quasi-linear. For each $R \in \mathcal{R}^N$, $f(R) = (x,t) \in F$ with (i) $x_i = 1$ if $R_i = \max_{j\in N} R_j$, and $x_i = 0$ otherwise, (ii) $t_i = max_{j\neq i} R_j$ if $x_i = 1$, and $t_i = 0$ otherwise. While the Vickrey rule is strategy-proof, it fails to be resilient as shown below.

Let $n = 2$ and fix a preference profile R with $R_1 > R_2$. The joint report (R_1, R_2) is a weakly dominant strategy: agent 1 receives the object and pays $t_2 = R_2$ while agent 2 pays nothing. However there is an infinity of joint misreports at which resilience fails –all of these are in fact Nash equilibrium reports– which are described by the following set,

$$\{(\tilde{R}_1, \tilde{R}_2) \in \mathcal{R}^N : 0 \leq \tilde{R}_1 \leq R_2 \text{ and } R_1 \leq \tilde{R}_2\}$$

Because agent 2 gets a non-negative payoff at R when the report is $(\tilde{R}_1, \tilde{R}_2)$, this joint lie is a Nash equilibrium at R. Within the above set, reverting to truthtelling from $(\tilde{R}_1, \tilde{R}_2) = (0, \tilde{R}_2)$ jointly is not profitable for both agents. $\diamond$

Based on the series of examples, it is clear that resilience is a very demanding condition. From Examples 3.5 and 3.6 two conclusions may however be drawn. The first and most important one is that, for many models and rules of interest there may exist a coalition of agents which can benefit by reverting to truthtelling even when resilience is violated. Hence, in many models, the requirement imposed by resilience is too strong if pre-play communication is allowed. Theorem 3.10 introduced in the next section will show under which conditions the above conclusion is true. The second conclusion is that the observation made in Example 3.5 that joint misreports are Pareto inefficient is not a general observation.

Let us now tie resilience to implementation. Saijo et al. (2007) show that the rectangular property is a necessary condition for secure implementation. The latter is motivated by the lack of dominant strategy play for rules that admit unwanted Nash equilibria in their direct revelation mechanisms. We provide a definition below.

Definition 3.8. [**Secure Implementation**] A decision rule f is secure implementable[6] if for each $R \in \mathcal{R}^N$,

(i) $R \in DS(\Gamma^*, R)$.

[6] To be precise, the definition of secure implementation is the one by direct mechanism. We are using this definition without loss of generality because Saijo et al. (2007)

	R_2	$\tilde{R}_2$
R_1	$f_1(R_1, R_2),\ f_2(R_1, R_2)$	$f_1(R_1, \tilde{R}_2),\ f_2(R_1, \tilde{R}_2)$
$\tilde{R}_1$	$f_1(\tilde{R}_1, R_2),\ f_2(\tilde{R}_1, R_2)$	$f_1(\tilde{R}_1, \tilde{R}_2),\ f_2(\tilde{R}_1, \tilde{R}_2)$

Table 1: The Rectangular Property

(ii) for each $\tilde{R} \in NE(\Gamma^*, R)$, $f(\tilde{R}) = f(R)$.

In words, rule f is secure implementable if truthtelling is a dominant strategy in the direct revelation mechanism associated to f, and for any Nash equilibrium report $\tilde{R}$ at profile R, it must be that $f(\tilde{R}) = f(R)$. Note that secure implementation is nothing more than the requirement of double implementation in dominant strategies and Nash equilibrium.

We first recall the intuition behind the necessity of the rectangular property for secure implementation. Our discussion is based on table 1 for which we have fixed a rule f. At the true preference profile (R_1, R_2), suppose that $f_1(R_1, \tilde{R}_2) I_1 f_1(\tilde{R}_1, \tilde{R}_2)$. By strategy-proofness, this implies that reporting $\tilde{R}_1$ is a best-response for agent 1 when agent 2 reports $\tilde{R}_2$. Assume next that $f_2(\tilde{R}_1, R_2) I_2 f_2(\tilde{R}_1, \tilde{R}_2)$. By the same token, $\tilde{R}_2$ is a best-response when agent 1 reports $\tilde{R}_1$. Hence $(\tilde{R}_1, \tilde{R}_2)$ is a Nash equilibrium at (R_1, R_2) with $f(\tilde{R}_1, \tilde{R}_2)$ as a Nash equilibrium outcome. By resilience we must have that $f(R_1, R_2) = f(\tilde{R}_1, \tilde{R}_2)$.

It is not difficult to see that a rule f is secure implementable if and only if f is resilient. For instance, if a preference report $\tilde{R}$ is a Nash equilibrium in the direct revelation game associated with a resilient rule f at profile R then $\tilde{R}$ must deliver the desired outcome $f(R)$. In addition, because f is strategy-proof by Lemma 3.3, truthtelling is a dominant strategy for the direct revelation game. Based on Theorem 3.4 and given Saijo et al. (2007)'s characterization in terms of strategy-proofness and the rectangular property, we obtain a new characterization of the class of secure implementable rules.

Theorem 3.9. ***Secure Implementation*** *A decision rule f is secure implementable if and only if f satisfies resilience.*

An interesting by-product of the theorem is a direct characterization of secure implementation in the form of a straightforward and intuitive condition. Indeed resilience is much easier to check in practice than both strategy-proofness and the rectangular property.

3.2 Group-Resilience

In this section, we identify the necessary and sufficient conditions for group-resilience and compare these conditions to the ones we obtained for resilience. For this purpose, we find useful to invoke the result from Saijo et al. (2007) that shows the equivalence between the rectangular property and the combination of the outcome rectangular property and non-bossiness in welfare. We define these conditions below.

Non-bossiness in welfare: A decision rule f satisfies non-bossiness in welfare if whenever $f_i(R)\ I_i\ f_i(\tilde{R}_i, R_{-i})$ for some $i \in N$, $R \in \mathcal{R}^N$ and $\tilde{R}_i \in \mathcal{R}_i$, then $f(R) = f(\tilde{R}_i, R_{-i})$.[7]

Outcome rectangular property: A decision rule f satisfies the outcome rectangular property if for each $R, R' \in \mathcal{R}^N$, if $f_i(R_i, R'_{-i}) = f_i(R')$ for each $i \in N$, then $f(R) = f(R')$.

We now show that group resilience is equivalent to strategy-proofness and non-bossiness in welfare.

Theorem 3.10. *A decision rule f satisfies group-resilience if and only if f satisfies both strategy-proofness and non-bossiness in welfare.*

Proof. We start with the necessity part of our proof. We already know from Lemma 3.3 that any group-resilient rule is strategy-proof. We show that any group resilient rule f satisfies non-bossiness in welfare. In contrast, suppose that this is not the case. Then there must exist $i \in N$, $R \in \mathcal{R}^N$ and $\tilde{R}_i \in \mathcal{R}_i$ such that

$$f(\tilde{R}_i, R_{-i})\ I_i\ f(R) \text{ and } f(\tilde{R}_i, R_{-i}) \neq f(R). \tag{2}$$

Then by group-resilience, there must exist S such that $f(R_S, R_{-S})\ P_j\ f(\tilde{R}_i, R_{-i})$ for all $j \in S$. Because the preferences for each $j \neq i$ are the same under both (R_S, R_{-S}) and $(\tilde{R}_i, R_{-i})$, it must be that $i \in S$. Consequently, $f(R_S, R_{-S}) = f(R)\ P_i\ f(\tilde{R}_i, R_{-i})$ which is a contradiction with (2).

