Partial Strategyproofness: Relaxing Strategyproofness for the Random Assignment Problem

Timo Mennle            Sven Seuken
University of Zurich            University of Zurich
January 11, 2017

Abstract

We present partial strategyproofness, a new, relaxed notion of strategyproofness for studying the incentive properties of non-strategyproof assignment mechanisms that involve some form of randomness. Informally, a mechanism is partially strategyproof if it makes truthful reporting a dominant strategy for those agents whose preference intensities differ sufficiently between any two objects. A single numerical parameter, the degree of strategyproofness, controls the extent to which these intensities must differ. We demonstrate that partial strategyproofness is a natural and useful relaxation of strategyproofness: It is axiomatically motivated, it allows a meaningful parametric and algorithmic comparison of mechanisms by their incentive properties, and it provides new insights across a wide range of popular mechanisms.

Keywords: Ordinal mechanisms, Random assignment, Strategyproofness, Stochastic dominance, Probabilistic Serial mechanism, Boston mechanism, HBS Draft mechanism

JEL: C78, D47, D82

*First version: January 13, 2014; Department of Informatics, University of Zurich, Switzerland. Email: {mennle, seuken}@ifi.uzh.ch. For updates see www.ifi.uzh.ch/ce/publications/PSP.pdf. We would like to thank (in alphabetical order) Atila Abdulkadiroğlu, Haris Aziz, Ivan Balbuzanov, Craig Boutilier, Eric Budish, Gabriel Carroll, Lars Ehlers, Katharina Huesmann, Bettina Klaus, Flip Klijn, Fuhito Kojima, Eric Maskin, Antonio Miralles, Bernardo Moreno, David Parkes, Baharak Rastegari, Alvin Roth, Steffen Schüldenzucker, Arunava Sen, Tayfun Sönmez, William Thomson, and Utku Unver for helpful comments on this work. Furthermore, we are thankful for the feedback we received from various participants at the following events (in chronological order): EC ’13, UECE Lisbon Meetings ’13, CIREQ Matching Conference ’14, EC ’14, SCW ’14, IWGTS ’14, ESWC ’15, GAMES ’16. Any errors remain our own. Some of the ideas presented in this paper were also described in a two-page abstract that was published in the conference proceedings of EC’14; see (Mennle and Seuken, 2014). This research was supported by the Hasler Foundation under grant #12078 and the SNSF (Swiss National Science Foundation) under grant #156836.
1. Introduction

The assignment problem is concerned with the allocation of indivisible objects to self-interested agents who have private preferences over these objects. Monetary transfers are not permitted, which makes this problem different from auctions and other settings with transferable utility. In practice, assignment problems often arise in situations that are of great importance to people’s lives; for example, when assigning students to seats at public schools (Abdulkadiroğlu and Sönmez, 2003), medical school graduates to entry level positions (Roth, 1984), or tenants to subsidized housing (Abdulkadiroğlu and Sönmez, 1998). Not surprisingly, since the seminal paper of Hylland and Zeckhauser (1979), the assignment problem has attracted much attention from mechanism designers.

In this paper, we study mechanisms for the assignment problem that take preference orders over objects as input. As mechanism designers, we care specifically about incentives for truthtelling under the mechanisms we design. A mechanism is *strategyproof* if truthtelling is a dominant strategy equilibrium of the associated preference revelation game. Participating in a strategyproof mechanism is simple for the agents because it eliminates the need to take the preferences or strategies of other agents into account. Strategyproofness thus yields a robust prediction of equilibrium behavior.\(^1\) This advantage makes strategyproofness the gold standard among incentive concepts.\(^2\)

The advantages of strategyproofness, however, come at a cost: Zhou (1990) showed that, in the assignment domain, it is impossible to achieve the optimum with respect to incentives, efficiency, and fairness simultaneously.\(^3\) This makes the assignment problem an interesting mechanism design challenge. For example, the Random Serial Dictatorship mechanism is strategyproof and anonymous, but only ex-post efficient. In fact, it is

---

\(^1\)In addition, strategyproofness yields a form of fairness as it *levels the playing field* between agents who have different levels of strategic ability (Pathak and Sönmez, 2008). Moreover, the truthful preference information that strategyproof mechanisms elicit can be useful beyond its role as input to the mechanism; for example, researchers rely on such information for computational experiments (see, e.g., (Abdulkadiroğlu, Pathak and Roth, 2009)), and in the context of course assignment, universities may wish to adjust course capacities in response to the true demand.

\(^2\)Recently, Li (2016) proposed *obvious strategyproofness*, a strict refinement of strategyproofness, which requires that agents can understand the optimality of truthful reporting without contingent reasoning. While the partial strategyproofness concept that we introduce in the present paper retains the ability to give honest and useful strategic advice to the agents (see Section 5.2), it is agnostic about the information-theoretic complexity of verifying the validity of this advice.

\(^3\)Specifically, Zhou (1990) showed that no (possibly random and possibly cardinal) assignment mechanism can satisfy strategyproofness, ex-ante efficiency, and equal treatment of equals.
conjectured to be the unique mechanism that satisfies all three properties (Lee and Sethuraman, 2011; Bade, 2016). The more demanding ordinal efficiency is achieved by the Probabilistic Serial mechanism, but any mechanism that achieves ordinal efficiency and equal treatment of equals cannot be strategyproof (Bogomolnaia and Moulin, 2001). Finally, rank efficiency, an even stronger efficiency concept, can be achieved with Rank Value mechanisms (Featherstone, 2011), but it is incompatible with strategyproofness, even without additional fairness requirements. Obviously, strategyproofness is in conflict with many other desiderata, and mechanism designers are therefore interested in studying non-strategyproof mechanisms. This highlights the need for good tools to capture the incentive properties of non-strategyproof mechanisms and to analyze what trade-offs are possible between incentives for truth-telling and other desiderata.  

