Sequential Preference Revelation in Incomplete Information Settings*

James Schummer[†] Rodrigo A. Velez[‡]

September 4, 2019

Abstract

Strategy-proof allocation rules incentivize truthfulness in *simultaneous move* games, but real world mechanisms sometimes elicit preferences *sequentially*. Surprisingly, even when the underlying rule is strategy-proof and non-bossy, sequential elicitation can yield equilibria where agents have a *strict* incentive to be untruthful. This occurs only under incomplete information, when an agent anticipates that truthful reporting would signal false private information about others' preferences. We provide conditions ruling out this phenomenon, guaranteeing all equilibrium outcomes to be welfare-equivalent to truthful ones.

JEL classification: C79, D78, D82.
Keywords: strategy-proofness; sequential mechanisms; implementation; market design.

1 Introduction

One of the most desirable incentive properties in mechanism design is that of strategy-proofness. It guarantees that when agents *simultaneously* report their preferences to a direct revelation mechanism, each agent has a weak incentive to be truthful. In practice, however, agents sometimes participate in

*We thank Vikram Manjunath, John Weymark, and seminar participants at CoEd15, Concordia University, TETC15, CRETE16, GAMES2016, SAET17, and SCW2018 for useful comments. In particular we thank referees and a Coeditor for helpful comments that have improved the paper. All errors are our own.

†Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston IL 60208; schummer@kellogg.northwestern.edu.

‡Department of Economics, Texas A&M University, College Station, TX 77843; rvelezca@tamu.edu.

such mechanisms *sequentially*: the planner collects preference reports from different agents at different times, sometimes even revealing individual or aggregated preference reports during the collection process. Our objective is to consider how the incentive properties of strategy-proof allocation rules extend to environments in which they are operated as sequential mechanisms.

The analysis of sequential mechanisms becomes increasingly important as technology increases our speed of communication. To illustrate this point, consider the standard school choice problem in which a school district assigns students to schools as a function of their reported preferences. In the past, the practical elicitation of preferences could be done only through the use of physical forms mailed through the postal service. Under such a system, agents (students or families) have little to no information about each others' reports at the time each mails in their own form. Even if mailings occur on different days, the agents are playing a simultaneous-move revelation game. More recently, however, preference elicitation occurs through electronic communication (e.g. email, web forms, or a smartphone app). The speed of such media opens up a new possibility when agents' reports are submitted asynchronously: the planner could choose to *publicly reveal* information about preference reports as they are being submitted. Such feedback occurs, for example, in the school district of Wake County, N.C., USA (Dur et al., 2018), where parents can see aggregated information about previously submitted preferences. In various municipalities of Estonia (Biró and Veski, 2016), the preferences of *individual* families over limited kindergarten slots are listed on a public web site. While this interim information revelation may provide logistical benefits for the agents,[1] the strategic impact of releasing this information is less clear.

Our main results identify conditions under which certain strategy-proof rules are strategically robust to sequential forms of implementation, even under incomplete information. As is already known, however, not *all* strategy-proof rules can be robust in this way. Consider a second-price auction where a fixed order of (private values) bidders sequentially announces their bids. Imagine that the last bidder to act follows the (optimal) strategy that is truthful except that, if any previous bid is higher than her value for the object, she "gives up" and bids zero. Given this strategy, truthful bidding would no longer be optimal for the earlier bidders, under a variety of informational assumptions. It is simple to construct non-truthful equilibria that involve such strategies; they differ from truthful ones in both revenue and efficiency.[2]

[1] By narrowing the list of obtainable schools, a family may require less time to determine their preferences.

[2] For a trivial example, consider the first bidder bidding above the support of the distribution of bidders' values, and all other bidders bidding zero.

This "bad equilibrium" problem of the second-price auction can be attributed to the fact it is *bossy*: a change in one bidder's report, ceteris paribus, can change another bidder's payoff without affecting her own. Anticipating that she will lose the auction anyway, the last bidder's non-truthful decision to bid zero is inconsequential to her. However it is *not* inconsequential to the previous bidders, who can strictly benefit from inducing this non-truthful behavior through non-truthful reports of their own.

The lesson from this example is that, generally speaking, we should not be surprised that the sequential implementation of *bossy*, strategy-proof rules could lead to incentives for non-truthful behavior. On the other hand, many prominent strategy-proof rules are *non-bossy*.[3] Examples include the Top Trading Cycles rule (Shapley and Scarf, 1974; Abdulkadiroğlu and Sönmez, 2003) for school choice and object allocation problems, the Median Voting rule (Moulin, 1980) used for the selection of a public goods level or alternative, and the Uniform Rationing rule (Benassy, 1982; Sprumont, 1991) used to fully allocate a divisible good when agents have satiable preferences. Since the intuition from the auction example does not apply to such rules, we are left with the question of whether the sequential implementation of such rules might preserve the incentive for non-truthful behavior in equilibrium.

There are at least two intuitive reasons to suspect a positive answer to this question. First, the "bossy effect" described above in the sequential second-price auction would no longer hold. If an agent's decision to be non-truthful is inconsequential to her, it must be inconsequential to everyone else, so it would seem that the incentive to induce non-truthful behavior from later-acting agents disappears.

Second, a related result of Marx and Swinkels (1997) also hints at a positive answer to our question, at least in the special case of complete information. Specifically, Marx and Swinkels provide results for general normal-form, complete information games that imply the following corollary. Suppose a (deterministic) sequential revelation mechanism is used to implement a rule that is both strategy-proof and satisfies their "TDI" condition (which is a strong version of non-bossiness). Then every subgame perfect equilibrium of the resulting game yields the same payoffs that would be obtained under truthful reporting. That is, a strong non-bossiness condition (TDI) rules out the kind of non-truthful equilibrium behavior that could occur in the second-price auction described above, as long as there is complete information. Indeed, as a

[3] Formalized in Section 3, *non-bossiness* requires that, if an agent's welfare is unaffected by a misreport of her preferences, then so is the welfare of all agents. Various, similar conditions are defined in the literature starting with Satterthwaite and Sonnenschein (1981).

corollary of our results, this complete information result turns out to be true even if we replace TDI with the weaker non-bossiness condition.

However, as we show in two examples below, this conclusion does *not* hold in the more general case of incomplete information. Specifically, in a standard Bayesian setup, we show that a sequential revelation game derived from a strategy-proof, non-bossy rule can yield sequential equilibria in which (i) payoffs differ from those obtained under truthful revelation, and (ii) a non-truthful agent would be *strictly* worse off by deviating to a truthful report. This is surprising since, as we have stated above, this phenomenon cannot occur under complete information.

This point is illustrated via Example 1, where preferences are solicited according to a fixed, deterministic ordering of the agents and an allocation rule is then applied. As implied by our Theorem 1, the critical feature driving the example is that the prior distribution of preference profiles has non-Cartesian support. A second example (Example 2) covers a broader range of scenarios in which the agents' reporting order is uncertain. Formally it considers a sequential revelation mechanism that randomizes the order in which agents anonymously (but publicly) report preferences, then applies the same allocation rule as in the first example. This example involves independently drawn preferences (hence Cartesian support) but from distributions with different supports, which turns out to be what drives the example.

The intuition behind the occurrence of non-truthful equilibrium outcomes in both examples is summarized by two observations. First is that an (out-of-equilibrium) truthful report by an early acting agent would induce a later agent to place zero weight (interim belief) on the true preference *profile* of all other agents, leading to her incorrect belief that misreporting is costless to her and to everyone else. Second, the earlier agent anticipates this, and finds that he is *strictly better off* avoiding this out-of-equilibrium report (even when it is truthful), in turn leading to a non-truthful outcome.[4]

Following these examples, we investigate the types of information structures that admit the intertwined phenomena in our examples, i.e. non-truthful equilibrium outcomes and a strict disincentive for truthfulness. First, we provide conditions that rule out a sequential revelation mechanism's "failure" to sustain only truthful outcomes (Theorem 1). Second, we show that these same

[4]That is, the phenomenon we highlight is a consequence of the fact that, in a sequential revelation game, preference reports also play a *signaling* role. This suggests that one might subject the sequential equilibria in our examples to the additional scrutiny of signaling games refinements in order to evaluate the sensibility of the belief systems sustaining them. We do this in an Online Appendix, showing that our equilibria satisfy the forward induction arguments introduced by Cho and Kreps (1987) and generalized to our setting by Govindan and Wilson (2009).

conditions also rule out the possibility that an agent faces a strict disincentive to be truthful in equilibrium (Theorem 2). For *fixed order* mechanisms (as in Example 1, where the agents' deterministic reporting order is commonly known) it is sufficient for the prior distribution over type profiles to satisfy a simple *Cartesian Support* (CS) condition, which is automatically satisfied, for instance, in complete information, i.i.d., and full-support settings. In lay terms, the CS condition requires that the private information of a coalition of agents never rule out some admissible preference relation for an agent who does not belong to the coalition. When there is uncertainty in the agents' reporting order (i.e. *random order* mechanisms as in Example 2), a similar condition suffices as long as the underlying rule also satisfies a weak anonymity condition.

All in all, our results document a surprising informational phenomenon that can derail incentives under strategy-proof, non-bossy rules when preference reports are revealed sequentially. Our sufficient conditions ruling out this phenomenon are naturally satisfied in many environments. In such settings, our results provide a valuable tool to planners deciding whether to operate such rules in a sequential form (see Section 5 for further discussion).

The motivation behind other work on extensive-form mechanisms and implementability, briefly discussed here, typically differs from ours. For instance, Moore and Repullo (1988) show how sequential mechanisms can help the planner by expanding the set of implementable social choice functions (SCFs) by explicitly considering *non-revelation* mechanisms under complete information.[5] Li (2017) considers whether a strategy-proof SCF can be implemented by such mechanisms in "obviously" dominant strategies. In contrast to such work, we are motivated by the observation that some planners have already committed to the use of a specific SCF as a revelation mechanism, but are making the *design choice* to elicit preferences sequentially rather than simultaneously. Indeed when preferences are elicited simultaneously in complete information environments, (Saijo et al., 2007) show that a strategy-proof, non-bossy SCF avoids the sort of "bad equilibrium" problem described above only when it also satisfies a particularly strong "rectangularity" property.[6]

The remainder of the paper is organized as follows. Section 2 presents the

[5] The experimental results of Klijn et al. (2019) also suggest that players can be better strategists in dynamic mechanisms that in static ones.

[6] None of the SCF's we mentioned earlier satisfies it. Indeed, certain forms of dictatorial rules are typically the only strategy-proof, non-bossy SCF's that do (Saijo et al., 2007; Bochet and Sakai, 2010; Fujinaka and Wakayama, 2011). On the other hand the equilibrium problem can be avoided by appealing to coalitional refinements (Bochet and Tumennassan, 2017) or "empirical plausibility" refinements (Velez and Brown, 2019).

examples described above. Section 3 introduces our model while Section 4 presents our results. In Section 5 we summarize and interpret our results. Proofs are relegated to the Appendix.

2 Two Examples

In the interest of brevity and simplicity, we provide our examples using minimal notation and definitions since the terminology is standard. Readers unfamiliar with the concepts may refer to Section 3 for formal definitions. The Appendix contains proofs that our examples are indeed sequential equilibria.

In both examples, we consider the *rationing problem* (Benassy, 1982), where an endowment Ω of a divisible good must be divided amongst four agents, $N = \{1,2,3,4\}$, each with single-peaked preferences. Each agent $i \in N$ has a privately known peak level of consumption, p_i, so that consuming x units of the good yields a payoff of $-|p_i - x|$.[7] Profiles of peaks are drawn from a prior distribution specific to each example.

A well-studied rule for this problem is the Uniform rationing rule (Benassy, 1982; Sprumont, 1991) which works as follows. In "deficit" cases (where the sum of the agents' peaks exceeds Ω), the Uniform-rationing rule allocates an equal share of the good, say λ, to all agents with the exception that any agent i for whom $p_i < \lambda$ receives her peak amount, p_i. Similarly in "surplus" cases (where Ω exceeds the sum of peaks), all agents receive some common share λ, with the exception that if $p_i > \lambda$, then i receives her peak amount, p_i.[8] The Uniform rule is both strategy-proof and non-bossy (see Section 3).

In our first example, the planner attempts to implement the Uniform rule by sequentially soliciting the agents' peaks according to a *fixed order*. Each agent observes the reports of previous agents, and uses this information to update beliefs about the remaining agents' (correlated) preferences. One can think of this procedure as representing roll-call voting, e.g. when preferences are revealed according to the order in which people are sitting around a table. In our second example, preferences are drawn independently but the planner randomizes the order in which he solicits the agents' reports. Each agent observes the previous reports—but not the identity of those who made them—and uses this information to update beliefs about the *identity* of the remaining agents. For each example we describe equilibrium payoffs which differ from payoffs under truth-telling behavior.

[7] Our restriction to piecewise linear payoff functions is unimportant and merely simplifies the example. It also means that peaks correspond to types in our general model.

[8] In both cases, this definition implicitly defines a unique λ.