We now prove the sufficiency part. Let $R, \tilde{R} \in \mathcal{R}^N$ be such that $f(R) \neq f(\tilde{R})$. Suppose that there exists some $i \in N$ and $\bar{R}_i \in \mathcal{R}_i$ such that $f(\bar{R}_i, \tilde{R}_{-i})\ P_i\ f(\tilde{R})$. This fact and strategy-proofness of f implies that $f(R_i, \tilde{R}_{-i})\ R_i\ f(\bar{R}_i, \tilde{R}_{-i})\ P_i\ f(\tilde{R})$. Next, suppose that there does not exist such an agent i for whom $f(\bar{R}_i, \tilde{R}_{-i})\ P_i\ f(\tilde{R})$ for some $\bar{R}_i \in \mathcal{R}_i$ –i.e. $\tilde{R} \in NE(\Gamma^*, R)$. Let $T \subset N$ with $|T| = t$ be the subset of agents who lie at $\tilde{R}$, i.e. $(\tilde{R}_T, R_{-T}) = (\tilde{R}_T, \tilde{R}_{-T}) = \tilde{R}$. The following claim is crucial for our proof.

Claim: Fix any nonnegative integer $s < t$. If $f(R_{\bar{S}}, \tilde{R}_{-\bar{S}}) = f(\tilde{R})$ for all $\bar{S} \subset T$ with $|\bar{S}| \leq s$, then for all $S \subseteq T$ with $|S| = s + 1$, it must be either (i) $f(R_S, \tilde{R}_{-S})\ P_i\ f(\tilde{R})$ for all $i \in S$ or (ii) $f(R_S, \tilde{R}_{-S}) = f(\tilde{R})$.

The claim above, which we prove below, immediately yields that either (a) there exists some $S \subseteq T$ with $f(R_S, \tilde{R}_{-S})\ P_i\ f(\tilde{R})$ or (b) $f(R_T, \tilde{R}_{-T}) = f(\tilde{R})$. In the latter case, $f(R) = f(\tilde{R})$ (recall $(R_T, \tilde{R}_{-T}) = (R_T, R_{-T})$). This would contradict our assumption that $f(R) \neq f(\tilde{R})$. Thus, the only possible case is (a) which concludes the proof of the

[7] Saijo et.al. (2007) labels this condition simply as non-bossiness.

sufficiency part.

Proof of the Claim. Let us prove the claim when $s = 0$. Fix any $S \subset T$ with $|S| = 1$. By construction, $S = \{i\}$ for some $i \in T$. By the Nash equilibrium assumption of $\tilde{R} = (\tilde{R}_S, R_{-S})$ at R, we have

$$f_i(\tilde{R}) \ R_i \ f_i(R_i, \tilde{R}_{-i}) = f_i(R_S, \tilde{R}_{-S}), \tag{3}$$

And by strategy-proofness,

$$f_i(R_i, \tilde{R}_{-i}) \ R_i \ f_i(\tilde{R}). \tag{4}$$

By combining the two relations above, we obtain that

$$f_i(R_i, \tilde{R}_{-i}) \ I_i \ f_i(\tilde{R}).$$

Because f satisfies non-bossiness in welfare, we have

$$f_i(R_i, \tilde{R}_{-i}) = f(R_S, \tilde{R}_{-S}) = f_i(\tilde{R}). \tag{5}$$

This means that (ii) in the claim is always satisfied if $s = 0$.

Now fix any s with $0 < s < t$. Pick any $S \subseteq T$ with $|S| = s + 1$. Because of the strategy-proofness of f, we know that $f(R_S, \tilde{R}_{-S}) \ R_i \ f(R_{S\setminus i}, \tilde{R}_i, \tilde{R}_{-S})$ for all $i \in S$. By construction, $|S \setminus i| = s$. By the assumption used in the claim, $f(R_{S\setminus i}, \tilde{R}_i, \tilde{R}_{-S}) = f(\tilde{R})$. Consequently,

$$f(R_S, \tilde{R}_{-S}) \ R_i \ f(R_{S\setminus i}, \tilde{R}_i, \tilde{R}_{-S}) = f(\tilde{R}) \text{ for all } i \in S. \tag{6}$$

If the relation above holds for everyone with a strict one, then we are in (i) of the claim. If not, there must be at least one agent i for whom $f(R_S, \tilde{R}_{-S}) \ I_i \ f(R_{S\setminus i}, \tilde{R}_i, \tilde{R}_{-S}) = f(\tilde{R})$. Then because f satisfies non-bossiness in welfare, we obtain $f(R_S, \tilde{R}_{-S}) = f(R_{S\setminus i}, \tilde{R}_i, \tilde{R}_{-S}) = f(\tilde{R})$ which is (ii) of the claim. This completes the proof.

□

While strategy-proofness is central in the mechanism design literature, non-bossiness in welfare is typically used to achieve different objectives, in particular to provide structure and tractability to different classes of rules in characterization theorems.[8] Here, non-bossiness in welfare inherits a somewhat unexpected strategic interpretation. While group resilience looks like a stringent requirement, it turns out to be no more demanding than two properties that have been regularly invoked in the literature.

[8]For a thorough discussion of non-bossiness and its use in the literature, see for instance the excellent survey by Thomson (2016).

4 Strategic Foundations of Group-Resilience

Our goal in this section is to provide some mechanism design foundations for group resilience. We first to connect group resilience with group strategy-proofness. Doing so achieves two objectives. First it allows us to identify well-known classes of rules which violate resilience but satisfy group resilience. Second, because group resilience requires coalitions to coordinate on joint reports –possibly via pre-play communication– nothing prevents coalitions from then coordinating on beneficial false reports. As such imposing group strategy-proofness is warranted, while the property seems very close to group resilience. Indeed, suppose that f is not group strategy-proof. Then there would exist a group of agents $S \subseteq N$ that gain by falsely reporting a subprofile $\tilde{R}_S$ when the true profile is R. Note that by group resilience this misreport is inherently unstable. This observation will lead us to the introduction of a new implementation notion that we call *group secure implementation*, a direct extension of secure implementation via direct mechanisms which requires double implementation in dominant strategies and in strong Nash equilibrium. Group strategy-proofness is a necessary condition for group secure implementation.

4.1 From Group Resilience to Group Strategy-Proofness

We pointed out earlier that the necessary and sufficient conditions for resilience are strategy-proofness, non-bossiness in welfare and the outcome rectangular property. On the other hand, group-resilience only requires strategy-proofness and non-bossiness in welfare. Due to this relaxation, some non-resilient rules that are prominent in the literature are group-resilient. This is already apparent in the previous series of examples, to the exception of Example 3.7. We first study in turn two models (domains), the allocation of private goods under strict preferences (e.g. Shapley-Scarf economies), and the allocation of a stock of resource under single-peaked preferences (Sprumont, 1991). Next, we provide a more general result under a notion of richness of the domain of preferences.

For this step, we find useful to introduce two new concepts. One is non-bossiness which is a weaker version of non-bossiness in welfare and the second one is group strategy-proofness, a strengthening of strategy-proofness once coalitions can form.

Group Strategy-Proofness: A decision rule f satisfies group strategy-proofness if for each $R \in \mathcal{R}^N$ and coalition $S \subseteq N$, there does not exist $R'_S \in \mathcal{R}^S$ such that $f(R'_S, R_{-S})\,\mathrm{wdom}[R,S] f(R)$.

Non-Bossiness: A decision rule f satisfies non-bossiness if whenever $f_i(R) = f_i(\tilde{R}_i, R_{-i})$ for some $i \in N$, $R \in \mathcal{R}^N$ and $\tilde{R}_i \in \mathcal{R}$, then $f(R) = f(\tilde{R}_i, R_{-i})$.

We now consider various models of interest and identify group-resilient rules.

4.1.1 Strict Preference Domain

In this domain, non-bossiness in welfare and its weaker version, non-bossiness, are equivalent. Thus, a rule is group-resilient if and only if it satisfies both strategy-proofness and non-bossiness. Furthermore, in the private goods model considered in Example 3.6, the combination of non-bossiness and strategy-proofness is equivalent to group strategy-proofness (Pápai, 2000).