In practice, non-strategyproof mechanisms, such as the Boston mechanism for the assignment of seats at public schools are ubiquitous. School choice administrators frequently have the explicit objective of assigning as many students as possible to their top choices (Dur, Mennle and Seuken, 2016). If students report their preferences truthfully, the Boston mechanism intuitively fares well with respect to this objective; however, it is known to be highly manipulable in theory (Abdulkadiroğlu and Sönmez, 2003), and it was found to be manipulated by students in practice (Calsamiglia and Güell, 2014). A second example is the Teach for America program, which used a mechanism that aimed at rank efficiency when assigning new teachers to positions at different schools (Featherstone, 2011). While this mechanism was not strategyproof, the organizers were confident that the majority of preferences were reported truthfully because participants lacked the information necessary to determine beneficial misreports. The prevalence of non-strategyproof mechanisms in practice raises a number of questions: What incentive guarantees do these mechanisms provide, if any? What honest and useful strategic advice can we give to the agents? How do these mechanisms compare in terms of their incentive

---

4Prior work has already led to some useful concepts for this purpose, e.g., the comparison by vulnerability to manipulation (Pathak and Sönmez, 2013) or strategyproofness in the large (Azevedo and Budish, 2015). We describe their relation to partial strategyproofness in Sections 5.4 and 6.3 respectively.

5Variants of the Boston mechanism are used, e.g., in Minneapolis, Seattle, Lee County (Kojima and Ünver, 2014), San Diego, Freiburg (Germany) (Dur, Mennle and Seuken, 2016), and throughout the German state of Nordrhein Westfalen (Basteck, Huesmann and Nax, 2015).

6Further interesting examples of the persistent use of non-strategyproof mechanisms in practice are the assignment of on-campus housing at MIT through a mechanism that minimizes the number of unassigned units (source: https://housing.mit.edu/), and the allocation of courses at Harvard Business School through the HBS Draft mechanism (Budish and Cantillon, 2012).
properties? and how can we quantify their “non-truthfulness”? To be useful for market design, an incentive concept should provide answers to these questions.

It should come as no surprise that researchers have been calling for useful relaxations of strategyproofness before (Budish, 2012). In this paper, we introduce partial strategyproofness, a new, relaxed notion of strategyproofness. Partial strategyproofness is particularly suited to the analysis of assignment mechanisms that involve some form of randomness. The randomness may come from any of three sources: First, it may be an intrinsic feature of the mechanism (e.g., of the Probabilistic Serial mechanism (Bogomolnaia and Moulin, 2001)), which corresponds to our typical understanding of the random assignment problem. Second, the randomness may be injected via random priorities; e.g., ties in school choice mechanisms are typically broken via lotteries (Erdil and Ergin, 2008). Third, the randomness may arise from the agents’ uncertainty about the other agents’ reports (which Azevedo and Budish (2015) call the interim perspective). We show that partial strategyproofness provides compelling answers to all the questions we have raised, independent of the particular source of randomness.

We now illustrate the definition of partial strategyproofness with a motivating example. Consider a setting in which three agents, conveniently named 1, 2, 3, compete for three objects, $a$, $b$, $c$, each in unit capacity. Suppose that the agents’ preferences are $P_1 : a > b > c$, $P_2 : b > a > c$, and $P_3 : b > c > a$; that agent 1 has utility 0 for its last choice $c$; and that the non-strategyproof Probabilistic Serial mechanism$^7$ is used to assign the objects. If all agents report their preferences truthfully, then agent 1 receives $a$, $b$, $c$ with probabilities $3/4$, $0$, $1/4$ respectively. If agent 1 reports $P'_1 : b > a > c$ instead, then these probabilities change to $1/2$, $1/3$, $1/6$. Observe that whether or not the misreport $P'_1$ increases agent 1’s expected utility depends on how intensely it prefers $a$ over $b$: If $u_1(a)$ is close to $u_1(b)$, then agent 1 would benefit from the misreport $P'_1$. If $u_1(a)$ is significantly larger than $u_1(b)$, then agent 1 would prefer to report truthfully. Specifically, agent 1 prefers to report truthfully if $3/4 \cdot u_1(a) \geq u_1(b)$.

Our definition of partial strategyproofness generalizes the intuition from this motivating example: For some bound $r$ in $[0, 1]$, we say that agent $i$’s utility function $u_i$ satisfies

---

$^7$The Probabilistic Serial mechanism (Bogomolnaia and Moulin, 2001) determines a random assignment via the Simultaneous Eating algorithm: All agents consume probability shares of their respective most-preferred objects at equal speeds; when an object is exhausted, the agents consuming this object move on and consume shares of their respective next most-preferred object. See Section 10 of the Online Appendix for a detailed definition.
uniformly relatively bounded indifference with respect to $r$ (URBI($r$)) if $r \cdot u_i(a) \geq u_i(b)$ holds whenever $i$ prefers $a$ to $b$ (after appropriate normalization). A mechanism is $r$-partially strategyproof if it makes truthful reporting a dominant strategy for any agent whose utility function satisfies URBI($r$). Varying the degree of strategyproofness $r$ between 0 and 1 gives rise to a new, parametrized class of incentive concepts. In fact, we find (in Section 8.2) that the Probabilistic Serial mechanism has a degree of strategyproofness of exactly $3/4$ in a setting with three agents and three objects.

We argue that partial strategyproofness is a natural and useful way to think about the incentive properties of non-strategyproof assignment mechanisms that involve some form of randomness. The following three arguments support our claim.

**Partial strategyproofness has a compelling axiomatic motivation.** Towards showing this, we first prove that full strategyproofness can be decomposed into three simple axioms. Each of the axioms restricts the way in which a mechanism can react when an agent swaps two consecutively ranked objects in its preference report (e.g., from $P_i : a > b$ to $P'_i : b > a$).