2.1 Example 1: Fixed-order revelation, correlated types

There are $\Omega = 4$ units to be divided and, for simplicity, the set of possible peaks is restricted to be $\{0, 1, 2, 2.5, 3\}$. We further restrict the possible *combinations* of agents' peaks via the common prior beliefs over types. Specifically, assume there are six equally likely profiles of peaks listed in Table 1.

p_1	p_2	p_3	p_4
2	1	2	2
2.5	1	0	0
3	1	0	0
2	2	2	2
2.5	2	0	0
3	2	0	0

Table 1: The common prior belief is that these six profiles of peaks are equally likely.

Observe that Agent 2's peak is equally likely to be $p_2 = 1$ or $p_2 = 2$ independently of the other agents' peaks. The list of the other three agents' peaks, (p_1, p_3, p_4), is one of three equally likely subprofiles, $(2, 2, 2)$, $(2.5, 0, 0)$, or $(3, 0, 0)$.

We consider the extensive form, incomplete information game in which Agents 1–4 sequentially (and in numerical order) publicly announce their peaks and the Uniform rule is applied to those announcements. A (mixed) strategy for player i maps the agent's peak p_i and observed history of $i-1$ previous reports into a lottery over reports.

Let σ be the (pure) strategy profile in which each agent always truthfully reports her peak, with two exceptions:

(i) if Agent 1's peak is $p_1 = 3$, then Agent 1 reports a peak of 2.5;

(ii) if Agent 1 has reported a peak of 3, then Agent 2 reports a peak of 2 (regardless of p_2).

Observe that when the profile of agents' peaks is $(3, 1, 0, 0)$, σ prescribes reports of $(2.5, 1, 0, 0)$. Thus the outcome under these reports, $(2.5, 1, 0.25, 0.25)$, differs from the Uniform allocation for these peaks, $(3, 1, 0, 0)$. Furthermore, if Agent 1 were to deviate to a truthful strategy when $p_1 = 3$, she is strictly worse off when $p_2 = 1$ (and otherwise indifferent): the resulting reports under σ_{-1} would be $(3, 2, 0, 0)$, yielding an allocation of $(2, 2, 0, 0)$.

Nevertheless, as we show in the Appendix, there exists a belief system β such that (σ, β) is a sequential equilibrium. What drives the example is the fact that, following an (off equilibrium path) report of 3 by Agent 1, Agent 2 (under β_2) believes with certainty that $p_1 = 2$, and hence that $(p_1, p_3, p_4) = (2, 2, 2)$. That is, after Agent 1 reports a "large" peak of 3, Agent 2 believes that the next two agents also will report relatively "large" peaks ($p_3 = p_4 = 2$), and so she will be allocated one unit of the good regardless of her report (as long as it is one or greater). She believes her misreport to be costless since her (possibly false) inference about p_1 gives her (possibly false) certainty about future reports. This occurs due to the extreme correlation across preferences, and gives an intuition behind our main results.

Backing up a step, if Agent 1 were to truthfully report a peak of $p_1 = 3$, she foresees being hurt by this because it could inflate the reported peak of Agent 2. In terms of beliefs, Agent 1 realizes that Agent 2's misreport will lead to her being "surprised" by the truthful reports of Agents 3 and 4 that do indeed (negatively) correlate with 1's report. Ex post, both Agents 1 and 2 would have been harmed by Agent 1's truthful report. Agent 1 foresees this, and averts this chain of events by misreporting her peak. Agent 1's incentive to do this is strict, despite the strategyproofness and non-bossiness of the Uniform rule.

2.2 Example 2: Random-order revelation, independent types

Our first example showed that, with correlated preferences, sequential revelation can lead to undesirable outcomes even when the underlying rule is both strategy-proof and non-bossy. Our second example shows that, even without correlation of preferences, undesirable outcomes can occur when the agents anonymously announce preference in *random* order.[9] In the example, an out-of-equilibrium truthful report would cause an agent to falsely infer the *identity* of an agent who has already reported and, because of an asymmetry in the sets of agents' possible types, to believe with certainty that a misrepresentation of preferences would be costless. This in turn strictly discourages some agents from making that report—even if it would be truthful—in the first place.

We consider a rationing problem identical to Example 1, except that there are now $\Omega = 8$ units of the good to be divided, and the distribution of peaks is changed. Peaks are drawn independently, but not identically, across agents ac-

[9] As a special case of such mechanisms, one can imagine a planner that periodically announces updated, aggregate statistics of reported preferences to date. Agents who have not yet reported can view information about previous reports without learning the identities of those who reported.

cording to the distributions in Table 2, where probability α is "somewhat close" to 1.[10]

	peak				
	0	2	2.5	2.9	3
Agents 1–3	0	0	α	$\frac{1-\alpha}{2}$	$\frac{1-\alpha}{2}$
Agent 4	α	$\frac{1-\alpha}{2}$	0	0	$\frac{1-\alpha}{2}$

Table 2: Each agent's peak is independently drawn according to the given distribution, where α is close to one.

Consider the extensive form game in which the agents report their peaks in a uniformly random order. Agents observe their own positions in the ordering and the previous agents' reports (but not who made them). The Uniform rule is then applied to those announcements. As before, a (mixed) strategy for player i maps the agent's peak p_i and observed history of reports into a lottery over reports.

We construct a non-truthful, sequential equilibrium that has positive probability of resulting in an inefficient outcome. The players are truthful on the equilibrium path with the exception that no agent, when chosen first to report, ever reports a peak of 3. This behavior is sustained by the possible "threat" of agents 1–3 responding with subsequent non-truthful reports of a peak of 3, regardless of their actual peaks. Let σ be the (pure) strategy profile in which each agent always truthfully reports her peak, with these exceptions:

(i) If Agent $i \in \{1,2,3\}$ is chosen to report first and her peak is $p_i = 3$, then she reports a peak of 2.9;

(ii) If Agent 4 is chosen to report first and her peak is $p_4 = 3$, then she reports a peak of 2;

(iii) If Agent $i \in \{1,2,3\}$ is not chosen to report first and all previous reports have been 3, then she also reports a peak of 3.

Observe that, for example, when the agents are asked to report in numerical order, and when the profile of peaks is $(3,2.5,2.5,0)$, the agents report peaks of $(2.9,2.5,2.5,0)$ under σ, yielding an allocation of $(2.9,2.5,2.5,0.1)$ to the agents. This differs from the Uniform allocation that would result under truthful reports, namely giving each agent her peak. In fact, when Agent $i \in \{1,2,3\}$ is

[10] Details are in the Appendix. The example does not depend at all on the same parameter α being applied to all four agents; this assumption merely simplifies exposition.

chosen to report first and $p_i = 3$, simple calculations show that the agent has a *strict* incentive to misreport her peak to be 2.9.[11]

Nevertheless, there exists a belief system β such that (σ, β) is a sequential equilibrium (see Appendix). What drives this example is an idea similar to that in Example 1, but that occurs through a different randomization device: the ordering of agents. Following any (off equilibrium path) report of 3 in the first round, any subsequent agent $i \neq 4$ believes (under β_i) with certainty that the report came from Agent 4. Given this belief in an "inflated" report from Agent 4, any other agent acting in, say, round 2 believes with certainty that *regardless of her report*, all subsequent reports will be peaks of 2.5 or greater, and hence any such report *she* makes will grant her an equal share (2 units) of Ω. It is the "certain belief" that Agent 4 acted first that leads to the certain belief about subsequent reports, which in turn leads to certain belief that she cannot harm herself with a misreport. This occurs in part due to the fact that the support of the marginal distributions of preferences vary across agents. Finally, as in Example 1, if we back up a step to Round 1, an agent $i \in \{1, 2, 3\}$ can foresee all of this, and realizes that by (truthfully) reporting a peak of 3, she potentially triggers exaggerated reports from the other two high-demand agents ($j, k \notin \{i, 4\}$). This strictly encourages a misreport in round 1.

3 Model

3.1 Environment

There is an arbitrary set of alternatives A and a set of agents $N \equiv \{1, \ldots, n\}$, $n \geq 2$. Each agent $i \in N$ has expected-utility preferences on the space of measures on A, $\Delta(A)$, represented by utility function u_i belonging to some domain of utility functions $\mathscr{U}$, which is assumed to be countable. Denote a profile of utility functions as $u \equiv (u_i)_{i \in N} \in \mathscr{U}^N$. Utility functions are private information, but are drawn from a commonly known prior probability measure,

[11] Intuitively, if i reports 2.9, then with large probability the other three agents truthfully report (2.5,2.5,0) in some order, giving i 2.9 units. If i reports 3, then with large probability agent 4 reports 0 and the remaining agents report 2.5 or 3, depending on whether they report after or before Agent 4, respectively. Thus with large probability the others' reports are equally likely to be (3,3,0), (3,0,2.5), or (0,2.5,2.5). Respectively, these reports give i either 2.67, 2.75, or 3 units of the good, a lottery worse than 2.9 units for sure when $p_i = 3$. As long as α is sufficiently large, a misreport of 2.9 is strictly better.

$\mu \in \Delta(\mathscr{U}^N)$.[12] We endow $\Delta(\mathscr{U}^N)$ with the l^1 norm. Assumptions about μ, which play a central role in our results, are described in Subsection 3.3.

For each pair of disjoint sets $S, T \subset N$, each $u_S \in \mathscr{U}^S$, and each $u_T \in \mathscr{U}^T$, we let (u_S, u_T) denote the sub-profile obtained by joining u_S and u_T. We write (u_{-T}, u_T) in place of $(u_{N\setminus T}, u_T)$ and write (u_{-i}, u_i) in place of $(u_{N\setminus\{i\}}, u_{\{i\}})$. For $S, T \subseteq N$ with $|S| = |T|$, we write $[u_S] = [u_T]$ to denote that $u_S \in \mathscr{U}^S$ is a relabeling of $u_T \in \mathscr{U}^T$, i.e.,

$$[u_S] = [u_T] \quad \Leftrightarrow \quad \exists \text{ bijection } \zeta : S \to T \text{ s.t. } \forall s \in S,\ u_s = u_{\zeta(s)}.$$

A social choice function (SCF), $f : \mathscr{U}^N \to A$, associates with each utility profile $u \in \mathscr{U}^N$ an alternative $f(u)$. The following four conditions on SCF's are central to the analysis. The first three are standard, and the fourth is a weakened version of a standard anonymity condition. For any SCF f on a given domain $\mathscr{U}^N$, we say that

- f is **strategy-proof** when for each $i \in N$, $u \in \mathscr{U}^N$, and $v_i \in \mathscr{U}$, $u_i(f(u_{-i}, u_i)) \geq u_i(f(u_{-i}, v_i))$.
- f is **non-bossy** (in welfare) when for each $i \in N$, $u \in \mathscr{U}^N$, and $v_i \in \mathscr{U}$,

$$u_i(f(u)) = u_i(f(u_{-i}, v_i)) \implies \forall j \in N,\ u_j(f(u)) = u_j(f(u_{-i}, v_i)).$$

- f is **non-bossy*** (in welfare/outcome) when for each $i \in N$, $u \in \mathscr{U}^N$, and $v_i \in \mathscr{U}$,

$$u_i(f(u)) = u_i(f(u_{-i}, v_i)) \implies f(u) = f(u_{-i}, v_i).$$

- f is **weakly anonymous** when it is welfare invariant under any permutation of the other agents' reports, i.e. when for each $i \in N$, $u_i \in \mathscr{U}$, and $v, v' \in \mathscr{U}^N$, we have

$$v_i = v'_i \text{ and } [v] = [v'] \implies u_i(f(v)) = u_i(f(v')).$$

Non-bossiness* plays a role in strengthening the conclusion of our main theorem that guarantees outcome equivalence. It is satisfied, for example, by the

[12] Our results generalize to a model of non-common priors at the expense of additional notation. As part of our analysis we prove that, under the assumptions of Theorem 1 and Theorem 2, each agent is behaviorally equivalent to a truthful agent. Thus even if agent types included a prior over others' types, an agent's expected utility would be calculated with respect to the induced distribution of payoff types.

Uniform rationing rule, Median Voting rule, and TTC in their respective standard environments. The weak anonymity condition is implied by the standard *anonymity* condition that requires an agent's consumption to be invariant to permutations of the other agents' reports. Since weak anonymity plays the role of a sufficient condition in our results, this weaker definition strengthens our results.[13] Most allocation rules considered in the literature satisfy the standard anonymity condition, including the Uniform-rationing rule discussed in our examples. To see that our condition weakens the standard one non-trivially, consider the so-called "impartial division of a dollar" problem (de Clippel et al., 2008). Many strategy-proof rules exist for that problem that are weakly anonymous yet violate the usual (stronger) anonymity condition.[14]

To describe the order in which agents sequentially report preferences we denote a generic permutation of N as $\pi\colon N \to N$, letting Π be the set of all permutations of N. We interpret π as a *mapping of positions to agents*, so $\pi(t)$ denotes the t-th agent in the ordering. Finally let $\Delta(\Pi)$ be the space of lotteries on Π with generic element $\Lambda \in \Delta(\Pi)$.

Given an SCF f, and a lottery $\Lambda \in \Delta(\Pi)$ over orderings, we consider the following **extensive game form with imperfect information**, denoted by $\Gamma(\Lambda, f)$.