Corollary 4.1. *Let $\mathcal{R}^N$ be the strict preference domain and let F determine a private good economy as defined in Example 3.6. A decision rule f satisfies group-resilience if and only if f satisfies group strategy-proofness.*

The above corollary implies that, in the direct revelation mechanism associated with any group-resilient rule f, no coalition of agents can profitably deviate from truthtelling, while there always exist some coalition that would gain by reverting to truthtelling from any report that is not prescribed by f. Truthtelling is an unequivocally focal strategy, i.e. truthtelling is very likely to arise in the direct revelation mechanism associated with f when communication is allowed among agents.

Another interesting feature of group resilient rules in this setting is self-enforcement in the spirit of coalition proof Nash equilibria (Bernheim et al., 1987). That is, any coalition which benefits by reverting to truthtelling from a collective misreport does not have to worry about further deflections by its members, thanks to group strategy-proofness.

Given that we have translated requirements for resilient and group-resilient rules in terms of conditions regularly used in the literature, we can compare the sets of resilient and group-resilient rules. Fujinaka and Wakayama (2011) show that the only secure implementable efficient rules in the so-called housing model are the *priority rules* which allocate objects based on a fixed ordering of the set of agents.[9] Therefore, efficiency – perhaps the most important criteria – along with resilience lead to an incredibly restrictive set of priority rules. On the other hand, Pycia and Ünver (forthcoming) show that any group-strategy proof and efficient rule is a trading cycle rule. The set of such rules includes Gale's top trading cycles rules – a prominent rule in the so-called house allocation literature.[10] Consequently, the set of group-resilient and efficient rules is the one of trading cycles rules – a considerable enlargement over the set of priority rules.

Out of the three necessary conditions for secure implementation or resilience, our results show that the outcome rectangular property can be completely dispensed with for group resilience.

[9] Priority rules are also known as serial dictatorships.

[10] Top trading cycles rules are also used in important practical problems: For example, on April 16, 2012, it was announced that the New Orleans Recovery School District would utilize a version of the top trading cycles allocation rule as the allocation rule for the centralized enrollment of children in public schools (Vanacore, 2012).

Before we go to our next preference domain, let us remark that non-bossiness is vacuously satisfied in public goods models in strict preference domains. As a result, group-resilience in such models is equivalent to strategy-proofness.

4.1.2 Single-Peaked Domain

Let us now turn our attention to the Sprumont model introduced in Examples 2.2 and 3.5. We first show that group-resilience implies group-strategy proofness in this setting. Recall that $\Omega \in \mathbb{R}_{++}$ is the stock of the resource. The feasible set of allocations is $F = \{x \in A : \sum_i x_i = \Omega\}$.

Lemma 4.2. *Let $\mathcal{R}^N$ be the single-peaked preferences domain and let F determine the feasible set of the Sprumont model. If a decision rule f satisfies group-resilience then it satisfies group strategy-proofness.*[11]

Proof. Because f is group-resilient, f satisfies strategy-proofness and non-bossiness in welfare. In contradiction to the lemma, let f satisfy both strategy-proofness and non-bossiness in welfare but not group strategy-proofness. Consequently, there exist $R \in \mathcal{R}^N$, $S \subset N$ and $\tilde{R}_S \in \mathcal{R}^S$ such that $f(\tilde{R}_S, R_{-S})\,\text{wdom}[R,S]\,f(R)$. By Lemma 6.1 (see the appendix), f is peak-only. Let $\hat{R} \in \mathcal{R}^N$ be a preference profile such that $p(\hat{R}_i) = f_i(\tilde{R}_S, R_{-S})$. Pick any $i \in S$. We will now show that $f(\hat{R}_i, R_{-i}) = f(R)$. Suppose otherwise. Non-bossiness[12] would yield $f(\hat{R}_i, R_{-i}) = f(R)$ if $f_i(\hat{R}_i, R_{-i}) = f_i(R)$. Thus, $f_i(\hat{R}_i, R_{-i}) \neq f_i(R)$. If $p(\hat{R}_i) = p(R_i)$, then by peak-onliness, $f_i(\hat{R}_i, R_{-i}) = f_i(R)$. Hence, we must have that $p(\hat{R}_i) \neq p(R_i)$. The proofs for the $p(\hat{R}_i) < p(R_i)$ and $p(\hat{R}_i) > p(R_i)$ cases are similar. Subsequently, let us only consider the $p(\hat{R}_i) > p(R_i)$ case. By Lemma 6.2 (see the appendix), the only possibility in which $f_i(\hat{R}_i, R_{-i}) \neq f_i(R)$ occurs if $p(R_i) \leq f_i(R) < f_i(\hat{R}_i, R_{-i}) \leq p(\hat{R}_i)$. However, this implies that $f_i(R)\ P_i\ p(\hat{R}_i) = f_i(\tilde{R}_S, R_{-S})$ because R_i is single-peaked. This contradicts that $i \in S$ and $f(\tilde{R}_S, R_{-S})\,\text{wdom}[R,S]f(R)$. Consequently, we have that $f(\hat{R}_i, R_{-i}) = f(R)$.

Now pick any $j \in S$ and $j \neq i$. Because $f(\hat{R}_i, R_{-i}) = f(R)$, by following the same steps as above, we obtain that $f(\hat{R}_{\{i,j\}}, R_{-\{i,j\}}) = f(\hat{R}_i, R_{-i}) = f(R)$. By continuing with the same logic, it must be that $f(\hat{R}_S, R_{-S}) = f(R)$.

We now move from $(\tilde{R}_S, R_{-S})$ to $(\hat{R}_S, R_{-S})$ by changing the preferences of agents in S, one at a time. We claim that at each step of this process, the allocation prescribed by f remains unaffected. To see this select any $i \in S$ and consider $(\hat{R}_i, \tilde{R}_{S\setminus\{i\}}, R_{-S})$. By the strategy-proofness of f, we must have that $f_i(\hat{R}_i, \tilde{R}_{S\setminus\{i\}}, R_{-S}) = f_i(\tilde{R}_S, R_{-S})$ because $f_i(\tilde{R}_S, R_{-S}) = p(\hat{R}_i)$. Thus, by non-bossiness we have that $f(\hat{R}_i, \tilde{R}_{S\setminus\{i\}}, R_{-S}) =$

[11] The proof of the Lemma relies on two auxiliary lemmas that we prove in the appendix.

[12] The stronger version of non-bossiness – non-bossiness in welfare – is not needed here.

$f(\tilde{R}_S, R_{-S})$. By employing similar arguments for the remaining steps of the process, we find that $f(\hat{R}_S, R_{-S}) = f(\tilde{R}_S, R_{-S})$. This contradicts the earlier conclusion that $f(\hat{R}_S, R_{-S}) = f(R)$ because $f(\tilde{R}_S, R_{-S}) \neq f(R)$. $\square$

The lemma above and Theorem 3.10 imply that truthtelling is again very likely to arise in the direct revelation games associated with group-resilient rules when pre-play communication is allowed. Not only no coalition can profitably deviate from truthtelling, they also cannot coordinate on an outcome not prescribed by the rule because some coalition will always find profitable to revert to truthtelling. In addition, if the rule satisfies efficiency – arguably the most desirable criteria in any resource allocation problem– then group strategy-proofness coincides with group-resilience.

Lemma 4.3. *Let $\mathcal{R}^N$ be the single-peaked preferences domain and let F determine the feasible set of the Sprumont model. If an efficient decision rule f satisfies group-resilience, then f satisfies group strategy-proofness.*

Proof. Let f be a group strategy-proof rule. First, let us prove that f satisfies non-bossiness. In contrast, let f violate the property. Then there must exist R, i and $\tilde{R}_i$ such that $f_i(R) = f_i(\tilde{R}_i, R_{-i})$ and $f(R) \neq f(\tilde{R}_i, R_{-i})$. Then because $f(\tilde{R}_i, R_{-i}) \neq f(R)$ and $\sum_{j\in N} f_j(R) = \sum_{j\in N} f_j(\tilde{R}_i, R_{-i}) = \Omega$, we must have two agents $j' \neq i$ and $i' \neq i$ with $f_{i'}(R) > f_{i'}(\tilde{R}_i, R_{-i})$ and $f_{j'}(R) < f_{j'}(\tilde{R}_i, R_{-i})$. Efficiency requires that both $f_{i'}(R)$ and $f_{j'}(R)$ to be on the same side of the respective peaks of these players. The same is true for $f_{i'}(\tilde{R}_i, R_{-i})$ and $f_{j'}(\tilde{R}_i, R_{-i})$. Given that the peaks of these agents do not change between profiles R and $(\tilde{R}_i, R_{-i})$, $f_{i'}(R)$ and $f_{i'}(\tilde{R}_i, R_{-i})$ are on the same side of $p(R_{i'})$. The same thing is true for agent j'. Consequently, by single-peakedness, one of i' and j' strictly prefers $f(\tilde{R}_i, R_{-i})$ to $f(R)$. Thus, one of i' and j' along with i can weakly block $f(R)$. Thus, f is not group strategy-proof, a contradiction.