1. A mechanism is *swap monotonic* if either the swap makes it more likely that the agent receives $b$ (the object that it claims to prefer) and less likely that the agent receives $a$, or the mechanism does not react to the swap at all. In other words, any swap must affect at least the agent’s probabilities for the objects that trade positions, and the change must be monotonic in the agent’s reported preferences.

2. A mechanism is *upper invariant* if an agent cannot improve its chances for objects that it likes more by misrepresenting its preferences for objects that it likes less. Precisely, the swap of $a$ and $b$ must not affect the agent’s chances for any object that it prefers strictly to $a$.

3. A mechanism is *lower invariant* if an agent cannot affect its chances for less-preferred objects by swapping $a$ and $b$. This axiom is the complement to upper invariance but for objects that the agent likes strictly less than $b$.

For our first main result, we show that strategyproofness can be decomposed into these three axioms: A mechanism is strategyproof if and only if it is swap monotonic, upper invariant, and lower invariant (Theorem 1). Intuitively, lower invariance is the least important of the three axioms (see Section 5.2 for a formal argument), and by dropping it, we arrive at the larger class of partially strategyproof mechanisms: We show that a
mechanism is $r$-partially strategyproof for some $r > 0$ if and only if it is swap monotonic and upper invariant (Theorem 2). Thus, partial strategyproofness describes the incentive properties of (possibly non-strategyproof) mechanisms that retain the two most important of the three axioms that constitute strategyproofness.

**Partial strategyproofness allows a meaningful parametric and algorithmic comparison of mechanisms by their incentive properties.** By construction, a greater degree of strategyproofness $r$ means that the mechanism is guaranteed to provide good incentives for a larger set of utility functions. An $r$ of 1 is equivalent to strategyproofness, and the lower limit concept of $r$-partial strategyproofness as $r$ approaches 0 is the weaker lexicographic strategyproofness (Theorem 3). The degree of strategyproofness thus parametrizes the spectrum of incentive concepts between these two concepts. In terms of *computability*, we show that partial strategyproofness can be verified algorithmically, and that the degree of strategyproofness can be computed (for any mechanism in any finite setting). Other approaches to studying the incentive properties of non-strategyproof mechanisms include *strategyproofness in the large* (Azevedo and Budish, 2015), or the comparison of mechanisms by their *vulnerability to manipulation* (Pathak and Sönmez, 2013). The insights from these other approaches are not in conflict with those from partial strategyproofness. However, in contrast to strategyproofness in the large, partial strategyproofness describes incentives in every finite setting separately. Moreover, neither of the two other approaches is parametric, and no algorithms are known to verify that they apply to a given mechanism.

**Partial strategyproofness provides new insights into the incentive properties of existing (and new) non-strategyproof mechanisms.** The Probabilistic Serial mechanism (PS) serves as a typical example of a partially strategyproof mechanism. Bogomolnaia and Moulin (2001) had already shown that PS is weakly strategyproof, and Balbuzanov (2015) strengthened this result by showing that it is convex strategyproof. Since partial strategyproofness implies convex strategyproofness, our result that PS is

---

8We prove *maximality* of the underlying domain restriction for swap monotonic, upper invariant, anonymous mechanisms (Proposition 1). This maximality implies that no other single parameter measure can be strictly more powerful at differentiating mechanisms by their incentive properties.

9This follows from a *dominance interpretation* of partial strategyproofness (Theorem 4), which generalizes the well-known equivalence of strategyproofness and SD-strategyproofness (Erdil, 2014).

10Convergence of the degree of strategyproofness $r$ to 1 in large settings implies strategyproofness in the large, but the converse may not hold (Proposition 4). The comparison by degree of strategyproofness and by vulnerability to manipulation never contradict each other (Proposition 7).
partially strategyproof (Proposition 5) is the strongest statement about its incentive properties in any finite setting that is known at present.\(^{11}\) A second insight from partial strategyproofness concerns the comparison of two popular non-strategyproof school choice mechanisms. Under the classical Boston mechanism (BM),\(^{12}\) all applications in the \(k^{th}\) round go to the students’ \(k^{th}\) choices. In particular, a student may waste one round by applying to a school with no more unfilled seats. This induces an obvious incentive to misreport by skipping exhausted schools in the application process. The adaptive Boston mechanism (ABM) (Mennle and Seuken, 2017d) works similarly to BM, except that students apply to their best available school in each round. Intuitively, this tweak removes part of the manipulability, but a formal understanding of this intuition has so far remained elusive.\(^{13}\) Partial strategyproofness now enables this formal understanding: With priorities determined by a single, uniform lottery, ABM is partially strategyproof, but BM is not.

We conclude that the axiomatic motivation, the parametric degree of strategyproofness, and the new insights about non-strategyproof mechanisms qualify partial strategyproofness as a natural and useful relaxed notion of strategyproofness for assignment mechanisms that involve some form of randomness.

2. Formal Model

A setting \((N, M, q)\) consists of a set of agents \(N\) \((n = \#N)\), a set of objects \(M\) \((m = \#M)\), and a vector \(q = (q_1, \ldots, q_m)\) of capacities (i.e., there are \(q_j\) units of object \(j\) available). We assume \(n \leq \sum_{j \in M} q_j\) (i.e., there are not more agents than the total number of units); otherwise we include a dummy object with capacity \(n\). Each agent \(i \in N\) has a strict preference order \(P_i\) over objects, where \(P_i : a > b\) indicates that agent \(i\) prefers object \(a\) to object \(b\). Let \(\mathcal{P}\) be the set of all possible preference orders. A preference profile \(P = (P_i)_{i \in N} \in \mathcal{P}^N\) is a collection of preference orders from all agents, and \(P_{-i} \in \mathcal{P}^{N \setminus \{i\}}\) is

\(^{11}\)We show that \(r\)-partial strategyproofness \((r > 0)\) implies weak SD-strategyproofness (Bogomolnaia and Moulin, 2001), convex strategyproofness (Balbuzanov, 2015), approximate strategyproofness (Carroll, 2013), and lexicographic dominance-strategyproofness (Cho, 2012).