Round 0: Nature randomly determines both a permutation $\pi \in \Pi$ according to Λ and a preference profile $u \in \mathscr{U}^N$ according to $\mu \in \Delta(\mathscr{U}^N)$; each agent i privately learns u_i.

Round 1: Agent $\pi(1)$ reports a preference relation $u'_{\pi(1)} \in \mathscr{U}$; all agents observe the report $u'_{\pi(1)}$ but not the identity of $\pi(1)$.

Round t ($2 \le t \le n$): Given the history of $t-1$ previous reports, which we denote by $h_{t-1} \in \mathscr{U}^{t-1}$, Agent $\pi(t)$ reports a preference relation $u'_{\pi(t)} \in \mathscr{U}$; all agents observe the report $u'_{\pi(t)}$ but not the identity of $\pi(t)$.

End: The outcome $f(u')$ is chosen.

Each agent implicitly knows her own position in the realized ordering π by

[13] In fact an even weaker definition could be used that requires invariance in permutations of reports, but only for the random orders of reports that could occur with positive probability in the sequential revelation game form (i.e. Λ defined below). Unfortunately, formalizing this condition would add complexity while offering little additional insight. For the sake of readability we therefore omit this definition.

[14] These strategy-proof rules, characterized by de Clippel et al. (2008, Theorem 1), determine an agent's share of the dollar using a personalized "aggregator" of the other agents' reports. Such rules are weakly anonymous as long as all aggregators are symmetric; they fail the stronger anonymity condition whenever the aggregators differ.

observing the length of history h_{t-1} when she makes her report in round t.[15] Of course, if Λ is deterministic, each agent knows the identity of each reporting agent. A deterministic Λ represents a simple roll-call voting procedure, e.g. agents sitting around a table, publicly announcing reports in a foreseeable order.

We denote the set of all histories by $H = \{h_t \in \mathscr{U}^t : 0 \leq t \leq n\}$. When we arbitrarily write $h_t \in H$ it is to be understood that the length of h_t is $t \in \{0, 1, \ldots, n\}$. At the beginning of Round 1, Agent $\pi(1)$ faces the trivial history denoted by $h_0 \equiv \emptyset$.

Depending on the distribution Λ, Agent i might have zero probability of being assigned to certain positions in the sequence, making it infeasible for that agent to see histories of certain lengths. Denote the set of "feasible" histories that Agent $i \in N$ could face by

$$H^i \equiv \{h_t \in H : \exists \pi \in \Pi \text{ s.t. } \Lambda(\pi) > 0 \text{ and } i = \pi(t+1)\}.$$

The dependence of H^i on Λ is dropped from the notation since Λ is typically fixed and clear from context. Of course some histories in H^i could be ruled out depending on the strategies of agents, but this is not relevant to this definition.

In the game $\Gamma(\Lambda, f)$, a (mixed, behavior) **strategy** for agent $i \in N$ is a function that maps her utility function and possible history into a randomized report, $\sigma_i : \mathscr{U} \times H^i \to \Delta(\mathscr{U})$, where $\Delta(\mathscr{U})$ is the space of countably additive probability measures on $\mathscr{U}$. Conditional on $u_i \in \mathscr{U}$ and $h \in H^i$, the probability that i makes a report of v_i under σ_i is denoted by $\sigma_i\langle u_i, h\rangle(v_i)$. A **strategy profile** is denoted by $\sigma \equiv (\sigma_i)_{i \in N}$. A **strategy subprofile** for $S \subseteq N$ is denoted by $\sigma_S \equiv (\sigma_i)_{i \in S}$; similarly for arbitrary $T \subseteq S$, $(\sigma_{S \setminus T}, \sigma'_T)$ denotes the strategy subprofile obtained by combining the lists $\sigma_{S \setminus T}$ and σ'_T. For each agent $i \in N$ the **truthful strategy**, denoted by τ_i, is the one that for each realized u_i and each history, reports u_i.

3.2 Sequential equilibria

We define the standard notion of sequential equilibrium (Kreps and Wilson, 1982). A belief function for agent $i \in N$ specifies, for each utility function and history seen by i, a distribution over player sequences (Π) and preferences of the other agents. Specifically it is a function $\beta_i : \mathscr{U} \times H^i \to \Delta(\Pi \times \mathscr{U}^{N \setminus \{i\}})$ where $\Delta(\Pi \times \mathscr{U}^{N \setminus \{i\}})$ is the set of countably additive measures on permutations of N and the other agents' types. Conditional on $u_i \in \mathscr{U}$ and $h \in H^i$, the probability

[15] The analysis of the corresponding game in which agents learn only the *set* of the previous reports (but not their relative order) is analogous to ours.

that β_i puts on a permutation and preference subprofile of the other agents is denoted by $\beta_i\langle u_i, h\rangle(\pi, u_{-i})$. A **belief system** is a profile of belief functions $\beta \equiv (\beta_i)_{i\in N}$.

An **assessment** is a pair (σ, β) of a strategy profile σ and belief system β. The assessment is **consistent** for $\langle\Gamma(\Lambda, f), \mathscr{U}^N, \mu\rangle$ if there is a sequence of assessments $\{(\sigma^k, \beta^k)\}_{k\in\mathbb{N}}$ such that for each $k \in \mathbb{N}$, (i) σ^k has full support,[16] (ii) β^k is obtained from Bayes' rule given Λ, μ, and σ^k, and (iii) as $k \to \infty$, $(\sigma^k, \beta^k) \to (\sigma, \beta)$.[17]

The definition of sequential rationality—that each agent is playing a best response at every possible information set—requires notation to denote the conditional probability of future paths of play. For $\pi \in \Pi$ and $t, s \in \{0, \dots, n\}$ such that $s > t$, we denote the set of predecessors and the set of $s-t$ successors of agent $\pi(t)$ in π by $\pi(1, ..., t-1)$ and $\pi(t+1, ..., s)$, respectively. Additionally, for types $u_{\pi(t+1,...,n)} \in \mathscr{U}^{\pi(t+1,...,n)}$, history $h_t \in \mathscr{U}^t$, and reports $v_{\pi(t+1,...,s)} \in \mathscr{U}^{\pi(t+1,...,s)}$, we let

$$\sigma(h_{\pi(t+1,...,s)} | h_t, \pi, u_{\pi(t+1,...,s)})$$

denote the probability of realizing final history $(h_t, v_{\pi(t+1,...,s)})$ under strategy profile σ conditional on: π being the selected permutation, agents $\pi(1, \dots, t)$ having selected actions h_t, and agents $\pi(t+1, \dots, s)$ having types $u_{\pi(t+1,...,s)}$.

Assessment (σ, β) is a **sequential equilibrium** of $\langle\Gamma(\Lambda, f), \mathscr{U}^N, \mu\rangle$ if it is consistent and for each $i \in N$, each $u_i \in \mathscr{U}$, each $t \in \{0, \dots, n-1\}$, and each $h_{t-1} \in H^i$, $\sigma_i\langle u_i, h_{t-1}\rangle$ is **sequentially rational**: for each $v_i \in \text{supp}(\sigma_i\langle u_i, h_{t-1}\rangle)$ and each deviation $w_i \in \mathscr{U}$, the expected utility from reporting v_i, namely

$$\sum_{\pi, u_{-i}} \sum_{h_{\pi(t+1,...,n)}} u_i\left(f\left(h_{t-1}, v_i, h_{\pi(t+1,...,n)}\right)\right) \sigma\left(h_{\pi(t+1,...,n)} | (h_{t-1}, v_i), \pi, u_{\pi(t+1,...,n)}\right) \beta_i\langle u_i, h_{t-1}\rangle(\pi, u_{-i})$$

is greater than or equal to that from reporting w_i, namely

$$\sum_{\pi, u_{-i}} \sum_{h_{\pi(t+1,...,n)}} u_i\left(f\left(h_{t-1}, w_i, h_{\pi(t+1,...,n)}\right)\right) \sigma\left(h_{\pi(t+1,...,n)} | (h_{t-1}, w_i), \pi, u_{\pi(t+1,...,n)}\right) \beta_i\langle u_i, h_{t-1}\rangle(\pi, u_{-i}).$$

Denote the set of sequential equilibria by $\text{SE}\langle\Gamma(\Lambda, f), \mathscr{U}^N, \mu\rangle$.

3.3 Information structure

Our results are centered around two conditions on the prior beliefs $\mu \in \Delta(\mathscr{U}^N)$. The first condition applies to our results on sequential reporting when the or-

[16] Profile σ has full support if for all $i \in N$, $u_i \in \mathscr{U}$, $h_t \in H^i$, and $v_i \in \mathscr{U}$, $\sigma_i\langle u_i, h_t\rangle(v_i) > 0$.

[17] Convergence is point-wise, i.e. fixing any $i \in N$, $u_i \in \mathscr{U}$, $h_t \in H^i$, $v_i \in \mathscr{U}$, and $(\pi, v_{-i}) \in \Pi \times \mathscr{U}^{N\setminus\{i\}}$, we have $\sigma_i^k\langle u_i, h_t\rangle(v_i) \to \sigma_i\langle u_i, h_t\rangle(v_i)$ and $\beta_i^k\langle u_i, h_t\rangle(\pi, v_{-i}) \to \beta_i\langle u_i, h_t\rangle(\pi, v_{-i})$.

der Λ is deterministic. It states that the support of μ can be written as a Cartesian product of n subsets of $\mathscr{U}$.

Definition 1. *A prior $\mu \in \Delta(\mathscr{U}^N)$ has* ***Cartesian support*** *if its support is the cross product of n non-empty subsets of $\mathscr{U}$, i.e. for any $u, v \in \text{supp}(\mu)$ and $j \in N$ we have $\mu(v_{-j}, u_j) \in \text{supp}(\mu)$. Denote the set of such priors by $\mathscr{M}_{\text{Cartesian}}$.*

This condition is violated in Example 1.[18] It is satisfied by any prior with full support on $\mathscr{U}^N$, but $\mathscr{M}_{\text{Cartesian}}$ also allows for asymmetric sets of possible utility functions across agents. For example, the prior in Example 2 has Cartesian support simply due to independence across agents' preferences. More generally, the Cartesian support condition merely requires the *support* for beliefs over agent j's preferences to be the same for any given realization of u_{-j}.

To obtain our results for the case in which the agents' reporting order, Λ, is non-deterministic (as in Example 2), we introduce the next condition. It rules out a preference report that could cause an agent to make "absolutely certain" conclusions (correctly or not!) about the identity of the agent who made the report. This requires strengthening the Cartesian support condition so that agents' possible preferences come from the same set.

Definition 2. *A prior $\mu \in \Delta(\mathscr{U}^N)$ has* ***symmetric Cartesian support*** *if, for some set $\mathscr{V} \subseteq \mathscr{U}$, $\text{supp}(\mu) = \mathscr{V}^N$. Denote the set of such priors by $\mathscr{M}_{\text{symm-Cartesian}}$.*

The condition requires that, if any profile of types occurs with positive probability, then so does any permutation of that profile. Obviously it implies the condition in Definition 1.

4 Results

We begin with the straightforward observation that, for any strategy-proof SCF f and any randomization of sequences Λ, there is an "always truthful" equilibrium in the sequential revelation game. The proof is obvious, but for completeness is provided in the Appendix along with all other proofs.

Proposition 1 (Truth-telling equilibrium)**.** *For any strategy-proof SCF f, any distribution $\Lambda \in \Delta(\Pi)$, and any prior $\mu \in \Delta(\mathscr{U}^N)$, truthful reporting is an equilibrium behavior: there exist beliefs β such that $(\tau, \beta) \in SE\langle\Gamma(\Lambda, f), \mathscr{U}^N, \mu\rangle$.*

We are interested in the converse question: When does an arbitrary sequential equilibrium of the sequential revelation game necessarily lead to

[18] E.g. there is no profile in the support of the prior where $p_1 = 3$ and $p_3 = 2$ simultaneously.

truthful outcomes, at least in welfare terms? The auction example discussed in the Introduction motivates our restriction to non-bossy SCF's. Even with this restriction, Example 1 and Example 2 show that for some information structures, a sequential revelation game for a strategy-proof, nonbossy f yields equilibrium outcomes that yield payoffs different from those under truthful reporting. Our main result is that, if we rule out those kinds of information structures, then *all* sequential equilibria must be payoff-equivalent to a truthful equilibrium.

Theorem 1 (Truth-equivalent payoffs)**.** *Let f be a strategy-proof and non-bossy SCF and let $\Lambda \in \Delta(\Pi)$. Suppose that at least one of the following conditions holds.*

1. *The prior has Cartesian support ($\mu \in \mathcal{M}_{\text{Cartesian}}$) and Λ is deterministic.*
2. *The prior has symmetric Cartesian support ($\mu \in \mathcal{M}_{\text{symm-Cartesian}}$) and f is weakly anonymous.*

Then equilibrium outcomes are welfare-equivalent to truthful ones: For each $(\sigma, \beta) \in SE\langle\Gamma(\Lambda, f), \mathcal{U}^N, \mu\rangle$, each $u \in \text{supp}(\mu)$, and each final history of reports $h_N \in \mathcal{U}^N$ reached with positive probability given σ, Λ, and u, we have $u(f(h_n)) = u(f(u))$. If f is also non-bossy then we also have $f(h_N) = f(u)$, i.e. equilibrium outcomes are equivalent to truthful ones.*

Theorem 1 establishes our main point, namely that even if a sequential equilibrium involves non-truthful strategies, all of the agents receive payoffs as if everyone had been truthful. Observe however that this statement still allows for an agent to face a strict incentive to misreport preferences in equilibrium; recall that this occurs in both examples of Section 2. It turns out that this phenomenon also can be ruled out under the same set of assumptions as in Theorem 1: in any sequential equilibrium, an agent's truthful strategy must be sequentially rational with respect to the other agents' strategy profile.