Finally, suppose that f violates non-bossiness in welfare. We have showed f satisfies non-bossiness. Thus, there must exist i, R and $\tilde{R}_i$ such that $f(\tilde{R}_i, R_{-i})\ I_i\ f(R)$ and $f_i(\tilde{R}_i, R_{-i}) \neq f_i(R)$. By single peakedness, we find that $f_i(\tilde{R}_i, R_{-i})$ and $f_i(R)$ are on the different sides of $p(R_i)$. Without loss of generality, we assume that $f_i(R) < p(R_i) < f_i(\tilde{R}_i, R_{-i})$. Efficiency now implies that for all j, $f_j(R) \leq p(R_j)$.

Strategy-proofness and single-peakedness imply that i cannot have any deviation R_i' with $f(R_i', R_{-i}) \in (f_i(R), f_i(\tilde{R}_i, R_{-i}))$. Consider a report $\bar{R}_i$ with $p(\bar{R}_i) = p(R_i)$ and $f_i(\tilde{R}_i, R_{-i})\ \bar{P}_i\ f_i(R)$. Then $f_i(\tilde{R}_i, R_{-i}) \notin (f_i(R), f_i(\tilde{R}_i, R_{-i}))$. Moreover, strategy-proofness and single-peakedness yield that $f_i(\bar{R}_i, R_{-i}) = f_i(\tilde{R}_i, R_{-i})$. If there exists any agent $j \neq i$ for whom $f(\tilde{R}_i, R_{-i})\ P_j\ f(R)$, then coalition $S = \{i, j\}$ can profit by deviating from R_S to $(\tilde{R}_i, R_j)$. Because $f_i(\bar{R}_i, R_{-i}) > f_i(R)$ and $\sum_{j\in N} f_j(R) = \sum_{j\in N} f_j(\bar{R}_i, R_{-i}) = \Omega$, there must exist an agent i' with $f_{i'}(R) > f_{i'}(\bar{R}_i, R_{-i})$. Recall that $f_{i'}(R) < p(R_{i'})$. Thus, $p(R_{i'}) > f_{i'}(\bar{R}_i, R_{-i})$. By construction, $f_i(\bar{R}_i, R_{-i}) < p(\bar{R}_i)$. This contradicts efficiency. $\square$

We here note that without imposing additional conditions, one cannot show that group strategy-proofness implies group resilience even in the Sprumont setting.

Example 4.4. *Group strategy-proofness does not imply group-resilience*
Let $N = \{1,2\}$ and consider the following rule which yields two outcomes: $x = (\Omega, 0)$ or $y = (0, \Omega)$. If x and y are Pareto comparable in the weak sense at any profile R then f assigns the better outcome. If both agents are indifferent between x and y, then f prescribes x. If none of these conditions are satisfied, f prescribes x or y depending on which one agent 1 prefers.

Clearly, f is group strategy-proof but it violates non-bossiness in welfare. Thus, f is not group-resilient. $\diamond$

To assess how demanding the requirements for group-resilience is, we further investigate the condition of non-bossiness in welfare. In the following result, we show that efficiency along with strategy proofness and non-bossiness imply non-bossiness in welfare.

Lemma 4.5. *Let $\mathcal{R}^N$ be the single-peaked preferences domain and let F determine the feasible set of the Sprumont model. If a decision rule f satisfies efficiency, strategy-proofness and non-bossiness, then it satisfies non-bossiness in welfare.*

Proof. Suppose f does not satisfy non-bossiness in welfare. Therefore, there must exist $R \in \mathcal{R}^N$, $i \in N$ and $R'_i \in \mathcal{R}_i$ such that $f(R'_i, R_{-i}) \neq f_i(R)$ and $f_i(R'_i, R_{-i})\ I_i\ f_i(R)$. In fact, because f satisfies non-bossiness, $f_i(R'_i, R_{-i}) \neq f_i(R)$. This means that $f_i(R'_i, R_{-i})$ and $f_i(R)$ are on the opposite sides of $p(R_i)$. Without loss of generality assume that $f_i(R) < p(R_i) < f_i(R'_i, R_{-i})$. Then by efficiency, $f_j(R) \leq p(R_j)$ for all $j \neq i$. Indeed, if $f_j(R) > p(R_j)$ for some j, we can improve i and j by taking tiny amount from j's allocation and by increasing i's by the the same amount. This implies that $\sum_j f_j(R) = \Omega < \sum_j p(R_j)$.

Fix a preference $\hat{R}_i \in \mathcal{R}_i$ such that $p(\hat{R}_i) = p(R_i)$ but $f_i(R)\ \hat{P}_i\ f_i(R'_i, R_{-i})$. By strategy-proofness, $f_i(\hat{R}_i, R_{-i}) = f_i(R'_i, R_{-i}) > p(\hat{R}_i)$. Then by efficiency, $\Omega > p(\hat{R}_i) + \sum_{j \neq i} p(R_j) = \sum_j p(R_j)$. This contradicts our earlier conclusion that $\Omega < \sum_j p(R_j)$. $\square$

Theorem 3.10 and Lemma 4.5 imply that any rule satisfying efficiency, strategy-proofness and non-bossiness is group-resilient. This class contains the whole family of fixed path rules by Moulin (1999), the sequential allotment rules Barbera et al. (1997), and these two classes each contain the celebrated uniform rule. In contrast, Bochet and Sakai (2010) show that the only secure implementable rules within the fixed path rules are the priority rules, i.e., serial dictatorships.

Let us finally comment on the public goods model of Moulin (1980). Unlike its counterpart in strict preferences domains, non-bossiness in welfare is not vacuously satisfied in this model. Thus, group-resilience is not equivalent to group strategy-proofness. However, group-resilience requires group strategy-proofness: the proof for this statement is identical to the one of Lemma 4.2. In addition, group strategy-proofness is equivalent to group-resilience for the class of efficient rules.[13] It is well-known that generalized median rules are group strategy-proof and efficient. Therefore, the set of group resilient rules includes the whole class of generalized median voter rules.

4.1.3 Rich Domains

We have discussed the desirability of rules that are both group-resilient and group strategy-proofness in some specific models. We explore now the relation between these two properties in a general environment. In Example 4.4, we showed that group strategy-proofness does not imply group resilience. At the same time, based on the previous subsections, one may conclude that group-resilience is more demanding than group strategy-proofness. We now show that this observation is not true in general environments. In fact, even resilience does not imply group strategy-proofness in general settings.