\(^{12}\)Under the Boston mechanism (Abdulkadiroğlu and Sönmez, 2003), students apply to schools in rounds. All students first apply to their first choice and applications are accepted according to priority. All students who did not receive their first choice apply to their second choice in the second round and are accepted into unfilled seats. This process continues until no new applications are received.

\(^{13}\)For fixed priorities, BM and ABM are as manipulable as each other, and for random priorities they cannot be compared by their vulnerability to manipulation (Dur, Mennle and Seuken, 2016).
a collection of preference orders of all agents except $i$. We extend agents’ preferences to lotteries via von Neumann–Morgenstern utility functions: A utility function $u_i : M \rightarrow \mathbb{R}^+$ is consistent with a preference order $P_i$ if $u_i(a) > u_i(b)$ whenever $P_i : a > b$, denoted $u_i \sim P_i$. $U_{P_i}$ denotes the set of all utility functions consistent with $P_i$.

A (random) assignment is represented by an $n \times m$-matrix $x = (x_{i,j})_{i \in N, j \in M}$, where no object is assigned beyond capacity (i.e., $\sum_{i \in N} x_{i,j} \leq q_j$ for all $j \in M$), and each agent receives some object with certainty (i.e., $\sum_{j \in M} x_{i,j} = 1$ for all $i \in N$ and $x_{i,j} \geq 0$ for all $i \in N, j \in M$). The value $x_{i,j}$ is the probability that agent $i$ gets object $j$. An assignment $x$ is deterministic if $x_{i,j} \in \{0, 1\}$ for all $i \in N, j \in M$. The $i^{th}$ row $x_i = (x_{i,j})_{j \in M}$ of $x$ is called the assignment vector of $i$ (or $i$’s assignment). The Birkhoff–von Neumann Theorem and its extensions (Budish et al., 2013) ensure that for any random assignment we can find a lottery over deterministic assignments that implements its marginal probabilities. Finally, let $X$ and $\Delta(X)$ denote the spaces of all deterministic and random assignments respectively.

A (random assignment) mechanism is a mapping $\varphi : \mathcal{P}^N \rightarrow \Delta(X)$ that selects an assignment based on a preference profile. $\varphi_i(P_i, P_{-i})$ denotes the assignment of agent $i$ when $i$ reports $P_i$ and the other agents report $P_{-i}$. The mechanism $\varphi$ is deterministic if it selects deterministic assignments (i.e., $\varphi : \mathcal{P}^N \rightarrow X$). Note that we only consider ordinal mechanisms, where the assignment only depends on the reported preference profiles but is independent of the underlying utility functions. If agent $i$ with utility function $u_i$ reports $P_i$ and the other agents report $P_{-i}$, then agent $i$’s expected utility is

$$E_{\varphi_i(P_i, P_{-i})}[u_i] = \sum_{j \in M} u_i(j) \cdot \varphi_{i,j}(P_i, P_{-i}).$$

(1)

Remark 1. Our formal model treats randomness as an intrinsic feature of mechanisms. This corresponds to the typical understanding of the random assignment problem as studied by Bogomolnaia and Moulin (2001). Randomness may also arise from random priorities, which play an important role for tie-breaking in school choice. Since random priorities are outside the control of the agents, the randomness that they induce can formally be treated as if it was an intrinsic feature of the mechanism. A third potential source of randomness is each agent’s uncertainty about the preference reports from the other agents; Azevedo and Budish (2015) call this the interim perspective. In Section 16 of the Online Appendix, we present an extension of our model to this interim perspective.
3. Strategyproofness and Incentive Axioms

In this section, we formally define strategyproofness and introduce the three axioms that we subsequently use to decompose and relax strategyproofness.

**Definition 1 (Strategyproofness).** A mechanism $\varphi$ is strategyproof if, for all agents $i \in N$, all preference profiles $(P_i, P_{-i}) \in \mathcal{P}^N$, all misreports $P'_i \in \mathcal{P}$, and all consistent utility functions $u_i \in U_i$, we have

$$E_{\varphi_i(P_i, P_{-i})}[u_i] \geq E_{\varphi_i(P'_i, P_{-i})}[u_i].$$

(2)

In words, strategyproofness requires that truthful reporting maximizes any agent’s expected utility, independent of that agent’s particular preferences, its utility function, or the preference reports of the other agents.

Alternatively, strategyproofness can be defined in terms of stochastic dominance: For a preference order $P_i \in \mathcal{P}$ (where $P_i : j_1 > \ldots > j_m$) and two assignment vectors $x_i, y_i$, we say that $x_i$ stochastically dominates $y_i$ at $P_i$ if, for all ranks $K \in \{1, \ldots, m\}$, we have

$$\sum_{k=1}^{K} x_{i,j_k} \geq \sum_{k=1}^{K} y_{i,j_k}.$$  

(3)

In words, $i$ has a weakly higher probability of obtaining one of its top-$K$ choices under $x_i$ than under $y_i$. This dominance is strict if inequality (3) is strict for at least one $K \in \{1, \ldots, m\}$. A mechanism $\varphi$ is stochastic dominance-strategyproof (SD-strategyproof) if, for all agents $i \in N$, all preference profiles $(P_i, P_{-i}) \in \mathcal{P}^N$, and all misreports $P'_i \in \mathcal{P}$, $\varphi_i(P_i, P_{-i})$ stochastically dominates $\varphi_i(P'_i, P_{-i})$ at $P_i$. Erdil (2014) showed that strategyproofness (Definition 1) and SD-strategyproofness are equivalent; and throughout the paper we will refer to this property simply as strategyproofness.