Theorem 2. *Under the same assumptions as in Theorem 1, truthful behavior is sequentially rational with respect to any sequential equilibrium: for any $(\sigma, \beta) \in SE\langle\Gamma(\Lambda, f), \mathcal{U}^N, \mu\rangle$ and any $i \in N$, the truth telling strategy τ_i is sequentially rational for i with respect to σ_{-i} and β_i.*

In the remainder of this section, we explain the reasoning behind the argument that proves Theorem 1, relegating the formalization to the Appendix. Theorem 2 follows from an an extension of these arguments that has essentially the same structure. Since the latter proof adds little insight to our main point, its formalization is provided in a separate Online Appendix.[19]

To explain the reasoning behind Theorem 1, we first point out how the conclusion would be quite straightforward in the two-agent case. (Indeed, the theorem's conclusions hold in the two-agent case whenever f is strategy-proof and non-bossy, even without the remaining assumptions on μ, Λ, and f.) In doing so, we observe a step in the reasoning at which the simplicity of the argument breaks down with $n \geq 3$ agents, which in turn reveals the point where the Cartesian support conditions play a role in deriving our results.

For the two-agent case, let f be a strategy-proof, non-bossy SCF and consider the deterministic sequential revelation game in which Agent 1 reports her preferences first. If Agent 1 reports $v_1 \in \mathscr{U}$, a best response for Agent 2 with utility u_2 is any report v_2 that maximizes $u_2(f(v_1, v_2))$. By strategy-proofness the truthful report u_2 is one best response, so for *any* best response v_2 we have $u_2(f(v_1, v_2)) = u_2(f(v_1, u_2))$. Since f is non-bossy, this implies $u_1(f(v_1, v_2)) = u_1(f(v_1, u_2))$. That is, *as long as Agent 2 is best-responding, Agent 1 receives a payoff as if Agent 2 were committed to being truthful* across the set of admissible utility functions for Agent 2.

Given this fact, a best response for Agent 1 with utility u_1 is any report v_1 that maximizes $E[u_1(f(v_1, u_2))]$, where the expectation is with respect to $\mu(\cdot|u_1)$. Strategy-proofness again implies that the truthful report u_1 is one such best response. However our results depend on the converse question: when is a non-truthful v_1 also optimal? This is where a subtle observation is relevant in extending the argument to more than two agents.

For v_1 to be an optimal report for u_1, we must have

$$\sum_{u_2 \in \mathscr{U}} u_1(f(v_1, u_2))\mu(u_2|u_1) \geq \sum_{u_2 \in \mathscr{U}} u_1(f(u_1, u_2))\mu(u_2|u_1)$$

Of course strategy-proofness implies the opposite inequality *pointwise*:

$$\forall u_2 \in \mathscr{U}, \quad u_1(f(v_1, u_2)) \leq u_1(f(u_1, u_2)).$$

Hence if v_1 is optimal we must have $u_1(f(v_1, u_2)) = u_1(f(u_1, u_2))$ for *any* u_2 *such that* $\mu(u_2|u_1) > 0$. The latter qualification is important: v_1 could be optimal for u_1 even though $u_1(f(v_1, u_2)) < u_1(f(u_1, u_2))$ for some u_2's that have zero probability under $\mu(\cdot|u_1)$. If such u_2's had positive probability, the first inequality above would be violated. At this point we reach the conclusion of Theorem 1

[19] In fact the Online Appendix provides a technically stronger result, stating that an agent's equilibrium strategy can be *replaced* with her truthful strategy in a way that, with appropriately modified beliefs, yields a new, payoff-equivalent equilibrium.

for the two-agent case. The previous equality, with non-bossiness, implies that $u_2(f(v_1, u_2)) = u_2(f(u_1, u_2))$ for unit mass of u (under μ). Thus for each $i \in N$, $u_i(f(v_1, v_2)) = u_i(f(u_1, u_2))$.

What made the argument simple in the two-agent case is the following: the only agent who needs to forecast another agent's type (i.e. Agent 1) is also the first agent to act. Thus this agent's beliefs are necessarily determined only by Bayesian updating the prior (μ) with respect to her private information (u_1). In particular this means that, fixing u_1, for any utility function u_2 of Agent 2 for which $\mu(u_1, u_2) > 0$, Agent 1 must anticipate u_2 with positive probability since his beliefs are precisely $\mu(\cdot|u_1)$.

When $n \geq 3$, however, an agent who acts neither first nor last has to form beliefs about later-to-act agents' types, based both on the prior *and* on the earlier agents' reports. This is what led to the phenomenon of Example 1: following an out-of-equilibrium report by Agent 1, Agent 2 would place zero weight on the chance of being "punished" by certain, future truthful reports (by Agents 3 and 4).

So let us reconsider the above argument for the general case $n \geq 3$, supposing again that agents report preferences in the deterministic sequence $1, 2, \ldots, n$. As before, for any reports v_{-n} of the first $n-1$ agents and for any realized utility u_n, truth-telling is a best response for Agent n and hence any best response provides Agent n with the same payoff as truth-telling. By non-bossiness each of the other agents also receives a payoff as if Agent n reported u_n. Thus earlier agents can behave *as if Agent n were committed to being truthful* across the set of Agent n's "foreseeable" utility functions u_n. Where the argument breaks down is in determining precisely which u_n's are foreseeable, since this is an implication of interim beliefs. Out of equilibrium beliefs of intermediate acting agents need not be a Bayesian update of μ. As illustrated in Example 1 and Example 2, following an off-equilibrium report, such an intermediate agent may place zero probability on the report being truthful and consequently could make a report that is, to the "surprise" of that agent, *not* payoff equivalent to a truthful one. In turn this can unravel the incentive for truthful reporting earlier in the game: earlier agents anticipate this agent's (mis)belief and are forced to report non-truthfully.

These arguments suggest that one *can* reach the same conclusion as in the two-agent case if one guarantees that, in a sequential equilibrium, each agent's beliefs place positive probability on types that force her to act as if she were truthful. In order to identify such conditions, we begin with two intuitive observations that apply to any consistent assessment, (σ, β). First, whenever an agent is called upon to report her preferences, she cannot anticipate an ex-ante impossible event with positive probability. The second statement concerns the beliefs of an agent—who acts at some interim stage of the sequential game—about the preferences of agents who have yet to act. Of course these beliefs are affected by the actions of previous agents: those actions are a function of their preferences, which correlate with later agents' preferences. What the second statement says is the converse of this idea: the history of play (h_{t-1}) influences a player's beliefs about future agents' preferences *only* to the extent that h_{t-1} influences that player's beliefs about previous agent's preferences.[20]

Lemma 1. *Fix a SCF f, a distribution of sequences $\Lambda \in \Delta(\Pi)$, and an assessment (σ, β) that is consistent for $\langle \Gamma(\Lambda, f), \mathscr{U}^N, \mu \rangle$.*

1. *Whenever an agent is asked to report her preferences, her beliefs place positive probability only on events that have positive prior probability, i.e., for each $u \in \text{supp}(\mu)$, $\pi \in \Pi$, $i \in N$, $h_{t-1} \in H^i$, and $u'_{-i} \in \mathscr{U}^{N \setminus \{i\}}$,*

$$\beta_i \langle u_i, h_{t-1} \rangle (\pi, u'_{-i}) > 0 \quad \Longrightarrow \quad \pi(t) = i,\ \Lambda(\pi) > 0,\ \mu_i(u'_{-i}, u_i) > 0$$

2. *At any history of play an agent does not learn anything about the types of the agents who she believes have yet to act, other than what she learns through her own type, and the history of play. That is, for each $u \in \mathscr{U}^N$, $\pi \in \Pi$, $i \in N$, and $h_{t-1} \in H^i$,*

$$\beta_i \langle u_i, h_{t-1} \rangle (\pi, u_{-i}) > 0 \quad \Longrightarrow \quad \beta_i \langle u_i, h_{t-1} \rangle \left(\cdot | \pi, u_{\pi(1,\ldots,t-1)} \right) = \mu_i \left(\cdot | u_{\pi(1,\ldots,t)} \right).$$

Property 2 in Lemma 1 suggests the types of conditions that would suffice to guarantee that, at any information set (on or off the equilibrium path) the agent will not "lose sight" of the true profile of types. First, in the deterministic-order case, the Cartesian support assumption turns out to guarantee that the realization of earlier-reporting agents' types—and their resulting reports—cannot lead an agent to rule out the actual realization of later-reporting agents' types. Second, in the random-order case, one also needs to rule out the possibility that an agent loses sight of the true profile because she rules out the true realized *sequence* of reports, π. For instance, in Example 2 Agent 1 fully believes (off the equilibrium path) that the first report came from Agent 4. Since Agents

[20] Theorems 1 and 2 generalize to all Perfect Bayesian equilibria satisfying the minimal consistency requirement in Lemma 1. In extensive-form games with perfect information in which agents' types are independent, this property is usually assumed as a basic consistency requirement of perfect Bayesian equilibria and is referred to as *beliefs being action-determined*, i.e., the marginal belief about the type of an agent can be updated only when this agent takes an action in the game, and this update exclusively depends on this agent's action (c.f., Osborne and Rubinstein, 1994, Def. 232.1). See Watson (2017) for a general definition of this minimal consistency condition with possibly correlated types.

2 and 3 are always high-demand types, Agent 1 believes the good will be rationed anyway and thus believes there to be no loss in exaggerating her report. However, if Agent 1 anticipated some chance of a future, truthful report from Agent 4, Agent 1 would strictly prefer not to exaggerate her claim. As we show, the requirements that the prior have *symmetric* Cartesian support *and* that the social choice function be weakly anonymous together guarantee that, even if the agent loses sight of the true sequence, she nevertheless envisions the possibility of another permutation and type profile that forces her to act as if she were truthful.

The following lemma, which applies to the deterministic-sequence case, states that in a consistent assessment, following any history, an agent must assign positive belief to any admissible "continuation profile," i.e. to any subprofile of preferences for the remaining agents that had positive prior probability under μ.[21]

Lemma 2. *Fix a prior $\mu \in \mathscr{M}_{\text{Cartesian}}$ exhibiting Cartesian support, a SCF f, and a deterministic distribution over sequences $\Lambda \in \Delta(\Pi)$, i.e. where $\Lambda(\pi) = 1$ for some $\pi \in \Pi$. Let assessment (σ, β) be consistent for $\langle \Gamma(\Lambda, f), \mathscr{U}^N, \mu \rangle$. Then for each $u \in \text{supp}(\mu)$, $t \in \{1, \dots, n\}$, and $h_{t-1} \in H_{t-1}$, there exists $v_{\pi(1,\dots,t-1)} \in \mathscr{U}^{\pi(1,\dots,t-1)}$ such that*

$$\beta_i \langle u_i, h_{t-1} \rangle (\pi, (v_{\pi(1,\dots,t-1)}, u_{\pi(t+1,\dots,n)})) > 0.$$

The next lemma extends the previous idea to the random-sequence case when the prior has symmetric Cartesian support. It states that in a consistent assessment, following any history, an agent must assign positive belief to *some reordering* of any admissible continuation profile, i.e. if it is possible for some set of later-reporting agents $\pi(t+1, \dots, n)$ to have a certain subprofile of preferences in u', then for any history h_{t-1} the tth reporting agent assigns positive probability to some set of remaining agents $\pi'(t+1, \dots, n)$ having (some ordering of) that subprofile.

Lemma 3. *Fix a prior $\mu \in \mathscr{M}_{\text{symm-Cartesian}}$, a SCF f, and $\Lambda \in \Delta(\Pi)$. Let (σ, β) be consistent for $\langle \Gamma(\Lambda, f), \mathscr{U}^N, \mu \rangle$. Then for each $u \in \text{supp}(\mu)$, $\pi \in \text{supp}(\Lambda)$, $t \in \{1, \dots, n\}$, and $h_{t-1} \in H_{t-1}$, there exist (i) $\pi' \in \Pi$ with $\pi'(t) = \pi(t)$ and (ii) $v \in \mathscr{U}^N$ with $[v_{\pi'(t+1,\dots,n)}] = [u_{\pi(t+1,\dots,n)}]$ such that*

$$\beta_{\pi(t)} \langle u_{\pi(t)}, h_{t-1} \rangle (\pi', v_{-\pi(t)}) > 0.$$

[21] The conclusion in the lemma cannot be strengthened to state that the agent necessarily places positive weight on the true profile. Consider for instance the agent who reports first. Conditional on an off-equilibrium report, the other agents' belief about this agent's type is unrestricted by consistency.