Example 4.6. *Narrow domain: strategy-proofness and non-bossiness in welfare*
Let $N = \{1, 2\}$ and $F = \{w, x, y, z\}$. Suppose that the set of preferences for each agent is $R_i \cup \tilde{R}_i$ where

$$c \; P_1 \; d \; P_1 \; a \; P_1 \; b \qquad\qquad b \; \tilde{P}_1 \; d \; \tilde{P}_1 \; a \; \tilde{P}_1 \; c$$
$$b \; P_2 \; d \; P_2 \; a \; P_2 \; c \qquad\qquad c \; \tilde{P}_2 \; d \; \tilde{P}_2 \; a \; \tilde{P}_2 \; b.$$

Now consider the following rule f:

	R_2	$\tilde{R}_2$
R_1	a	c
$\tilde{R}_1$	b	d

Observe here that f satisfies strategy-proofness and non-bossiness in welfare. In fact, f satisfies the rectangular property and is thus both resilient and group-resilient. However, f is not group strategy-proof since $f(\tilde{R}) = d \; P_i \; a = f(R)$ for both $i = 1, 2$. ◇

A key feature for the failure of group strategy-proofness in the above example is the narrowness of the preference domain. When the preference domain is sufficiently

[13] In public good models, non-bossiness is vacuously satisfied. Thus, one needs to prove that group-strategy proofness requires f be non-bossy in welfare. If $f(R) I_i f(\tilde{R}_i, R_{-i})$ and $f(R) \neq f(\tilde{R}_i, R_{-i})$, then $f(R)$ and $f(\tilde{R}_i, R_{-i})$ are on different sides of $p(R_i)$. Assume without loss of generality, $f(R) < p(R_i) < f(\tilde{R}_i, R_{-i})$. Fix a report $\bar{R}_i$ with $p(\bar{R}_i) = p(R_i)$ and $f(\tilde{R}_i, R_{-i}) \bar{P}_i f(R)$. Strategy proofness and single-peakedness imply that $f(\bar{R}_i, R_i) \notin (f(R), f(\tilde{R}_i, R_{-i}))$. In fact, $f(\bar{R}_i, R_i) = f(\tilde{R}_i, R_{-i})$. Thus, $p(\bar{R}_i) < f(\bar{R}_i, R_i)$. At the same time, group strategy-proofness implies that $p(R_j) < f(\tilde{R}_i, R_{-i})$ for all $j \neq i$; otherwise, such an agent j and i together weakly improves from truthtelling by reporting $(\tilde{R}_i, R_j)$ at profile R. Subsequently, $p(R_j) < f(\tilde{R}_i, R_{-i}) = f(\bar{R}_i, R_{-i})$ for all $j \neq i$. Recall that $p(\bar{R}_i) < f(\bar{R}_i, R_i)$. Thus, f is not efficient at $(\bar{R}_i, R_{-i})$.

"large", the combination of strategy-proofness and non-bossiness in welfare implies group strategy-proofness. We introduce below a richness condition on the preference domain which guarantees that the two conditions imply group strategy-proofness.

Rich domain: Domain $\mathcal{R}^N$ is *rich* if for each $i \in N$, $x_i \neq y_i \in A_i$, there exists $R_i \in \mathcal{R}_i$ such that $x_i \; P_i \; y_i \; P_i \; z_i$ for all $z_i \in A_i$ with $z_i \neq x_i$ and $z_i \neq y_i$.

If the domain $\mathcal{R}^N$ is rich, then for any given agent $i \in N$ and for any two alternatives $x_i, y_i \in A_i$, there must exist preferences for agent i which place these alternatives as the top two alternatives. For instance, in private good economies the strict preference domain is rich. Clearly, any preference domain containing the strict preference domain is also rich. Furthermore, some domains not containing the strict preference domain are rich. For instance, suppose that the agents only care about their top two alternatives. Such domains would be rich as long any two alternatives are the top two at some point. However, some important domains of interest such as the single peaked domain in the Sprumont setting are not rich.

Theorem 4.7. *Let $\mathcal{R}^N$ be a rich domain. If a decision rule f satisfies group-resilience then it satisfies group strategy-proofness.*

Proof. Pick $R \in \mathcal{R}^N$ and suppose that f violates group strategy-proofness at R. Hence, there exists $S \subseteq N$, $R'_S \in \mathcal{R}^S$ such that $f(R'_S, R_{-S}) \,\text{wdom}[R, S]\; f(R)$. Let $\hat{R}_S \in \mathcal{R}^S$ be the preference for S such that

(i) for any $i \in S$ for whom $f_i(R) = f_i(R'_S, R_{-S})$, $f_i(R'_S, R_{-S})$ is the most preferred alternative for agent i under $\hat{R}_i$

(ii) for any $i \in S$ for whom $f_i(R) \neq f_i(R'_S, R_{-S})$, $f_i(R'_S, R_{-S}) \; \hat{P}_i \; f_i(R) \; \hat{P}_i \; z_i$ for $z_i \neq f_i(R'_S, R_{-S})$ and $z_i \neq f_i(R)$.

The existence of such a profile of preferences for S is guaranteed because our domain is rich. We now change R to $(\hat{R}_S, R_{-S})$, one agent's preference at a time. We show that the initial selection operated by f, $f(R)$, does not change at any step of this process. Pick any $i \in S$. If $f_i(R'_S, R_{-S}) = f_i(R)$, then by strategy-proofness we must have $f_i(\hat{R}_i, R_{-i}) = f_i(R)$. Otherwise, agent i would have a profitable deviation at $(\hat{R}_i, R_{-i})$ because $f_i(R)$ is the most preferred alternative for i at $\hat{R}_i$ (by construction). Suppose that $f_i(R'_S, R_{-S}) \neq f_i(R)$. Then strategy-proofness of f implies that $f_i(\hat{R}_i, R_{-i})$ is either $f_i(R'_S, R_{-S})$ or $f_i(R)$. Otherwise, agent i would have a profitable deviation at $(\hat{R}_i, R_{-i})$ because by construction, $f_i(R'_S, R_{-S})$ and $f_i(R)$ are the two most preferred alternatives for i at $\hat{R}_i$. Because $f(R'_S, R_{-S}) \,\text{wdom}[R, S]\; f(R)$ and $i \in S$, we must have that $f_i(R'_S, R_{-S}) \; R_i \; f_i(R)$. If $f_i(R'_S, R_{-S}) \; P_i \; f_i(R)$, then strategy-proofness implies that $f_i(\hat{R}_i, R_{-i}) \neq f_i(R'_S, R_{-S})$. If $f_i(R'_S, R_{-S}) \; I_i \; f_i(R)$, then non-bossiness in welfare implies that $f_i(\hat{R}_i, R_{-i}) \neq f_i(R'_S, R_{-S})$. Thus, in all cases $f_i(\hat{R}_i, R_{-i}) = f_i(R)$. Then by non-bossiness, we obtain that $f(\hat{R}_i, R_{-i}) = f(R)$. Now pick any $j \neq i \in S$. By applying the same arguments as above we obtain that $f(\hat{R}_{\{i,j\}}, R_{-\{i,j\}}) = f(\hat{R}_i, R_{-i}) = f(R)$. The same reasoning applies for the remaining agents in S. Hence, we obtain that $f(\hat{R}_S, R_{-S}) = f(R)$.

We now reach $(\hat{R}_S, R_{-S})$ from (R'_S, R_{-S}) by sequentially changing preferences of agents in S, one at a time. We claim that the initial selection operated by f, $f(R'_S, R_{-S})$, does not change at any step of this process. Pick any $i \in S$. By construction of $\hat{R}_S$, $f_i(R'_S, R_{-S})$ is the most preferred alternative for agent i at $\hat{R}_i$. Then strategy-proofness of f yields that $f_i(\hat{R}_i, R'_{S\setminus\{i\}}, R_{-S}) = f_i(R'_S, R_{-S})$. Now because f satisfies non-bossiness we get that $f(\hat{R}_i, R'_{S\setminus\{i\}}, R_{-S}) = f(R'_S, R_{-S})$. Similar arguments apply for the remaining agents in S. Consequently, we have that $f(\hat{R}_S, R_{-S}) = f(R'_S, R_{-S})$. Recall that earlier we showed that $f(\hat{R}_S, R_{-S}) = f(R)$. Thus, $f(R) = f(R'_S, R_{-S})$ which contradicts $f(R'_S, R_{-S})\,\text{wdom}[R, S]\; f(R)$. □

4.2 Group-Resilience vs. Group Secure Implementation

We showed that resilience is more demanding than strategy-proofness but equivalent to secure implementation. Do the same relations hold for the group versions of these notions? We study this question in this section and introduce the group counterpart of secure implementation.