Next, we introduce three axioms. To state the axioms formally, we require two auxiliary concepts: For any preference order $P_i \in \mathcal{P}$, the neighborhood of $P_i$ is the set of all preference orders that differ from $P_i$ by a swap of two consecutively ranked objects, denoted $N_{P_i}$ (e.g., the neighborhood of $P_i : a > b > c$ includes $P'_i : b > a > c$ but not $P''_i : c > a > b$). For any object $j \in M$, the upper contour set of $j$ at $P_i$ is the set of objects that $i$ prefers strictly to $j$, denoted $U(j, P_i)$. Conversely, the lower contour set of $j$ at $P_i$ is the set of objects that $i$ likes strictly less than $j$, denoted $L(j, P_i)$. For example,
the upper contour set of $b$ at $P_i: a > b > c > d$ is $\{a\}$ and the lower contour set is $\{c,d\}$.

Swapping two consecutive objects in the true preference order (or equivalently, reporting a preference order from the neighborhood of the true preference order) is a basic kind of misreport. The axioms we define limit the way in which a mechanism can change an agent’s assignment when this agent uses such a basic kind of misreport.

**Axiom 1 (Swap Monotonicity).** A mechanism $\varphi$ is swap monotonic if, for all agents $i \in N$, all preference profiles $(P_i, P_{-i}) \in \mathcal{P}^{N}$, and all misreports $P'_i \in NP_i$ from the neighborhood of $P_i$ with $P_i: a > b$ but $P'_i: b > a$, one of the following holds:

- either: $\varphi_i(P_i, P_{-i}) = \varphi_i(P'_i, P_{-i})$,
- or: $\varphi_{i,a}(P_i, P_{-i}) > \varphi_{i,a}(P'_i, P_{-i})$ and $\varphi_{i,b}(P_i, P_{-i}) < \varphi_{i,b}(P'_i, P_{-i})$.

Swap monotonicity requires that the mechanism reacts to the swap in a direct and responsive way: The swap reveals information about the agent’s relative preference over $a$ and $b$. Thus, if anything changes about that agent’s assignment, then the assignment for $a$ and $b$ must be affected directly. Moreover, the mechanism must respond to the new information by increasing the assignment for the object that the agent reports to like more and reducing the assignment for the object that the agent reports to like less.

For deterministic mechanisms, swap monotonicity is equivalent to strategyproofness (see Proposition 6 in Section 11 of the Online Appendix). For the more general class of random mechanisms, swap monotonicity is weaker than strategyproofness but prevents a certain obvious kind of manipulability: Consider a mechanism that assigns an agent’s reported first choice with probability 1/3 and its reported second choice with probability 2/3. The agent is unambiguously better off by ranking its second choice first. Swap monotonicity precludes such opportunities for manipulation. Nevertheless, even swap monotonic mechanisms may be manipulable in a stochastic dominance sense, as the following example shows.

**Example 1.** Consider a mechanism where reporting $P_i: a > b > c > d$ leads to an assignment of $(0, 1/2, 0, 1/2)$ for $a, b, c, d$ respectively and swapping $b$ and $c$ (i.e., reporting $P'_i: a > c > b > d$) leads to $(1/2, 0, 1/2, 0)$. This is consistent with swap monotonicity; yet, the latter assignment stochastically dominates the former at $P_i$.\(^{14}\)

\(^{14}\)In Section 12 of the Online Appendix, we provide a swap monotonic continuation of this mechanism.
Note that while the swap monotonic mechanism in Example 1 is manipulable in a stochastic dominance sense, the manipulation leads to changes in the assignment for an object that the agent prefers to both \( b \) and \( c \). Our next axiom restricts the changes in the assignment for these objects under a swap.

**Axiom 2** (Upper Invariance). A mechanism \( \varphi \) is *upper invariant* if, for all agents \( i \in N \), all preference profiles \((P_i, P_{-i}) \in \mathcal{P}^N\), and all misreports \( P'_i \in N_{P_i} \) with \( P_i : a > b \) but \( P'_i : b > a \), we have that \( \varphi_{i,j}(P_i, P_{-i}) = \varphi_{i,j}(P'_i, P_{-i}) \) for all \( j \in U(a, P_i) \).

Intuitively, upper invariance ensures that agents cannot influence their chances of obtaining more-preferred objects by changing the order of less-preferred objects. This axiom was introduced by Hashimoto et al. (2014) as one of the central axioms to characterize the Probabilistic Serial mechanism. If an outside option is available and if the mechanism is individually rational, then upper invariance is equivalent to truncation robustness.\(^{15}\) Many assignment mechanisms satisfy upper invariance, including Random Serial Dictatorship, Probabilistic Serial, the Boston mechanism, Deferred Acceptance (for the proposing agents), Top-Trade Cycles, and the HBS Draft mechanism.

Finally, our third axiom is symmetric to upper invariance but restricts the changes in the assignment for objects that the agent likes less than those objects that get swapped.

**Axiom 3** (Lower Invariance). A mechanism \( \varphi \) is *lower invariant* if, for all agents \( i \in N \), all preference profiles \((P_i, P_{-i}) \in \mathcal{P}^N\), and all misreports \( P'_i \in N_{P_i} \) with \( P_i : a > b \) but \( P'_i : b > a \), we have that \( \varphi_{i,j}(P_i, P_{-i}) = \varphi_{i,j}(P'_i, P_{-i}) \) for all \( j \in L(b, P_i) \).

In words, a mechanism is lower invariant if changing the order of two consecutive objects does not affect the agents’ chances of obtaining any less-preferred objects.

### 4. Decomposition of Strategyproofness

Each of the three axioms that we have introduced affects incentives by preventing manipulations from being beneficial in particular ways. We are now ready to formulate our first main result that strategyproofness can be decomposed into these axioms.

\(^{15}\) *Truncation robustness* is a type of robustness to manipulation: It requires that agents cannot increase their chances for preferred objects by bringing the outside option up in their preference report.
Theorem 1. A mechanism $\varphi$ is strategyproof if and only if it is swap monotonic, upper invariant, and lower invariant.