That is, if it is ex-ante possible for a set of "true" types, $[u_{\pi(t+1,\ldots,n)}]$, to be realized by the agents yet to act after round t, then agent $\pi(t)$ must anticipate with positive probability this same set of types to be realized by *some* ordering of *some* set of agents yet to act, $[v_{\pi'(t+1,\ldots,n)}]$.

5 Discussion and concluding remarks

We have demonstrated that the use of strategy-proof, non-bossy rules as sequential revelation mechanisms may lead to equilibria that yield non-truthful outcomes and that give agents strict disincentives to be truthful. This phenomenon occurs in Bayesian settings; in contrast it is ruled out in complete information settings such as those considered in previous literature. The non-truthful equilibria in our motivating examples are robust to forward induction arguments. Therefore the issue that we address is not simply an artifact of some construction of "unreasonable" equilibrium beliefs, but rather is a consequence of the structure of the *primitives* of the model: the prior beliefs. Our results state that the phenomena in our examples can be ruled across *all* sequential equilibria whenever prior beliefs satisfy our Cartesian Support (CS) conditions (along with a weak anonymity assumption when reporting sequences are random).

The degree to which our results have a positive or negative interpretation heavily depends on the planner's interpretation of the CS conditions.[22] In many applications, particularly when the agents are unfamiliar with each other, such conditions are innocuous.[23] One can imagine scenarios in which the CS conditions do not apply, however. For example, imagine that Agent 2 knows (with certainty) that Agents 1 and 3 have identical preferences (e.g. they are siblings or partners) but is unsure what those common preferences are. In this case, Agent 1's private information rules out certain preference relations for Agent 3 *with certainty*, so the CS condition breaks down. Similar stories can be constructed in which Agents 1 and 3 are certain to have opposed preferences (e.g. they are competitors), violating the CS condition.

If a planner views our CS conditions to be applicable,[24] then our results should be interpreted as a positive robustness check on strategy-proof, non-

[22] We thank referees for raising this point.

[23] In a finite model, where the set of all priors is a simplex, Definitions 1 and 2 are satisfied for almost every prior with respect to any measure over the set of all priors that is absolutely continuous with respect to the Lebesgue measure.

[24] The applicability of CS is determined by the planner's perception of the players' prior beliefs. Something that could aid this determination is historical data on the distribution of agents' preferences. For example in the school choice environment, there is a growing literature addressing

bossy rules. It is "safe" to use such rules as sequential revelation mechanisms in the sense that sequential equilibria are guaranteed to yield the desired outcome.

If instead players are sufficiently informed about each others' preferences that the CS conditions fail, then our work highlights a channel through which sequential revelation could lead to payoffs that differ from those prescribed by the underlying allocation rule. On the other hand, a reduction in agents' private information can make Bayesian implementation (with *indirect* mechanisms) easier to achieve. In an exchange economy setting, Postlewaite and Schmeidler (1986) show that (monotonic) SCF's are Bayesian implementable whenever beliefs satisfy a condition implying the following: for any realization of types for $N\setminus\{j\}$, agent j's type can be deduced with certainty. Since this condition has the opposite flavor of CS, this loosely suggests an agenda for future work: the structure of information could be used to guide the planner's decision whether to use sequential revelation mechanisms or simultaneous (possibly indirect) mechanisms.[25]

When the CS condition fails, one might wonder how "far" equilibrium payoffs might be from truthful ones. Within a general allocation model such as ours, it is not possible to provide any general answer to this question. The primary reason for this is the disconnection between our CS conditions (which pertain only to prior beliefs) and the agents' payoffs (which depend on the specification and structure of type- and outcome-spaces). Indeed in some settings, there may be *no* gap between equilibrium outcomes and truthful ones even when the CS condition fails. In settings where truthful reports are strictly dominant actions, for example, any sequential equilibrium outcome is a truthful one, regardless of whether beliefs have Cartesian Support.

In other settings, non-truthful sequential equilibria can be arbitrarily "bad" depending on the scale of payoffs. To see this, observe that the equilibria we construct in Example 1 and Example 2 would be preserved under any affine transformation of the agent's payoffs. Therefore any absolute or relative measure of welfare loss or inefficiency in terms of payoffs can be achieved with an appropriately rescaled version of these examples. Since our objective is to address *when* undesirable outcomes occur, rather than to describe *what* they are, we leave further consideration of the latter question to future work that is specialized to specific problems or domains.

---

preference estimation even when agents are not truth-telling or when the underlying SCF is not strategy-proof. See Fack et al. (2019) and the references therein.

[25] We are grateful to a referee for this insight.

Appendix

Sequential Equilibrium of Example 1

We provide the formal arguments of the claims given in Subsection 2.1, namely that (σ, β) is a sequential equilibrium with the properties described earlier.

Agent 1's beliefs under β are, trivially, the conditional measures $\beta_1\langle p_1\rangle = \mu(\cdot|p_1)$, i.e. a Bayesian update of Table 1 given p_1. Following any report from Agent 1 (i.e. any history h_2) and independent of her own peak p_2, Agent 2's beliefs, β_2, are described in Table 3. Denote the possible subprofiles (p_1, p_3, p_4) by $u_{-2} = (2,2,2)$, $v_{-2} = (2.5,0,0)$, and $w_{-2} = (3,0,0)$.

Agent 1's report (h_2)	Agent 2's belief
0	$\frac{1}{3}u_{-2}, \frac{1}{3}v_{-2}, \frac{1}{3}w_{-2}$
1	u_{-2}
2	u_{-2}
2.5	$\frac{1}{2}v_{-2}, \frac{1}{2}w_{-2}$
3	u_{-2}

Table 3: Beliefs for agent 2 are independent of her type. The table shows the distribution on $\mathscr{U}^{\{1,3,4\}}$ that agent 2 believes is true when she observes the respective histories of play.

Finally, since agents 3 and 4 always report their peaks truthfully and the Uniform rule is strategy-proof, it is straightforward to specify beliefs β_3, β_4 in a way that satisfies consistency and sequential rationality. We omit these details.

To show that σ is sequentially rational for Agent 1, observe Table 4 which shows the distribution of outcomes that Agent 1 predicts conditional on her report and her peak (which informs her about the truthful reports of Agents 3 and 4). Agent 1's strategy, σ_1, places positive probability only on actions that maximize her expected payoff conditional on her type. Of particular note is that, when $p_1 = 3$, agent 1's *unique* best response is to misreport her peak as 2.5. That is, Agent 1 strictly prefers to misreport her preferences in this equlibrium.

Similarly Table 5 shows the distribution of outcomes that agent 2 predicts, conditional on her report and on the report of Agent 1. (Recall that agent 2's own peak does not influence her beliefs and thus is omitted from the table.) Agent 2's strategy, σ_2, places positive probability only on actions that maximize her expected payoff conditional on the history of play.

To complete the argument that (β, σ) is a sequential equilibrium we observe that β_2 is easily seen to be the limit (as $\varepsilon \to 0$) of beliefs obtained by

Agent 1's report	Agent 1's peak		
	$p_1 = 2$	$p_1 = 2.5$	$p_1 = 3$
0	0	$\frac{1}{2}(2/3), \frac{1}{2}(1)$	$\frac{1}{2}(2/3), \frac{1}{2}(1)$
1	**1**	1	1
2	**1**	2	2
2.5	**1**	$\mathbf{\frac{1}{2}(2), \frac{1}{2}(2.5)}$	$\mathbf{\frac{1}{2}(2), \frac{1}{2}(2.5)}$
3	**1**	2	2

Table 4: For each combination of Agent 1's report and peak, the table provides the distribution of the amount of good received by agent 1 under σ. For each peak, the payoff-maximal outcomes are boldfaced. Observe that when $p_1 = 3$, the payoff-maximal outcome is obtained *only* by (mis)reporting a peak of 2.5.

Agent 2's report	Agent 1's report, h_1				
	0	1	2	2.5	3
0	$\frac{1}{3}(0), \frac{2}{3}(1)$	0	0	$\frac{1}{2}$	0
1	1	1	1	1	1
2	$\frac{1}{3}(\frac{4}{3}), \frac{2}{3}(2)$	1	1	2	1
2.5	$\frac{1}{3}(\frac{4}{3}), \frac{2}{3}(2.5)$	1	1	2	1
3	$\frac{1}{3}(\frac{4}{3}), \frac{2}{3}(3)$	1	1	2	1

Table 5: Agent 2's consumption under σ, conditional on the reports of Agents 1 and 2. The consumption may be a lottery whose outcome depends on the reports of Agents 3 and 4 (independent of agent 2's type). E.g. when Agents 1 and 2 both report 0, Agent 2 receives 0 with probability 1/3 and receives 1 otherwise.

Bayesian updating when Agent 1 plays the (full support) mixed strategy σ_1^ε defined in Table 6.

Sequential Equilibrium in Example 2

We provide the formal arguments of the claims given in Subsection 2.2, namely that (σ, β) is a sequential equilibrium with the properties described earlier.

Define strategy profile σ^ε as follows. First, whenever Agent 1, 2, or 3 is chosen to report first, that agent randomizes her report with the distribution in Table 7.

Whenever Agent 4 is chosen to report first, she randomizes according to the distribution in Table 8.

	Agent 1's Type		
Action	2	2.5	3
0	ε	ε	ε
1	ε	ε^2	ε^2
2	$1-4\varepsilon$	ε	ε
2.5	ε	$1-2\varepsilon-2\varepsilon^2$	$1-2\varepsilon-2\varepsilon^2$
3	ε	ε^2	ε^2

Table 6: Fully mixed strategies σ_1^ε, converging to σ_1, whose associated Bayesian beliefs define β_2. Each column provides a distribution over actions for the respective type.

	Peak of agent 1, 2, or 3		
Report	2.5	2.9	3
0	ε	ε	ε
2	ε	ε	ε
2.5	$1-3\varepsilon-\varepsilon^2$	ε	ε
2.9	ε	$1-3\varepsilon-\varepsilon^2$	$1-3\varepsilon-\varepsilon^2$
3	ε^2	ε^2	ε^2

Table 7: First-round strategies ($\sigma_i^\varepsilon(p_i;\emptyset)$, $1 \le i \le 3$) for Agents 1–3. Each column is a distribution over reports given the agent's peak.

Finally, when any agent $i \in N$ is *not* the first to report her peak, her strategy falls into one of two cases.

- First, if both $i \in \{1,2,3\}$ and all previous reports have been "3," the agent reports a peak of 3 with probability $1-4\varepsilon$ and reports any one of the other four admissible peaks with probability ϵ (independently of her true peak).
- Otherwise (if either $i = 4$ or at least one previous report was not "3") the agent reports her true peak with probability $1-4\varepsilon$ and reports any one of the other four admissible peaks with probability ϵ.

This completes the description of σ^ε, which has full support on the set of all reports. Let β^ε be the unique Bayesian belief based on σ^ε. Our example is the assessment (σ,β) which is the limit of $(\sigma^\varepsilon,\beta^\varepsilon)$ as $\varepsilon \to 0$.

We argue that if an agent $i \in \{1,2,3\}$ observes any history of reports with a first entry of 3, then β_i assigns probability one to the event that Agent 4 has

	Peak of agent 4		
Report	0	2	3
0	$1-3\varepsilon-\varepsilon^2$	ε	ε
2	ε	$1-4\varepsilon$	$1-4\varepsilon$
2.5	ε	ε	ε
2.9	ε	ε	ε
3	ε^2	ε	ε

Table 8: First-round strategies ($\sigma_4^\varepsilon(p_4;\emptyset)$) for Agent 4.

already reported her preferences. To see this, one can compute the probability of this event induced by σ^ε for any ε, using Bayes' rule. Informally, however, first consider the history (3), i.e. where i is reporting second. The probability that an agent different from Agent 4 is chosen first and reports 3 is of order ε^2, while the probability that Agent 4 is chosen first and reports 3 is of order ε, yielding the claim. Next consider any history $(3,p)$ where $p \neq 3$. The probability of this pair of reports coming from two agents *other than* Agent 4 is of order ε^3; The probability of this pair of reports coming from a pair *containing* Agent 4 is of order ε^2, proving the claim. Next consider history $(3,3)$. The probability of this pair of reports coming from two agents *other than* Agent 4 is of order ε^2; The probability of this pair of reports coming from a pair *containing* Agent 4 is of order ε, proving the claim. Finally, the claim is trivial for such a history of length 3.

To show that (σ,β) is a sequential equilibrium, first observe that it is obviously consistent, being defined as the limit of full-support assessments. To show sequential rationality, first consider any history under σ in which at least one agent has already reported some peak other than 3. In this case, regardless of the realization of peaks, σ prescribes truthful behavior for all subsequent agents regardless of what future reports are made. Thus no subsequent agent's report can affect which reports will be made by the agents who follow. Hence by the strategy-proofness of the Uniform rule no agent can achieve a higher payoff than the one obtained by making a truthful report, proving sequential rationality of σ following such histories.