Definition 4.8. [**Group Secure Implementation**] A decision rule f is group secure implementable if[14] for each $R \in \mathcal{R}^N$,

(i) $R \in DS(\Gamma^*, R) \cap SNE(\Gamma^*, R)$

(ii) $DS(\Gamma^*, R) = f(R)$

(iii) $SNE(\Gamma^*, R) = f(R)$

Observe here that we require truthtelling to be both a dominant strategy and a strong Nash equilibrium. At the same time both $DS(\Gamma^*, R)$ and $SNE(\Gamma^*, R)$ should coincide with $f(R)$. Hence no coalition can gain by deviating from truthtelling, and whenever $f(\tilde{R}) \neq f(R)$, some coalition S can deviate from $\tilde{R}$ and weakly improve when going from $\tilde{R}$ to $(R_S, \tilde{R}_{-S})$. We need to introduce a weakening of non-bossiness in welfare that is necessary for the full implementation in dominant strategies requirement in (ii). It is

[14] As in the definition of secure implementation used in this paper, we are stating group secure implementation in terms of direct mechanisms. However, unlike in the case of secure implementation, group secure implementation via general mechanisms is less demanding than its counterpart using direct mechanisms. We can provide an example demonstrating this phenomenon upon request.

obvious from the definition below that weak non-bossiness in welfare is weaker than non-bossiness in welfare. We also define a notion of preference reversal, the group reversal property.

Weak Non-Bossiness in Welfare: A decision rule f satisfies weak non-bossiness in welfare if whenever $f_i(R_i, \hat{R}_{-i})\ I_i\ f_i(\tilde{R}_i, \hat{R}_{-i})$ for some $i \in N$, $R \in \mathcal{R}^N$, some R_i, $\tilde{R}_i \in \mathcal{R}_i$, and for all $\hat{R}_{-i} \in \mathcal{R}^{N\setminus i}$, then $f(\tilde{R}_i, R_{-i}) = f(R)$.

Definition 4.9. [**Group Reversal Property**] A decision rule f satisfies the group reversal property if for each $R, \tilde{R} \in \mathcal{R}^N$ with $f(R) \neq f(\tilde{R})$, there exist $S \subseteq N$ and $R'_S \in \mathcal{R}^S$ such that $f(R'_S, \tilde{R}_{-S})\,\mathrm{wdom}[R, S]\ f(\tilde{R})$.

We need one last definition, a weakening of non-bossiness in welfare that is necessary for the full implementation in dominant strategies requirement in (ii).

The group reversal property says that whenever a "bad" outcome arises in the direct revelation game associated with the rule in question some coalition can profitably deviate. Group-resilience requires some coalition to profit by reverting to truthtelling from a bad outcome. Thus, group-resilience is a more demanding condition that the group reversal property.

Theorem 4.10. *A rule f is group secure implementable if and only if*

- *f is group-strategy-proof*
- *f satisfies the group reversal property*
- *f satisfies weak non-bossiness in welfare*

Proof. If f is not group-strategy proof, then there exist $R \in \mathcal{R}^N$, $S \subseteq N$ and $\tilde{R}_S \in \mathcal{R}^S$ such that $f(\tilde{R}_S, R_{-S})\,\mathrm{wdom}[R, S] f(R)$. Clearly, $R \notin SNE(\Gamma^*, R)$. Thus, group strategy-proofness is a necessary condition for group secure implementation via the direct mechanism of f.

If f does not satisfy the group reversal property, then there exist $R, \tilde{R} \in \mathcal{R}^N$ such that (i) $f(R) \neq f(\tilde{R})$ and (ii) for no $S \subseteq N$ and no $R'_S \in \mathcal{R}^S$, $f(R'_S, \tilde{R}_{-S})\,\mathrm{wdom}[R, S] f(\tilde{R})$. Thus, $f(\tilde{R}) \in SNE(\Gamma^*, R)$, but $f(\tilde{R}) \neq f(R)$ which is a contradiction. Thus, the group reversal property is a necessary condition for group secure implementation via the direct mechanism of f.

Let us show that f must also also satisfy weak non-bossiness in welfare. Pick $i \in N$, $R \in \mathcal{R}^N$, $R'_i \in \mathcal{R}_i$, and assume that $f(R) \neq f(R'_i, R_{-i})$. By the group secure implementation requirement, we know that $(R'_i, R_{-i}) \notin DS(\Gamma^*, R)$. Hence R'_i must be dominated by R_i for agent i at preference R_i so that there exists $\hat{R}_{-i} \in \mathcal{R}^{N\setminus -i}$ such that $f_i(R_i, \hat{R}_{-i})\ P_i\ f_i(R'_i, R_{-i})$. For the sufficiency part, Consider Γ^* the direct revelation mechanism of f, and let $R \in \mathcal{R}^N$ be the true preference profile. By group strategy-proofness, profile R is a both a dominant strategy and a strong Nash equilibrium of (Γ^*, R). Suppose there is $i \in N$, and $R_i' \in \mathcal{R}_i$ that is also a dominant strategy for i at R_i. If $f(R_i', R_{-i}) \neq f(R)$, the group reversal property ensures that (R_i', R_{-i}) is not a strong Nash equilibrium at R. By weak non-bossiness in welfare, there exists $\hat{R}_{-i} \in \mathcal{R}^{N\setminus -i}$ such that $f_i(R_i, \hat{R}_{-i})\ P_i\ f_i(R_i', R_{-i})$. Hence, R_i' cannot in fact be dominant after all, a contradiction, so $f(R_i', R_{-i}) = f(R)$. Finally if there is $\tilde{R} \in \mathcal{R}^N$ such that $f(R) \neq f(\tilde{R})$, the group reversal property ensures that $\tilde{R}$ cannot be a strong Nash equilibrium. Hence f is group secure implemented by Γ^*. $\square$

Now that we have identified the necessary and sufficient conditions for group secure implementation, we can now compare group-resilience to group secure implementation. The group reversal property is obviously less demanding than group resilience. However group-resilience is neither stronger nor weaker than group strategy-proofness as we pointed out in the previous subsection. Therefore, it is perhaps expected that group resilience and group secure implementation are not comparable in general domains and models. In fact, the rule in Example 4.6 is group-resilient but not group secure implementable. We now present an example of rule that is group secure implementable but not group resilient.

Example 4.11. *Group resilience is not necessary*
Let $N = \{1, 2\}$ and $F = \{a, b, c, d, e, f\}$. Suppose that the set of preferences is $R_1 \cup \tilde{R}_1$ for agent 1 and $R_2 \cup \tilde{R}_2 \cup \bar{R}_2$ for agent 2 where

$$
\begin{array}{lll}
b\ P_1\ c\ I_1\ f\ P_1 a\ P_1 d\ P_1 e & d\ \tilde{P}_1\ e\ \tilde{P}_1\ f\ \tilde{P}_1\ \cdots & \\
a\ P_2\ b\ P_2\ d\ I_2\ f\ P_2 e\ P_2 c & b\ \tilde{P}_2\ e\ \tilde{P}_2\ h\ \tilde{P}_2\ \cdots & c\ \bar{P}_2\ b\ \bar{P}_2\ f\ \cdots
\end{array}
$$

Consider the following rule f:

	R_2	$\tilde{R}_2$	$\bar{R}_2$
R_1	a	b	c
$\tilde{R}_1$	d	e	f

One can easily see that f satisfies non-bossiness in welfare. For instance, at profile $(R_1, \bar{R}_2)$, agent 1 is indifferent between $f(R_1, \bar{R}_2) = c$ and $f(\tilde{R}_1, \bar{R}_2) = f$. Hence, f is not group-resilient. On the other hand, it is easy to see that f is group-strategy proof. In addition, f satisfies the group-reversal property. To see this, let us concentrate on the Nash equilibria that yield a bad outcome at each profile. At profile R, the report $(\tilde{R}_1, \bar{R}_2)$ is a bad Nash equilibrium. However, agents 1 and 2 gain by deviating to $(R_1, \tilde{R}_2)$. At profile $(R_1, \bar{R}_2)$, the report $(\tilde{R}_1, \bar{R}_2)$ is a bad Nash equilibrium. Again agents 1 and 2 profit by deviating to $(R_1, \tilde{R}_2)$. At the remaining profiles, there are no bad Nash equilibria. Thus, any profile leading to a bad outcome is blocked by some group or an individual. Consequently, we find that f satisfies the group reversal property. As a result, f is group secure implementable but not group-resilient. $\diamond$

Finally, we note that when a rule f satisfies group strategy-proofness and group-resilience then f is group secure implementable. Our results from the previous sections identify many such rules. This is in stark contrast to the negative results obtained with secure implementation. Recall that if the domain of preferences is "large" enough, any group-resilient rule is group secure implementable. We close with a recall of our results under the light of group secure implementation.