Proof. “$\Rightarrow$”. Assume towards contradiction that $\varphi$ is strategyproof but not upper invariant. Then there exists some agent $i \in N$, some preference profile $P = (P_i, P_{-i}) \in \mathcal{P}^N$, some misreport $P'_i \in N_{P_i}$ with $P_i : a > b$ but $P'_i : b > a$, and an object $j$ which $i$ prefers strictly to $a$ and $b$ with $\varphi_{i,j}(P'_i, P_{-i}) \neq \varphi_{i,j}(P_i, P_{-i})$. Choose $j$ to be $i$’s most-preferred such object. Without loss of generality, $\varphi_{i,j}(P'_i, P_{-i}) > \varphi_{i,j}(P_i, P_{-i})$ (otherwise, we can reverse the roles of $P'_i$ and $P_i$). Then $\varphi_i(P_i, P_{-i})$ does not stochastically dominate $\varphi_j(P'_i, P_{-i})$, a contradiction to strategyproofness. Lower invariance follows by an analogous argument, except that we take $j$ to be the least-preferred choice for which $i$’s assignment changes.

By upper and lower invariance, any swap of two consecutive objects (e.g., from $P_i : a > b$ to $P'_i : b > a$) only leads to a re-distribution of probability between $a$ and $b$. If misreport $P'_i$ leads to a strictly higher assignment for $a$, then $\varphi_i(P'_i, P_{-i})$ strictly stochastically dominates $\varphi_i(P_i, P_{-i})$ at $P_i$, a contradiction; and swap monotonicity follows.

“$\Leftarrow$”. We invoke a result from (Carroll, 2012) that strategyproofness can be shown by verifying that no agent can benefit from swapping two consecutive objects. Let $\varphi$ be a swap monotonic, upper invariant, and lower invariant mechanism, and consider any agent $i \in N$, any preference profile $(P_i, P_{-i}) \in \mathcal{P}^N$, and any misreport $P'_i \in N_{P_i}$ with $P_i : a > b$ but $P'_i : b > a$. Observe that $\varphi_i(P_i, P_{-i})$ stochastically dominates $\varphi_i(P'_i, P_{-i})$ at $P_i$: By upper and lower invariance, $i$’s assignment for all objects remains constant under the misreport, except possibly its assignment for $a$ and $b$; and, by swap monotonicity, $i$’s assignments for $a$ and $b$ can only decrease and increase respectively.

Theorem 1 illustrates why strategyproofness is such a restrictive requirement: If an agent swaps two consecutive objects (e.g., from $P_i : a > b$ to $P'_i : b > a$), the only way in which a strategyproof mechanism can react is by increasing that agent’s assignment for $b$ and decreasing its assignment for $a$ by the same amount.

In his seminal paper, Gibbard (1977) gave a similar decomposition of strategyproofness for random voting rules. He showed that any rule is strategyproof if and only if it is localized and non-perverse. Of course, all decompositions of strategyproofness are (by definition) equivalent. The appeal of our Theorem 1 lies in the choice of axioms, which are simple and easily motivated. They make our decomposition particularly useful, e.g., when verifying strategyproofness of new mechanisms, or when encoding strategyproofness.
as constraints under the automated mechanism design paradigm (Sandholm, 2003).

Remark 2. While we focus on strict preferences in the present paper, Theorem 1 can be extended to the case of weak preferences; see (Mennle and Seuken, 2017a).

Remark 3. To obtain strategyproofness in Theorem 1, we have used a local sufficiency result for strategyproofness by Carroll (2012). Cho (2012) proved an analogous local sufficiency result for lexicographic strategyproofness. In (Mennle and Seuken, 2017c), we have proven a local sufficiency result for partial strategyproofness that unifies these two prior local sufficiency results.

5. Partial Strategyproofness

Theorem 1 serves as another reminder that strategyproofness is a restrictive requirement, and that it may be worthwhile to think about the incentive properties of non-strategyproof mechanisms. In this section, we define partial strategyproofness, a new, relaxed notion of strategyproofness, and we show how this concept arises by dropping lower invariance from the decomposition of strategyproofness in Theorem 1.

5.1. URBI Domain Restriction and Partial Strategyproofness

Recall the motivating example from the Introduction, where agent 1 is contemplating a misreport under the Probabilistic Serial mechanism. \( r = 3/4 \) is the pivotal degree of indifference, which determines whether the misreport is beneficial to the agent or not. The following domain restriction generalizes this intuition of agents who differentiate sufficiently between different objects.

Definition 2 (URBI). A utility function \( u_i \) satisfies uniformly relatively bounded indifference with respect to bound \( r \in [0,1] \) (URBI(r)) if, for all objects \( a, b \in M \) with \( u_i(a) > u_i(b) \), we have

\[
 r \cdot \left( u_i(a) - \min_{j \in M} u_i(j) \right) \geq u_i(b) - \min_{j \in M} u_i(j). \tag{4}
\]

If \( \min_{j \in M} (u_i(j)) = 0 \) (i.e., the last choice yields utility 0), then inequality (4) simplifies to \( r \cdot u_i(a) \geq u_i(b) \). In words, agent \( i \) must value \( b \) at least a factor \( r \) less than \( a \) whenever
Figure 1: Geometric interpretation of uniformly relatively bounded indifference.

For a geometric interpretation, consider Figure 1: The condition means that $i$’s utility function (e.g., the point labeled $u_i$) cannot be arbitrarily close to the indifference hyperplane $H(P_i, P'_i)$, but it must lie within the shaded triangular area. $r$ is the slope of the dashed line that bounds this area at the top. Any utility function in $U_{P_i}$ that lies outside the shaded area (e.g., the point labeled $\tilde{u}_i$) violates URBI($r$). For convenience we denote by URBI($r$) the set of utility functions that satisfy this constraint.