Second consider an agent $i \in N$ who must act following some history $h \in \{(3),(3,3),(3,3,3)\}$ in which all previous reports are 3. Recall that if $i \neq 4$, β_i assigns probability 1 to the event that Agent 4 already reported, i.e. will not act in a future round. Of course if $i = 4$, this agent also knows that she, herself, will not act in a future round. Thus all future reports following i's (if any) will be 2.5

or higher. This leads to two possibilities. If i reports a peak of 0, the Uniform rule assigns zero units to agent i. If i reports any other peak (2 or higher), the Uniform rule assigns two units to agent i. Thus at each of these histories, a report of 3 is sequentially rational when $i \in \{1,2,3\}$ and truthful reporting is sequentially rational for $i = 4$, i.e. σ is sequentially rational for agents acting in round 2 or later.

Finally consider the anticipated distribution of consumption of an agent who is selected to report her peak first. The simplest case is Agent 4, since under σ, Agents 1–3 always report peaks of 2.5 or higher. Therefore if Agent 4 reports a peak of 2 or more, she receives 2 units with certainty. If she reports a peak of 0, her consumption is at most 0.5 units. When Agent 4's peak is 0, a report of 0 maximizes her expected payoff; otherwise a report of 2 does. Therefore σ_4 is sequentially rational.

For an agent $i \in \{1,2,3\}$, it is readily checked that a reported peak of 0 is never optimal given σ_{-i}: reporting a peak of 0 yields at most 1.5 units of the good, while reporting 2 guarantees her 2 units of the good, which is always preferable since her peak is above 2. By reporting 2.5, i receives either 2.5 units (if 4's peak is 0) or 2 units (otherwise). Calculations for reports of 2.9 and 3 are more tedious, but can be seen intuitively by considering the fact that α is close to 1. Given this assumption, the approximate distribution over i's consumption following any first round report is provided in Table 9. For instance, a first round report of 0 triggers truthful behavior by the remaining agents, yielding Agent i 1.5 units of the good with significant probability (at least α^3). Similarly, a report of 2.9 yields 2.9 units of the good with probability α^3. If i reports 3, however, then the outcome depends on which of the next three reports will be made by (truthful) Agent 4, since agents $\{1,2,3\} \setminus \{i\}$ are truthful only if they report *after* Agent 4. It can be checked that when the other agents' peaks are $(2.5, 2.5, 0)$ (probability α^3), Agent i is equally likely to receive 8/3, 11/4, or 3 units of the good.

Therefore it should be clear that, when i's peak is 2.5 or 2.9, i should report truthfully. However when i's peak is 3 (and due to piecewise linear utility in outcomes), she strictly prefers misreporting her peak to be 2.9, as prescribed by σ when α is reasonably close to 1.[26]

Observe in particular that, when chosen to act first, agents $i \in \{1,2,3\}$ have a *strict* disincentive to truthfully reveal a peak of 3. Doing so would lead the remaining of those agents to (falsely) believe that Agent 4 made that report, and thus (falsely) allow them to believe that over-reporting their peaks would have

[26]E.g. with some rounding, an agent with peak at 3 is indifferent between outcome distributions $\frac{1}{3}\delta_{8/3} + \frac{1}{3}\delta_{11/4} + \frac{1}{3}\delta_3$ and $\delta_{2.805}$.

First round report	Distribution of consumption for agent i	
	$i \in \{1,2,3\}$	$i = 4$
0	$\approx \delta_{1.5}$	$\approx \delta_{0.5}$
2	δ_2	δ_2
2.5	$(1-\alpha)\delta_2 + \alpha\delta_{2.5}$	δ_2
2.9	$\approx \delta_{2.9}$	δ_2
3	$\approx \frac{1}{3}\delta_{8/3} + \frac{1}{3}\delta_{11/4} + \frac{1}{3}\delta_3$	δ_2

Table 9: Under σ, any first round report by Agent i yields a distribution over consumption as given in the table. The notation δ_x represents a distribution with probability 1 on receiving x units of the good. The approximation ($\approx$) improves as α becomes close to one.

no repercussions. Finally, observe that Agent 4 indeed could be considered the "most likely" agent to make such an out-of-equilibrium report, since, given σ, that is the unique agent who is merely indifferent between using the first round report prescribed by σ and reporting 3.

Proofs

Proof of Proposition 1. Let f, Λ, and μ be as in the statement of the proposition. To construct β, fix an arbitrary mixed strategy profile with full support, σ', and define $\sigma^\epsilon \equiv (1-\epsilon)\tau + \epsilon\sigma'$ for $\epsilon > 0$. Let β^ϵ be the belief system defined by Bayesian updating σ^ϵ, and let β be the (well-defined) limit of β^ϵ as $\epsilon \to 0$. Clearly $\sigma^\epsilon \to \tau$, and (τ, β) is consistent.

Sequential rationality follows immediately from the strategy-proofness of f since, for any $i \in N$, τ_{-i} prescribes truthful behavior from the other agents *regardless* of the history of play. Thus, when Agent i reports $v_i \in \mathscr{U}$, the outcome ends up being $f(u_{-i}, v_i)$ for any realization of true preferences $u \in \mathscr{U}^N$ and for any realization of sequence π according to Λ. By strategy-proofness the truthful report $v_i = u_i$ maximizes i's payoff ex-post for any such realization, so such a report also maximizes i's expected payoff at any information set. Hence (τ, β) is a sequential equilibrium. □

Proof of Lemma 1. Fix f, Λ, and (σ, β) as in the statement of the lemma. Since (σ, β) is consistent, fix a sequence of assessments $\{(\sigma^k, \beta^k)\}_{k \in \mathbb{N}}$ such that for each $k \in \mathbb{N}$, (i) σ^k has full support, (ii) β^k is obtained by Bayesian updating (w.r.t. Λ, μ, σ^k), and (iii) as $k \to \infty$, $(\sigma^k, \beta^k) \to (\sigma, \beta)$.

Statement 1. Let $u \in \text{supp}(\mu)$, $\pi \in \Pi$, $i \in N$, $h_{t-1} \in H^i$, and $u'_{-i} \in \mathscr{U}^{N \setminus \{i\}}$.

Suppose that either $\pi(t) \neq i$, $\Lambda(\pi) = 0$, or $\mu_i(u'_{-i}, u_i) = 0$. Since for each $k \in \mathbb{N}$ each β^k is obtained by Bayesian updating a full-support strategy profile σ^k, we must have $\beta_i^k\langle u_i, h_{t-1}\rangle(\pi, u'_{-i}) = 0$, obtaining a limit of $\beta_i\langle u_i, h_{t-1}\rangle(\pi, u'_{-i}) = 0$.

Statement 2. Let $u \in \mathscr{U}^N$, $\pi \in \Pi$, $i \in N$, and $h_{t-1} \in H^i$ be such that $\beta_i\langle u_i, h_{t-1}\rangle(\pi, u_{-i}) > 0$. By Statement 1, $\mu(u) > 0$, so $\mu_i(\cdot|u_{\pi(1,\dots,t)})$ is well-defined. Since $\beta_i^k\langle u_i, h_{t-1}\rangle(\pi, u_{-i}) \to \beta_i\langle u_i, h_{t-1}\rangle(\pi, u_{-i})$, there is $K \in \mathbb{N}$ such that for all $k \geq K$, $\beta_i^k\langle u_i, h_{t-1}\rangle(\pi, u_{-i}) > 0$. Thus, for each $k \geq K$, $\beta_i^k\langle u_i, h_{t-1}\rangle(\cdot|\pi, u_{\pi(1,\dots,t-1)})$ is well-defined, and specifically

$$\beta_i^k\langle u_i, h_{t-1}\rangle(u_{\pi(t+1,\dots,n)}|\pi, u_{\pi(1,\dots,t-1)}) = \frac{\beta_i^k\langle u_i, h_{t-1}\rangle(\pi, u_{-i})}{\sum\limits_{v_{\pi(t+1,\dots,n)} \in \mathscr{U}^{\pi(t+1,\dots,n)}} \beta_i^k\langle u_i, h_{t-1}\rangle(\pi, u_{\pi(1,\dots,t-1)}, v_{\pi(t+1,\dots,n)})}$$

which, as $k \to \infty$, converges to

$$\frac{\beta_i\langle u_i, h_{t-1}\rangle(\pi, u_{-i})}{\sum\limits_{v_{\pi(t+1,\dots,n)} \in \mathscr{U}^{\pi(t+1,\dots,n)}} \beta_i\langle u_i, h_{t-1}\rangle(\pi, u_{\pi(1,\dots,t-1)}, v_{\pi(t+1,\dots,n)})}$$

which equals $\beta_i\langle u_i, h_{t-1}\rangle(u_{\pi(t+1,\dots,n)}|\pi, u_{\pi(1,\dots,t-1)})$. Since $\mu(u) > 0$, each $\beta_i^k\langle u_i, h_{t-1}\rangle(u_{\pi(t+1,\dots,n)}|\pi, u_{\pi(1,\dots,t-1)})$ is equal to:

$$\frac{\Lambda(\pi)\mu(u)\sigma^k(h_t|h_0, \pi, u)}{\sum\limits_{v_{\pi(t+1,\dots,n)} \in \mathscr{U}^{\pi(t+1,\dots,n)}} \Lambda(\pi)\mu(u_{\pi(1,\dots,t)}, v_{\pi(t+1,\dots,n)})\sigma^k(h_t|h_0, \pi, u_{\pi(1,\dots,t)}, v_{\pi(t+1,\dots,n)})}.$$

For each $v_{\pi(t+1,\dots,n)} \in \mathscr{U}^{\pi(t+1,\dots,n)}$, we have $\sigma^k(h_t|\pi, u_{\pi(1,\dots,t)}, v_{\pi(t+1,\dots,n)}) = \sigma^k(h_t|\pi, u)$ (i.e. the probability of seeing h_t is only a function of the first t agent's types), so the previous equation reduces to

$$\beta_i^k\langle u_i, h_{t-1}\rangle(u_{\pi(t+1,\dots,n)}|\pi, u_{\pi(1,\dots,t-1)}) = \mu(u_{\pi(t+1,\dots,n)}|u_{\pi(1,\dots,t)}).$$

Thus,

$$\beta_i\langle u_i, h_{t-1}\rangle(u_{\pi(t+1,\dots,n)}|\pi, u_{\pi(1,\dots,t-1)}) = \mu(u_{\pi(t+1,\dots,n)}|u_{\pi(1,\dots,t)}). \quad \square$$

Proofs of Lemma 2 and Lemma 3. Suppose $\mu \in \mathscr{M}_{\text{Cartesian}}$, and let u, π, t, and h_{t-1} be as in the statement of Lemma 3. Denote $i \equiv \pi(t)$ and choose any (π', v_{-i}) in the support of $\beta_i\langle u_i, h_{t-1}\rangle$. By statement 1 in Lemma 1, $\Lambda(\pi') > 0$, $\mu(u_i, v_{-i}) > 0$, and $\pi'(t) = \pi(t) = i$. By statement 2 in Lemma 1,

$$\beta_i\langle u_i, h_{t-1}\rangle(\cdot|\pi', v_{\pi'(1,\dots,t-1)}) = \mu(\cdot|v_{\pi'(1,\dots,t-1)}, u_i). \tag{1}$$

We consider two (non-exhaustive) cases. The first case is implied when Λ is deterministic, proving Lemma 2. The second adds the symmetric Cartesian assumption, proving Lemma 3.

Case 1: Λ is deterministic. By Lemma 1, $\pi' = \pi$. Since $\mu(u) > 0$ and $\mu(u_i, v_{-i}) \equiv \mu(u_{\pi(t)}, v_{N\setminus\pi(t)}) > 0$, Cartesian support implies $\mu(u_{\pi(t,t+1)}, v_{N\setminus\pi(t,t+1)}) > 0$. Repeating the argument implies $\mu(u_{\pi(t,t+1,t+2)}, v_{N\setminus\pi(t,t+1,t+2)}) > 0$, etc., concluding with

$$\mu(u_{\pi(t,\ldots,n)}, v_{N\setminus\pi(t,\ldots,n)}) > 0.$$

Therefore, with (1) we have

$$\beta_i\langle u_i, h_{t-1}\rangle(u_{\pi(t+1,\ldots,n)}|\pi, v_{\pi(1,\ldots,t-1)}) = \mu(u_{\pi(t+1,\ldots,n)}|v_{\pi(1,\ldots,t-1)}, u_i) > 0.$$

Since $(\pi', v_{-i}) \equiv (\pi, v_{-i})$ is in the support of $\beta_i\langle u_i, h_{t-1}\rangle$,

$$\beta_i\langle u_i, h_{t-1}\rangle(\pi, (v_{\pi(1,\ldots,t-1)}, u_{\pi(t+1,\ldots,n)})) > 0.$$

Case 2: $\mu \in \mathscr{M}_{\text{symm–Cartesian}}$. Let $v'_{\pi'(t+1,\ldots,n)} \in \mathscr{U}^{\pi'(t+1,\ldots,n)}$ be such that $[v'_{\pi'(t+1,\ldots,n)}] = [u_{\pi(t+1,\ldots,n)}]$. Since $\mu(u) > 0$ and $\mu(u_i, v_{-i}) > 0$, symmetric Cartesian support implies

$$\mu(u_i, v_{\pi'(1,\ldots,t-1)}, v'_{\pi'(t+1,\ldots,n)}) > 0.$$

Thus by (1), $\beta_i\langle u_i, h_{t-1}\rangle(v'_{\pi'(t+1,\ldots,n)}|\pi', v_{\pi'(1,\ldots,t-1)}) > 0$. Since (π', v_{-i}) is in the support of $\beta_i\langle u_i, h_{t-1}\rangle$,

$$\beta_i\langle u_i, h_{t-1}\rangle(\pi', v_{\pi'(1,\ldots,t-1)}, v'_{\pi'(t+1,\ldots,n)}) > 0.$$

Thus with $\pi', (v_{\pi'(1,\ldots,t-1)}, v'_{\pi'(t+1,\ldots,n)})$ satisfies part (ii) of Lemma 3. □

Proof of Theorem 1. Fix notation as in the statement of the theorem: let f be a strategy-proof, non-bossy SCF, and let $\Lambda \in \Delta(\Pi)$, $\mu \in \mathscr{M}_{\text{Cartesian}}$, and $(\sigma, \beta) \in \text{SE}\langle\Gamma(\Lambda, f), \mathscr{U}^N, \mu\rangle$.