Corollary 4.12. *Let $\mathcal{R}^N$ be the strict preference domain and let F determine a private good economy. If a decision rule f satisfies group-resilience, then f is group secure implementable.*

Corollary 4.13. *Let $\mathcal{R}^N$ be the single-peaked domain and let F determine the feasible set of the Sprumont model. If a decision rule f satisfies efficiency and group-resilience, then f is group secure implementable.*

Corollary 4.14. *Let $\mathcal{R}^N$ be a rich domain. If a decision rule f satisfies group-resilience, then f is group secure implementable.*

5 Discussion and Extensions

5.1 Weak Group-Resilience

In this subsection, we investigate the implications of amending the concept of group-resilience in the sense that any truth-reverting group from a "bad" outcome improves weakly.

Definition 5.1. A rule f is weak group-resilient if whenever $f(R) \neq f(\tilde{R})$, there exists S such that

$$f(R_S, \tilde{R}_{-S}) R_j f(R)$$

for all $j \in S$ and with a strict relation for at least one member of S.

Clearly, each group-resilient rule is weak group-resilient. Therefore, non-bossiness in welfare and strategy-proofness are sufficient conditions for weak group-resilience. It turns out that only one of these sufficient conditions, strategy-proofness, is necessary for weak group-resilience.

Theorem 5.2. *If a rule f satisfies weak group resilience, then f is strategy-proof.*

Proof. In contrast to the lemma suppose that f is not strategy-proof. This means that there must exist $i \in N$, $R \in \mathcal{R}^N$ and $\tilde{R}_i \in \mathcal{R}_i$ such that

$$f(\tilde{R}_i, R_{-i}) \ P_i \ f(R). \tag{7}$$

Consequently, $f(\tilde{R}_i, R_{-i}) \neq f(R)$. Then by weak group-resilience, there must exist S such that $f(R_S, R_{-S}) \ R_j f(\tilde{R}_i, R_{-i})$ for all $j \in S$ and with strict relation for at least one $j \in S$. Because the preferences for each $j \neq i$ are the same under both (R_S, R_{-S}) and $(\tilde{R}_i, R_{-i})$, it must be that $i \in S$. Consequently, $f(R_S, R_{-S}) = f(R) \ R_i f(\tilde{R}_i, R_{-i})$ which is a contradiction with (7). $\square$

Example 5.3. *Non-bossiness in welfare is not necessary*
Consider a house allocation problem with three alternatives, a, b, and c. Let there be two agents, 1 and 2. The set of preferences for agent 1 is unrestricted while the one for agent 2 consist of only one strict preference which ranks a ahead of c. The rule in this case is as follows: $f(R) = (a, c)$ if $a \ P_1 b$ but $f(R) = (b, a)$ if $b \ R_1 a$. It is easy to see that f is strategy-proof. Next, fix a preference profile R in which aI_1b. In this case $f(R) = (b, a)$. Consider another preference $\tilde{R}_1$ of agent 1 in which $a\tilde{P}_1b$. Now $f(\tilde{R}_1, R_2) = (a, c) \neq f(R)$ but $f_1(R) = b \ I_1 \ a = f_1(\tilde{R}_1, R_2)$. This shows that f violates non-bossiness in welfare.
However, f is weak group-resilient. To show this, let us consider any R with $f(R) = (a, c)$. By construction, it must be that $a \ P_1 \ b$. For any $\tilde{R}$ with $f(\tilde{R}) = (b, a)$, agent 1 can revert to truthtelling, R_1, and obtain a. Thus, agent 1 can improve by reverting to truthtelling whenever $f(\tilde{R}) \neq f(R)$ at profile R. Consider any $\bar{R}$ with $f(\bar{R}) = (b, a)$. In addition, we must have $b\bar{R}_1a$. If $b\bar{P}_1a$, then one can see easily that agent 1 will revert back to truthtelling from any report $\tilde{R}$ with $f(\tilde{R}) \neq f(\bar{R})$. If $b\bar{I}_1a$, agent 1 can revert back to truthtelling from any report $\tilde{R}$ with $f(\tilde{R}) \neq f(\bar{R})$ and weakly improve group $\{1, 2\}$. Thus, f is group-resilient. $\diamond$

6 Conclusion

We introduce two new robustness requirements on decision rules which are equivalent to the notions of prevalence of truthtelling. While the resilience conditions seems to be, from the outset, a very demanding condition, our results show that only the individual version of resilience is hard to reach in general. Group-resilience, by not imposing any size on the deviating coalitions, turns out to be much less demanding. Only strategy-proofness and non-bossiness in welfare are required for its characterization. As such making truthtelling salient in direct mechanisms is not more daunting than invoking two well-known conditions from the mechanism design literature. While the combination of strategy-proofness and non-bossiness in welfare may seem strong, our investigation shows that there exist large classes of group-resilient rules in many models of interest.

A disappointing observation however is that auction models, or in general any VCG mechanisms violate group-resilience. From a practical perspective, we expect that group-resilient rules should work well in practice if pre-play communication is possible. Indeed, among all possible deviations, the reversion to truthtelling is the easiest to coordinate on, and therefore the least taxing mentally. A thorough experimental investigation would be called for to strengthen and validate this conjecture.

A by-product of our results is a significant step towards the understanding of the limitations of secure implementation. As we emphasize, if pre-play communication is possible, many of the observed failures of secure implementation just vanish.

References

R. Aumann. Acceptable points in general cooperative n-person games. In *Contributions to the Theory of Games IV.* Princeton University Press, 1959.

S. Barbera, M. Jackson, and A. Neme. Strategy-proof allotment rules. *Games and Economic Behavior*, 18(1):1–21, 1997.

B.Douglas Bernheim, Bezalel Peleg, and Michael D Whinston. Coalition-proof nash equilibria i. concepts. *Journal of Economic Theory*, 42(1):1 – 12, 1987.

O. Bochet and T. Sakai. Secure implementation in allotment economies. *Games and Economic Behavior*, 68(1):35 – 49, 2010.

T. N. Cason, T. Saijo, T. Sjöström, and T. Yamato. Secure implementation experiments: Do strategy-proof mechanisms really work? *Games and Economic Behavior*, 57(2): 206–235, 2006.

B. Dutta and A. Sen. Nash implementation with partially honest individuals. *Games and Economic Behavior*, 74:154–169, 2012.

Y. Fujinaka and T. Wakayama. Secure implementation in shapley–scarf housing markets. *Economic Theory*, 48(1):147–169, 2011.

S. Li. Obviously strategy-proof mechanisms. working paper, Harvard University, 2016.

D. Moreno and J. Wooders. An Experimental Study of Communication and Coordination in Noncooperative Games. *Games and Economic Behavior*, 24(1-2):47–76, July 1998.

H. Moulin. On strategy-proofness and single peakedness. *Public Choice*, 35(4):437–455, 1980.

H. Moulin. Rationing a Commodity along Fixed Paths. *Journal of Economic Theory*, 84 (1):41–72, January 1999.

S. Pápai. Strategyproof assignment by hierarchical exchange. *Econometrica*, 68(6), 2000.

M. Pycia and P. Troyan. Obviously dominance and random priority. working paper, UCLA, 2016.

M. Pycia and U. Ünver. Incentive Compatible Allocation and Exchange of Discrete Resources. *Theoretical Economics*, forthcoming.