**Remark 4.** To gain some intuition about the size of the set URBI($r$), consider a setting with $m = 3$ objects. Suppose that $\min_{j \in M} (u_i(j)) = 0$ and that the utilities for the first and second choice are determined by drawing a point uniformly at random from $(0,1)^2 \setminus H(P_i, P'_i)$ (i.e., from the open unit square excluding the indifference hyperplane). Then the share of utilities that satisfy URBI($r$) is exactly $r$. For example, if $r = 3/4$, the probability of drawing a utility function that satisfies URBI($3/4$) is $3/4$.

We are now ready to define our new relaxed incentive concept.

**Definition 3** (Partial Strategyproofness). Given a setting $(N, M, q)$ and a bound $r \in [0, 1]$, a mechanism $\varphi$ is $r$-partially strategyproof (in the setting $(N, M, q)$) if, for all agents $i \in N$, all preference profiles $(P_i, P_{-i}) \in \mathcal{P}^N$, all misreports $P'_i \in \mathcal{P}$, and all utility functions $u_i \in U_{P_i} \cap \text{URBI}(r)$, we have

$$\mathbb{E}_{\varphi_i(P_i, P_{-i})}[u_i] \geq \mathbb{E}_{\varphi_i(P'_i, P_{-i})}[u_i].$$

We say that $\varphi$ is partially strategyproof (in $(N, M, q)$) if it is $r$-partially strategyproof in $(N, M, q)$ for some positive bound $r > 0$. 

14
In words, an \( r \)-partially strategyproof mechanisms makes truthful reporting a dominant strategy, but only for agents whose utility functions satisfy \( \text{URBI}(r) \). Since all utility functions satisfy \( \text{URBI}(1) \), 1-partial strategyproofness is equivalent to full strategyproofness, and since the set \( \text{URBI}(r) \) is smaller for smaller values of \( r \), the \( r \)-partial strategyproofness requirement is less demanding for smaller \( r \).

5.2. Decomposition of Partial Strategyproofness

For our second main result, we prove that partial strategyproofness can be decomposed into the axioms swap monotonicity and upper invariance.

**Theorem 2.** Given a setting \((N, M, q)\), a mechanism \( \varphi \) is partially strategyproof (i.e., \( r \)-partially strategyproof for some \( r > 0 \)) if and only if it is swap monotonic and upper invariant.

**Proof.** We use the following lemma, which is proven in Appendix A.

**Lemma 1.** Given a setting \((N, M, q)\) and a mechanism \( \varphi \), the following are equivalent:

a. For all agents \( i \in N \), all preference profiles \((P_i, P_{-i}) \in \mathcal{P}^N \), and all misreports \( P'_i \in \mathcal{P} \) with \( \varphi_i(P_i, P_{-i}) \neq \varphi_i(P'_i, P_{-i}) \), there exists an object \( a \in M \) such that \( \varphi_{i,a}(P_i, P_{-i}) > \varphi_{i,a}(P'_i, P_{-i}) \) and \( \varphi_{i,j}(P_i, P_{-i}) = \varphi_{i,j}(P'_i, P_{-i}) \) for all \( j \in U(a, P_i) \),

b. \( \varphi \) is swap monotonic and upper invariant.

Lemma 1 establishes a connection between the behavior of mechanisms under arbitrary misreports (i.e., statement a) and local properties (i.e., swap monotonicity and upper invariance in statement b). It replaces the local sufficiency result of Carroll (2012) which we have used to prove Theorem 1.

Towards Theorem 2, we must show that partial strategyproofness and statement a of Lemma 1 are equivalent. For the fixed setting \((N, M, q)\) let \( \delta \) be the smallest non-zero variation in the assignment resulting from any change of report by any agent; formally,

\[
\delta = \min \left\{ \left| \varphi_{i,j}(P_i, P_{-i}) - \varphi_{i,j}(P'_i, P_{-i}) \right| \right\}_{i \in N, j \in M, (P_i, P_{-i}) \in \mathcal{P}^N, P'_i \in \mathcal{P}, \text{s.t. } \left| \varphi_{i,j}(P_i, P_{-i}) - \varphi_{i,j}(P'_i, P_{-i}) \right| \neq 0}.
\]  

\( \delta \) is strictly positive for any non-constant \( \varphi \), and well-defined as \( N, M, \) and \( \mathcal{P} \) are finite.
Necessity (partial strategyproofness ⇐ a). For any agent \( i \in N \) and any preference profile \((P_i, P_{-i}) \in \mathcal{P}^N\), suppose that \( i \) considers some misreport \( P'_i \in \mathcal{P} \). If \( \varphi_i(P_i, P_{-i}) \neq \varphi_i(P'_i, P_{-i}) \), then a implies that there exists some object \( a \) for which \( i \)'s assignment strictly decreases, and \( i \)'s assignment for all more-preferred objects \( j \in U(a, P_i) \) remains unchanged. Since \( i \)'s assignment for \( a \) changes, it must decrease by at least \( \delta \). Let \( b \) be the object that \( i \) ranks directly below \( a \) in \( P_i \). Then \( i \)'s gain in expected utility from misreporting is greatest if, first, \( i \) receives \( a \) with probability \( \delta \) and its last choice with probability \((1 - \delta)\) when being truthful, and second, \( i \) receives \( b \) with certainty when misreporting. The gain is thus bounded from above by

\[
\left( \delta \cdot u_i(a) + (1 - \delta) \cdot \min_{j \in M} u_i(j) \right),
\]

which is weakly negative if

\[
\delta \cdot \left( u_i(a) - \min_{j \in M} u_i(j) \right) \geq u_i(b) - \min_{j \in M} u_i(j).
\]

Inequality (8) holds for all utility functions in \( U_{P_i} \cap \text{URBI}(\delta) \). Thus, if \( i \)'s utility function satisfies \( \text{URBI}(\delta) \), then no misreport increases \( i \)'s expected utility, or equivalently, \( \varphi \) is \( \delta \)-partially strategyproof.