Case 1: Λ is deterministic. Without loss of generality, let $\Lambda(\pi) = 1$ where $\pi(i) \equiv i$. By Lemma 1, any player's beliefs $\beta_i\langle\rangle()$ must assign probability 1 to the sequence π. Obesrve that, throughout Case 1, $f(h_{t-1}, v_{\pi(t,\ldots,n)})$ is the outcome anticipated by agent $\pi(t)$ after history h_{t-1}, when the agent anticipates the remaining agents to make reports $v_{\pi(t,\ldots,n)}$.[27]

We prove that for any feasible history (even off the equilibrium path), equilibrium continuation strategies are welfare-equivalent to truthful continuation

[27] In contrast, when π is uncertain, if agent $\pi(t)$ sees history h_{t-1} and anticipates future reports $v_{\pi(t,\ldots,n)}$, the agent still cannot necessarily anticipate a particular outcome of f since the agent does not know which agents made which reports in $(h_{t-1}, v_{\pi(t,\ldots,n)})$.

strategies. That is, for any $t \in \{1, \ldots, n\}$, $u \in \text{supp}(\mu)$, $h_{t-1} \in H^{\pi(t)}$, and (equilibrium continuation) $v_{\pi(t,\ldots,n)} \in \mathscr{U}^{\pi(t,\ldots,n)}$ with $\sigma(v_{\pi(t,\ldots,n)} | h_{t-1}, \pi, u_{\pi(t,\ldots,n)}) > 0$, we have

$$u(f(h_{t-1}, v_{\pi(t,\ldots,n)})) = u(f(h_{t-1}, u_{\pi(t,\ldots,n)})). \tag{2}$$

Moreover if f is also non-bossy* in outcomes we have $f(h_{t-1}, v_{\pi(t,\ldots,n)}) = f(h_{t-1}, u_{\pi(t,\ldots,n)})$. Applying (2) to the case $t = 1$ yields the theorem. The proof of (2) is by backward induction on t.

Initial step $t = n$. Fix $u \in \text{supp}(\mu)$, $h_{n-1} \in H^{\pi(n)}$, and an equilibrium report $v_{\pi(n)} \in \mathscr{U}$ in the support of $\sigma_{\pi(n)}(u_{\pi(n)}, h_{n-1})$. There is no strategic uncertainty for player $\pi(n) = n$, so sequential rationality of (σ, β) immediately implies

$$u_{\pi(n)}(f(h_{n-1}, v_{\pi(n)})) \geq u_{\pi(n)}(f(h_{n-1}, u_{\pi(n)})) \tag{3}$$

while strategy-proofness of f implies the reverse inequality,

$$u_{\pi(n)}(f(h_{n-1}, v_{\pi(n)})) \leq u_{\pi(n)}(f(h_{n-1}, u_{\pi(n)})). \tag{4}$$

yielding $u_{\pi(n)}(f(h_{n-1}, v_{\pi(n)})) = u_{\pi(n)}(f(h_{n-1}, u_{\pi(n)}))$. Since f is non-bossy,

$$u(f(h_{n-1}, v_{\pi(n)})) = u(f(h_{n-1}, u_{\pi(n)})). \tag{5}$$

Moreover if f is non-bossy* in outcomes we have $f(h_{n-1}, v_{\pi(n)}) = f(h_{n-1}, u_{\pi(n)})$.

Inductive step $t < n$. To prove the inductive step, fix $t \in \{0, \ldots, n-1\}$ and assume the induction hypothesis: for any $u \in \text{supp}(\mu)$, $h_t \in H^{\pi(t+1)}$, and (equilibrium continuation) $v_{\pi(t+1,\ldots,n)} \in \mathscr{U}^{\pi(t+1,\ldots,n)}$ with $\sigma(v_{\pi(t+1,\ldots,n)} | h_t, \pi, u_{\pi(t+1,\ldots,n)}) > 0$, we have

$$u(f(h_t, v_{\pi(t+1,\ldots,n)})) = u(f(h_t, u_{\pi(t+1,\ldots,n)})) \tag{6}$$

(and that, moreover, if f is non-bossy* in outcomes, $f(h_t, h_{\pi(t+1,\ldots,n)}) = f(h_t, u_{\pi(t+1,\ldots,n)})$).

Fix $u \in \text{supp}(\mu)$, $h_{t-1} \in H^{\pi(t)} \equiv H^t$, and an equilibrium report $v_t \in \mathscr{U}$ with $\sigma_t(v_t | u_t, h_{t-1}) > 0$. Sequential rationality of (σ, β) implies that the expected payoff from reporting v_t, namely

$$\sum_{v_{-t}} \sum_{v'_{\pi(t+1,\ldots,n)}} u_t(f(h_{t-1}, v_t, v'_{\pi(t+1,\ldots,n)}))\sigma(v'_{\pi(t+1,\ldots,n)} | (h_{t-1}, v_t), v_{-t})\beta_t\langle u_t, h_{t-1}\rangle(\pi, v_{-t}), \tag{7}$$

is greater than or equal to the expected payoff from truth-telling, namely

$$\sum_{v_{-t}} \sum_{v'_{\pi(t+1,\ldots,n)}} u_t(f(h_{t-1}, u_t, v'_{\pi(t+1,\ldots,n)}))\sigma(v'_{\pi(t+1,\ldots,n)} | (h_{t-1}, u_t), v_{-t})\beta_t\langle u_t, h_{t-1}\rangle(\pi, v_{-t}) \tag{8}$$

Now consider any v_{-t} such that $\beta_t\langle u_t, h_{t-1}\rangle(\pi, v_{-t}) > 0$. By Lemma 1 we have $(v_{-t}, u_t) \in \text{supp}(\mu)$, so the induction hypothesis (6) applies: for any report v'_t and any $v'_{\pi(t+1,\ldots,n)}$ in the support of $\sigma(\cdot|(h_{t-1}, v'_t), \pi, v_{\pi(t+1,\ldots,n)})$,

$$u(f(h_{t-1}, v'_t, v'_{\pi(t+1,\ldots,n)})) = u(f(h_{t-1}, v'_t, v_{\pi(t+1,\ldots,n)})).$$

Thus, expression (7) is equal to

$$\sum_{v_{-t}} u_t(f(h_{t-1}, v_t, v_{\pi(t+1,\ldots,n)}))\beta_t\langle u_t, h_{t-1}\rangle(\pi, v_{-t}), \tag{9}$$

and expression (8) is equal to

$$\sum_{v_{-t}} u_t(f(h_{t-1}, u_t, v_{\pi(t+1,\ldots,n)}))\beta_t\langle u_t, h_{t-1}\rangle(\pi, v_{-t}). \tag{10}$$

While summation (9) is greater than or equal to summation (10), the strategy-proofness of f implies a point-wise inequality in the reverse direction: for any arbitrary $v_{\pi(t+1,\ldots,n)} \in \mathscr{U}^{\pi(t+1,\ldots,n)}$,

$$u_t(f(h_{t-1}, v_t, v_{\pi(t+1,\ldots,n)})) \le u_t(f(h_{t-1}, u_t, v_{\pi(t+1,\ldots,n)})). \tag{11}$$

Thus, not only is summation (9) equal to summation (10), but it must be that for each v_{-t} satisfying $\beta_t\langle u_t, h_{t-1}\rangle(\pi, v_{-t}) > 0$, we have

$$u_t(f(h_{t-1}, v_t, v_{\pi(t+1,\ldots,n)})) = u_t(f(h_{t-1}, u_t, v_{\pi(t+1,\ldots,n)})). \tag{12}$$

In words, after any history and for any subprofile of types *that agent t anticipates for the future agents* ($\beta_t > 0$), any report by agent t leads to "as-if-truthful" behavior from the future agents. What remains to be shown is that agent t in fact *does* anticipate any admissible subprofile (or when Λ is random, anticipates subprofiles that are "strategically equivalent" to admissible subprofiles and sequences of future agents). This would rule out the phenomenon in Example 1, where an agent (off the equilibrium path) rules out what still could be the actual types of agents yet to act. That is, in the deterministic case we need to show that (12) holds for any v_{-t} such that $(u_t, v_{-t}) \in \text{supp}(\mu)$.

Therefore reconsider u_{-t} which was assumed to satisfy $(u_t, u_{-t}) \in \text{supp}(\mu)$. By Lemma 2 there exists $v'_{\pi(1,\ldots,t-1)}$ such that $\beta_t\langle u_t, h_{t-1}\rangle(\pi, (v'_{\pi(1,\ldots,t-1)}, u_{\pi(t+1,\ldots,n)})) > 0$. Thus we can invoke (12) with respect to the subprofile $(v'_{\pi(1,\ldots,t-1)}, u_{\pi(t+1,\ldots,n)})$ to conclude

$$u_t(f(h_{t-1}, v_t, u_{\pi(t+1,\ldots,n)})) = u_t(f(h_{t-1}, u_t, u_{\pi(t+1,\ldots,n)})). \tag{13}$$

Since f is non-bossy,

$$u(f(h_{t-1}, v_t, u_{\pi(t+1,\dots,n)})) = u(f(h_{t-1}, u_t, u_{\pi(t+1,\dots,n)})). \tag{14}$$

Moreover if f is also non-bossy* in outcomes we have

$$f(h_{t-1}, v_t, u_{\pi(t+1,\dots,n)}) = f(h_{t-1}, u_t, u_{\pi(t+1,\dots,n)}). \tag{15}$$

Finally consider any $v'_{\pi(t+1,\dots,n)}$ in the support of $\sigma(\cdot|(h_{t-1}, v_t), \pi, u_{\pi(t+1,\dots,n)})$. By the induction hypothesis (6) with respect to history (h_{t-1}, v_t),

$$u(f(h_{t-1}, v_t, v'_{\pi(t+1,\dots,n)})) = u(f(h_{t-1}, v_t, u_{\pi(t+1,\dots,n)}))$$

and if f is non-bossy* in outcomes,

$$f(h_{t-1}, v_t, v'_{\pi(t+1,\dots,n)}) = f(h_{t-1}, v_t, u_{\pi(t+1,\dots,n)}).$$

Since $(v_t, v'_{\pi(t+1,\dots,n)})$ can be chosen to be arbitrary equilibrium reports in the support of $\sigma(\cdot|h_{t-1}, \pi, u_{\pi(t,\dots,n)}) > 0$, these equalities combined with (14) and (15) yield (2) and prove Case 1.

Case 2: $\mu \in \mathscr{M}_{\text{symm-Cartesian}}$ and f is weakly anonymous. The idea of the proof is similar to Case 1, but the arguments and notation need to account for the fact that each acting agent remains uncertain about the identity of agents who have already reported. In order to evaluate an agent's interim expected payoff, this requires an additional piece of notation. Given a permutation π, an "anonymous" history h_t, and a "non-anonymous" sequence of reports $v_{\pi(t+1,\dots,n)}$, let

$$f(h_t, v_{\pi(t+1,\dots,n)}|\pi)$$

denote the outcome of f when (i) agents $\pi(1), \dots, \pi(t)$ report the types listed in h_t, and (ii) agents $\pi(t+1), \dots, \pi(n)$ report types $v_{\pi(t+1,\dots,n)}$.

As in Case 1, we begin by showing that, following any feasible history, equilibrium continuation strategies are welfare-equivalent to truthful continuation strategies, regardless of which agents have yet to act. That is, for any $t \in \{1, \dots, n\}$, $\pi \in \text{supp}(\Lambda)$, $u \in \text{supp}(\mu)$, $h_{t-1} \in H^{\pi(t)}$, and (equilibrium continuation) $v_{\pi(t,\dots,n)} \in \mathscr{U}^{\pi(t,\dots,n)}$ with $\sigma(v_{\pi(t,\dots,n)}|h_{t-1}, \pi, u_{\pi(t,\dots,n)}) > 0$, we have

$$u(f(h_{t-1}, v_{\pi(t,\dots,n)}|\pi)) = u(f(h_{t-1}, u_{\pi(t,\dots,n)}|\pi)). \tag{2'}$$

Moreover if f is also non-bossy* in outcomes we have $f(h_{t-1}, v_{\pi(t,\dots,n)}|\pi) = f(h_{t-1}, u_{\pi(t,\dots,n)}|\pi)$.