Tatsuyoshi Saijo, Tomas Sjöström, and Takehiko Yamato. Secure implementation. *Theoretical Economics*, 2(3):203–229, 2007.

A. Saporiti. Secure implementation with partially honest agents. *Journal of Economic Theory*, 154:216–228, 2014.

Y. Sprumont. The division problem with single-peaked preferences: A characterization of the uniform allocation rule. *Econometrica*, 59(2), 1991.

W. Thomson. Non-bossiness. *Social Choice and Welfare*, 47(3), 2016.

A. Vanacore. Centralized enrollment in recovery school district gets first tryout. *Times-Picayune*, April, 16 2012.

Appendix

We need the following two lemmas for the proof of Lemma 4.2.

Lemma 6.1. *Let $\mathcal{R}^N$ be the single-peaked preferences domain and let F determine the feasible set of the Sprumont model. If a decision rule f satisfies strategy-proofness and non-bossiness in welfare, then it satisfies peak-onliness, i.e., for any $R, \tilde{R} \in \mathcal{R}^N$ with $p(R_i) = p(\tilde{R}_i)$, it must be that $f(R) = f(\tilde{R})$*

Proof. We first show that for all $R \in \mathcal{R}^N$, $i \in N$ and $R'_i \in \mathcal{R}_i$ with $p(R_i) = p(R'_i)$, we have $f_i(R) = f_i(R'_i, R_{-i})$. Suppose otherwise. Let us denote $p(R'_i) = p(R_i) = \bar{p}$. Then because f is strategy-proof and the preferences are single peaked we have $f_i(R) \neq \bar{p}$ and $f_i(R'_i, R_i) \neq \bar{p}$. Without loss of generality let us assume that $f_i(R) < \bar{p}$. If $f_i(R'_i, R_i) < \bar{p}$, then f cannot be strategy-proof because R_i and R'_i are single-peaked and $f_i(R'_i, R_i) \neq f_i(R)$ (by assumption). Thus, we have that $f_i(R'_i, R_i) > \bar{p} > f_i(R)$. Then we can find a single peaked preference $\tilde{R}_i \in \mathcal{R}_i$ such that $p(\tilde{R}_i) = \bar{p}$ and $f_i(R'_i, R_i)\tilde{I}_i f_i(R)$. Again strategy-proofness and the single-peakedness of preferences imply that either $f_i(\tilde{R}_i, R_i) = f_i(R)$ or $f_i(\tilde{R}_i, R_i) = f_i(R'_i, R_i)$. Without loss of generality assume that $f_i(\tilde{R}_i, R_i) = f_i(R)$. As a result when agent i's preference is $\tilde{R}_i$, he can deviate to R'_i and obtain

$f_i(R'_i, R_{-i})$. By construction, $f_i(R'_i, R_i)\tilde{I}_i f_i(R) = f_i(\tilde{R}_i, R_i)$. Then non-bossiness in welfare implies that $f(R'_i, R_i) = f(\tilde{R}_i, R_{-i})$. This contradicts $f_i(\tilde{R}_i, R_i) = f_i(R) \neq f_i(R'_i, R_i)$. Therefore, $f_i(R) = f_i(R'_i, R_{-i})$. The rest of the proof is a simple consequence of non-bossiness in welfare. □

Lemma 6.2. *Let $\mathcal{R}^N$ be the single-peaked preferences domain and let F determine the feasible set of the Sprumont model. Let decision rule f satisfy strategy-proofness and peak-onliness. For any $R \in \mathcal{R}^N$ and $\tilde{R}_i \in \mathcal{R}_i$ with $p(R_i) < p(\tilde{R}_i)$, one of the following cases must occur:*

(i) $f_i(R) = f_i(\tilde{R}_i, R_{-i}) \leq p(R_i) < p(\tilde{R}_i)$.

(ii) $p(R_i) \leq f_i(R) \leq f_i(\tilde{R}_i, R_{-i}) \leq p(\tilde{R}_i)$, $p(R_i) < f_i(\tilde{R}_i, R_{-i})$ *and* $f_i(R) < p(\tilde{R}_i)$.

(iii) $p(R_i) < p(\tilde{R}_i) \leq f_i(R) = f_i(\tilde{R}_i, R_{-i})$.

Proof. The lemma is a direct consequence of the following results which we prove next.

(a) $f_i(R) \leq f_i(\tilde{R}_i, R_{-i})$.

(b) If either $f_i(R) < p(R_i)$ or $f_i(\tilde{R}_i, R_{-i}) > p(\tilde{R}_i)$, then $f_i(\tilde{R}_i, R_{-i}) = f_i(R)$.

(c) If either $p(\tilde{R}_i) \leq f_i(R)$ or $p(R_i) \geq f_i(\tilde{R}_i, R_{-i})$, then $f_i(\tilde{R}_i, R_{-i}) = f_i(R)$.

(a) On the contrary, assume $f_i(R) > f_i(\tilde{R}_i, R_{-i})$. If $f_i(\tilde{R}_i, R_{-i}) \geq p(R_i)$, then by the single-peakedness of f, i gains by reporting $\tilde{R}_i$ at R, a contradiction with the strategy-proofness of f. Hence, $f_i(\tilde{R}_i, R_{-i}) < p(R_i)$. A similar argument gives $p(\tilde{R}_i) < f_i(R)$. Consequently, $f_i(\tilde{R}_i, R_{-i}) < p(R_i) < p(\tilde{R}_i) < f_i(R)$. We next show that strategy-proofness is violated if $f_i(\tilde{R}_i, R_{-i})$ and $f_i(R)$ in on the opposite sides of $p(R_i)$. Fix any $\bar{R}_i$ with $p(\bar{R}_i) = p(R_i)$ such that i prefers $f_i(\tilde{R}_i, R_{-i})$ to $f_i(R)$ under $\bar{R}_i$. By peak-onliness, $f_i(R) = f_i(\bar{R}_i, R_{-i})$. Thus, i prefers $f_i(\tilde{R}_i, R_{-i})$ to $f_i(\bar{R}_i, R_{-i})$ under $(\bar{R}_i, R_{-i})$, a contradiction with the strategy-proofness of f. This completes the proof of (a).

(b) We concentrate on the $f_i(R) < p(R_i)$ case because the proof of the other case is a mirror image of the current case. By (a), $f_i(R) \leq f_i(\tilde{R}_i, R_{-i})$. Thus, in contradiction to (b), let $f_i(R) < f_i(\tilde{R}_i, R_{-i})$. If $f_i(\tilde{R}_i, R_{-i}) \leq p(R_i)$, then by the single-peakedness of f, i manipulates f at $(\tilde{R}_i, R_{-i})$ by reporting R_i. Consequently, $f_i(R) < p(R_i) < f_i(\tilde{R}_i, R_{-i})$. Because $f_i(R)$ and $f_i(\tilde{R}_i, R_{-i})$ are on the opposite sides of $p(R_i)$, as in the proof of (a), we reach a contradiction with strategy-proofness. Therefore, $f_i(R) = f_i(\tilde{R}_i, R_{-i})$.

(c) As with case (b) we only concentrate on the $p(\tilde{R}_i) \leq f_i(R)$ case. By (a), $f_i(R) \leq f_i(\tilde{R}_i, R_{-i})$. Hence, in contradiction to (c), let $f_i(R) < f_i(\tilde{R}_i, R_{-i})$. Then $p(\tilde{R}_i) \leq f_i(R) < f_i(\tilde{R}_i, R_{-i})$. By the single-peakedness of f, i manipulates f at $(\tilde{R}_i, R_{-i})$ by reporting R_i, a contradiction to the strategy-proofness of f. □

We are now ready to prove Lemma 4.2.