Initial step $t = n$. Fix $\pi \in \text{supp}(\Lambda)$, $u \in \text{supp}(\mu)$, $h_{n-1} \in H^{\pi(n)}$, and an equilibrium report $v_{\pi(n)} \in \mathscr{U}$ in the support of $\sigma_{\pi(n)}(u_{\pi(n)}, h_{n-1})$. Since f is weakly anonymous, there is no strategic uncertainty for player $\pi(n)$: player $\pi(n)$'s payoff depends only on her own report and the *unordered* list of reports $[h_{n-1}]$. Formally, weak anonymity implies that $u_{\pi(n)}(f(h_{n-1}, v_{\pi(n)}|\pi')$ is independent of π' (subject to $\pi'(n) = \pi(n)$), so her beliefs over sequences become irrelevant.

Therefore sequential rationality implies

$$u_{\pi(n)}(f(h_{n-1}, v_{\pi(n)}|\pi)) \geq u_{\pi(n)}(f(h_{n-1}, u_{\pi(n)}|\pi)) \tag{3'}$$

while strategy-proofness of f implies the reverse inequality,

$$u_{\pi(n)}(f(h_{n-1}, v_{\pi(n)}|\pi)) \leq u_{\pi(n)}(f(h_{n-1}, u_{\pi(n)}|\pi)). \tag{4'}$$

yielding $u_{\pi(n)}(f(h_{n-1}, v_{\pi(n)}|\pi)) = u_{\pi(n)}(f(h_{n-1}, u_{\pi(n)}|\pi))$. Since f is non-bossy,

$$u(f(h_{n-1}, v_{\pi(n)}|\pi)) = u(f(h_{n-1}, u_{\pi(n)}|\pi)). \tag{5'}$$

Moreover if f is non-bossy* in outcomes we have $f(h_{n-1}, v_{\pi(n)}|\pi) = f(h_{n-1}, u_{\pi(n)}|\pi)$.

Inductive step $t < n$. To prove the inductive step, fix $t \in \{0, \ldots, n-1\}$ and assume the induction hypothesis: for any $\pi \in \text{supp}(\Lambda)$, $u \in \text{supp}(\mu)$, $h_t \in H^{\pi(t+1)}$, and (equilibrium continuation) $v_{\pi(t+1,\ldots,n)} \in \mathscr{U}^{\pi(t+1,\ldots,n)}$ with $\sigma(v_{\pi(t+1,\ldots,n)}|h_t, \pi, u_{\pi(t+1,\ldots,n)}) > 0$, we have

$$u(f(h_t, v_{\pi(t+1,\ldots,n)}|\pi)) = u(f(h_t, u_{\pi(t+1,\ldots,n)}|\pi)) \tag{6'}$$

(and that, moreover, if f is non-bossy* in outcomes, $f(h_t, h_{\pi(t+1,\ldots,n)}|\pi) = f(h_t, u_{\pi(t+1,\ldots,n)}|\pi)$).

Fix an agent $j \in N$ who could feasibly act in round t, i.e. for whom $\tilde{\pi}(t) = j$ for some $\tilde{\pi} \in \text{supp}(\Lambda)$. Fix $u \in \text{supp}(\mu)$, $h_{t-1} \in H^j$, and (equilibrium) report $v_j \in \mathscr{U}$ with $\sigma_j(v_j|u_j, h_{t-1}) > 0$. By sequential rationality, the expected payoff from reporting v_j, namely

$$\sum_{v_{-t},\pi} \sum_{v'_{\pi(t+1,\ldots,n)}} u_j(f(h_{t-1}, v_j, v'_{\pi(t+1,\ldots,n)}|\pi))\sigma(v'_{\pi(t+1,\ldots,n)}|(h_{t-1}, v_j), v_{-t})\beta_j\langle u_j, h_{t-1}\rangle(\pi, v_{-t}), \tag{7'}$$

is greater than or equal to the expected payoff from truth-telling, namely

$$\sum_{v_{-t},\pi} \sum_{v'_{\pi(t+1,\ldots,n)}} u_j(f(h_{t-1}, u_j, v'_{\pi(t+1,\ldots,n)}|\pi))\sigma(v'_{\pi(t+1,\ldots,n)}|(h_{t-1}, u_j), v_{-t})\beta_j\langle u_j, h_{t-1}\rangle(\pi, v_{-t}) \tag{8'}$$

Now consider any $(\tilde{\pi}, \tilde{v}_{-t})$ such that $\beta_t\langle u_t, h_{t-1}\rangle(\tilde{\pi}, \tilde{v}_{-t}) > 0$. By Lemma 1 we have $\tilde{\pi} \in \text{supp}(\Lambda)$ and $(\tilde{v}_{-t}, u_t) \in \text{supp}(\mu)$, so the induction hypothesis (6′) applies: for any report v'_j and any $v'_{\tilde{\pi}(t+1,\dots,n)}$ in the support of $\sigma(\cdot|(h_{t-1}, v'_j), \tilde{\pi}, \tilde{v}_{\tilde{\pi}(t+1,\dots,n)})$,

$$u(f(h_{t-1}, v'_j, v'_{\tilde{\pi}(t+1,\dots,n)}|\tilde{\pi})) = u(f(h_{t-1}, v'_j, \tilde{v}_{\tilde{\pi}(t+1,\dots,n)}|\tilde{\pi})).$$

Thus, expression (7′) is equal to

$$\sum_{v_{-t},\pi} u_j(f(h_{t-1}, v_j, v_{\pi(t+1,\dots,n)}|\pi))\beta_j\langle u_j, h_{t-1}\rangle(\pi, v_{-t}), \tag{9′}$$

and expression (8′) is equal to

$$\sum_{v_{-t},\pi} u_j(f(h_{t-1}, u_j, v_{\pi(t+1,\dots,n)}|\pi))\beta_j\langle u_j, h_{t-1}\rangle(\pi, v_{-t}). \tag{10′}$$

While (9′) is greater than or equal to (10′), the strategy-proofness of f implies a point-wise inequality in the reverse direction: for arbitrary $\tilde{\pi}$ and $\tilde{v}_{\tilde{\pi}(t+1,\dots,n)} \in \mathcal{U}^{\tilde{\pi}(t+1,\dots,n)}$,

$$u_j(f(h_{t-1}, v_j, \tilde{v}_{\tilde{\pi}(t+1,\dots,n)}|\tilde{\pi})) \le u_j(f(h_{t-1}, u_j, \tilde{v}_{\tilde{\pi}(t+1,\dots,n)}|\tilde{\pi})). \tag{11′}$$

Thus, not only is summation (9′) equal to summation (10′), but it must be that for each $(\tilde{\pi}, \tilde{v}_{-j})$ satisfying $\beta_j\langle u_j, h_{t-1}\rangle(\tilde{\pi}, \tilde{v}_{-j}) > 0$, we have

$$u_j(f(h_{t-1}, v_j, \tilde{v}_{\pi(t+1,\dots,n)}|\tilde{\pi})) = u_j(f(h_{t-1}, u_j, \tilde{v}_{\pi(t+1,\dots,n)}|\tilde{\pi})). \tag{12′}$$

Therefore reconsider u_{-j} which was assumed to satisfy $(u_j, u_{-j}) \in \text{supp}(\mu)$, and consider an arbitrary $\pi \in \text{supp}(\Lambda)$ such that $\pi(t) = j$. By Lemma 3 there exist π' and v'_{-j}, with $\pi'(t) = j$ and $[v'_{\pi'(t+1,\dots,n)}] = [u_{\pi(t+1,\dots,n)}]$, such that $\beta_j\langle u_j, h_{t-1}\rangle(\pi', v'_{-j}) > 0$. Thus we can invoke (12′) with respect to π' and v'_{-j} to conclude

$$u_j(f(h_{t-1}, v_j, v'_{\pi'(t+1,\dots,n)}|\pi')) = u_j(f(h_{t-1}, u_j, v'_{\pi'(t+1,\dots,n)}|\pi')).$$

Since f is weakly anonymous and $[v'_{\pi'(t+1,\dots,n)}] = [u_{\pi(t+1,\dots,n)}]$ this implies

$$u_j(f(h_{t-1}, v_j, u_{\pi(t+1,\dots,n)}|\pi)) = u_j(f(h_{t-1}, u_j, u_{\pi(t+1,\dots,n)}|\pi)) \tag{13′}$$

which is analogous to (13). The remainder of the proof is identical to the proof of Case 1 following (13). □

References

Abdulkadiroğlu, A., Sönmez, T., 2003. School choice: A mechanism design approach. American Economic Review 93 (3), 729–747.
URL http://www.jstor.org/stable/3132114

Benassy, J. P., 1982. The economics of market disequilibrium. New York: Academic Press.

Biró, P., Veski, A., 2016. Efficiency and fair access in kindergarten allocation policy design, mimeo.

Bochet, O., Sakai, T., 2010. Secure implementation in allotment economies. Games and Economic Behavior 68 (1), 35 – 49.
URL http://dx.doi.org/10.1016/j.geb.2009.04.023

Bochet, O., Tumennassan, N., 2017. One truth and a thousand lies: Focal points in mechanism design, mimeo.

Cho, I.-K., Kreps, D. M., 1987. Signaling games and stable equilibria. The Quarterly Journal of Economics 102 (2), 179–221.
URL http://www.jstor.org/stable/1885060

de Clippel, G., Moulin, H., Tideman, N., 2008. Impartial division of a dollar. Journal of Economic Theory 139 (1), 176 – 191.
URL http://dx.doi.org/10.1016/j.jet.2007.06.005

Dur, U., Hammond, R. G., Morrill, T., February 2018. Identifying the harm of manipulable school-choice mechanisms. American Economic Journal: Economic Policy 10 (1), 187–213.
URL http://www.aeaweb.org/articles?id=10.1257/pol.20160132

Fack, G., Grenet, J., He, Y., April 2019. Beyond truth-telling: Preference estimation with centralized school choice and college admissions. American Economic Review 109 (4), 1486–1529.
URL http://www.aeaweb.org/articles?id=10.1257/aer.20151422

Fujinaka, Y., Wakayama, T., 2011. Secure implementation in Shapley-Scarf housing markets. Economic Theory 48 (1), 147–169.
URL http://dx.doi.org/10.1007/s00199-010-0538-x

Govindan, S., Wilson, R., 2009. On forward induction. Econometrica 77 (1), 1–28.
URL http://www.jstor.org/stable/40056520

Klijn, F., Pais, J., Vorsatz, M., 2019. Static versus dynamic deferred acceptance in school choice: Theory and experiment. Games and Economic Behavior 113, 147 – 163.
URL https://doi.org/10.1016/j.geb.2018.08.009

Kreps, D. M., Wilson, R., 1982. Sequential equilibria. Econometrica 50 (4), 863–894.
URL http://www.jstor.org/stable/1912767

Li, S., November 2017. Obviously strategy-proof mechanisms. American Economic Review 107 (11), 3257–87.
URL http://www.aeaweb.org/articles?id=10.1257/aer.20160425

Marx, L. M., Swinkels, J. M., 1997. Order independence for iterated weak dominance. Games and Economic Behavior 18 (2), 219 – 245.
URL http://dx.doi.org/10.1006/game.1997.0525

Moore, J., Repullo, R., 1988. Subgame perfect implementation. Econometrica 56 (5), 1191–1220.
URL http://www.jstor.org/stable/1911364

Moulin, H., 1980. On strategy-proofness and single peakedness. Public Choice 35 (4), 437–455.
URL http://www.jstor.org/stable/30023824

Osborne, M., Rubinstein, A., 1994. A course in game theory. MIT Press, Cambridge, MA.

Postlewaite, A., Schmeidler, D., 1986. Implementation in differential information economies. Journal of Economic Theory 39 (1), 14 – 33.
URL https://doi.org/10.1016/0022-0531(86)90018-9

Saijo, T., Sjöström, T., Yamato, T., 2007. Secure implementation. Theoretical Economics 2 (3), 203–229.
URL http://econtheory.org/ojs/index.php/te/article/view/20070203/0

Satterthwaite, M. A., Sonnenschein, H., 1981. Strategy-proof allocation mechanisms at differentiable points. Review of Economic Studies 48 (4), 587–597.
URL http://www.jstor.org/stable/2297198

Shapley, L., Scarf, H., 1974. On cores and indivisibility. Journal of Mathematical Economics 1 (1), 23 – 37.
URL http://dx.doi.org/10.1016/0304-4068(74)90033-0

Sprumont, Y., 1991. The division problem with single-peaked preferences: A characterization of the uniform allocation rule. Econometrica 59 (2), 509–519.
URL http://www.jstor.org/stable/2938268

Velez, R. A., Brown, A. L., 2019. Empirical strategy-proofness, mimeo, Texas A&M University.

Watson, J., 2017. A general, practicable definition of perfect bayesian equilibrium, mimeo, UCSD, accessed April 21, 2019.
URL https://econweb.ucsd.edu/~jwatson/PAPERS/WatsonPBE.pdf