Sufficiency (partial strategyproofness ⇒ a). Let \( \varphi \) be \( r \)-partially strategyproof for some \( r > 0 \), and assume towards contradiction that \( \varphi \) violates \( a \). This is equivalent to saying that there exists an agent \( i \in N \), a preference profile \((P_i, P_{-i}) \in \mathcal{P}^N\), a misreport \( P'_i \in \mathcal{P} \) with \( \varphi_i(P_i, P_{-i}) \neq \varphi_i(P'_i, P_{-i}) \), and an object \( a \in M \), such that \( \varphi_{i,a}(P_i, P_{-i}) < \varphi_{i,a}(P'_i, P_{-i}) \) but \( \varphi_{i,j}(P_i, P_{-i}) = \varphi_{i,j}(P'_i, P_{-i}) \) for all \( j \in U(a, P_i) \). Again, let \( b \) be the object that \( i \) ranks directly below \( a \) in \( P_i \). Since \( i \)'s assignment for \( a \) increases, it must increase by at least \( \delta \). Thus, \( i \)'s gain in expected utility is smallest if, first, \( i \) receives \( b \) with certainty when being truthful, and second, \( i \) receives \( a \) with probability \( \delta \) and its last choice with probability \((1 - \delta)\) when misreporting. This makes

\[
\left( \delta \cdot u_i(a) + (1 - \delta) \cdot \min_{j \in M} u_i(j) \right) - u_i(b)
\]
a lower bound on \( i \)'s gain from misreporting. This bound is strictly positive if

\[
\delta \cdot \left( u_i(a) - \min_{j \in M} u_i(j) \right) > u_i(b) - \min_{j \in M} u_i(j),
\]

which holds for all utility functions in \( U_P \cap \text{URBI}(r) \) for \( r < \delta \). Therefore, \( \varphi \) cannot be \( r \)-partially strategyproof for any \( r > 0 \), a contradiction.

Theorem 2 provides an axiomatic motivation for the definition of partial strategyproofness: The class of partially strategyproof mechanisms consists exactly of those mechanisms that are swap monotonic and upper invariant, but may violate lower invariance. The equivalence also allows the formulation of straightforward and honest strategic advice that can be given to agents who participate in swap monotonic and upper invariant mechanisms: They are best off reporting their preferences truthfully as long as their preference intensities for any two different objects are sufficiently different.\(^{16}\)

Remark 5. In light of the interpretation of upper invariance as robustness to manipulation by truncation (Hashimoto et al., 2014) and the fact that, for deterministic mechanisms, swap monotonicity is equivalent to strategyproofness (see Proposition 6 in Section 11 of the Online Appendix), dropping lower invariance suggests itself as a sensible approach to relaxing strategyproofness. Alternatively, one may consider mechanisms that violate a different axiom instead (i.e., swap monotonicity or upper invariance). However, neither of these approaches admits the construction of interesting ordinally efficient mechanisms.\(^{17}\)

5.3. Maximality of the URBI Domain Restriction

In this section, we study how well the URBI(\( r \)) domain restriction captures the incentive properties of non-strategyproof assignment mechanisms. By definition, \( r \)-partial strategyproofness implies that the set URBI(\( r \)) must be contained in the set of utility functions for which truthful reporting is a dominant strategy. However, the two sets may not be exactly equal, as the following example shows.

\(^{16}\)As we show in Section 6.2, this advice can be refined further: An agent who is close to being indifferent between some objects may potentially gain from misreporting, but this gain is bounded.

\(^{17}\)There exist mechanisms that are swap monotonic (SM), upper invariant (UI), ordinally efficient (OE), anonymous (AN), neutral (NE), and non-bossy (NB) (e.g., Probabilistic Serial); however, as we have shown in (Mennle and Seuken, 2017e), no mechanism is SM, LI, OE, AN, NE, and NB. Moreover, no mechanism is UI, LI, AN, and OE. Thus, dropping either UI or SM (instead of LI) from strategyproofness does not enable the design of new and appealing ordinally efficient mechanisms.
Example 2. Consider a setting with four agents and four objects $a$, $b$, $c$, $d$ in unit capacity. In this setting, the adaptive Boston mechanism (see Section 8.4) with priorities determined by a single, uniform lottery is $1/3$-partially strategyproof but not $r$-partially strategyproof for any $r > 1/3$. However, it is a simple (though tedious) exercise to verify that an agent with utilities $u_i(a) = 6$, $u_i(b) = 2$, $u_i(c) = 1$, $u_i(d) = 0$ cannot benefit from misreporting, independent of the reports from the other agents. But $u_i$ violates URBI($1/3$), since
\[
\frac{u_i(c) - \min_{j \in M} u(j)}{u_i(b) - \min_{j \in M} u(j)} = \frac{1 - 0}{2 - 0} = 1/2 > 1/3.
\] (11)

Example 2 shows that for some specific $r$-partially strategyproof mechanism there may exist utility functions that violate URBI($r$), but for which truthful reporting is nonetheless a dominant strategy. This raises the question whether URBI($r$) is too small in the sense that it excludes some utility functions for which all $r$-partially strategyproof mechanisms make truthful reporting a dominant strategy. Our next proposition, which shows maximality of the URBI($r$) domain restriction, dispels this concern.

Proposition 1. For all settings $(N, M, q)$ with $m \geq 3$ objects, all bounds $r \in (0, 1)$, all preference orders $\hat{P}_i \in \mathcal{P}$, and all utility functions $\hat{u}_i \in U_{\hat{P}_i}$ that violate URBI($r$), there exists a mechanism $\hat{\varphi}$ such that

1. $\hat{\varphi}$ is $r$-partially strategyproof,
2. $\mathbb{E}_{\hat{\varphi}(\hat{P}_i, P_{-i})}[\hat{u}_i] < \mathbb{E}_{\hat{\varphi}(P'_i, P_{-i})}[\hat{u}_i]$ for some misreport $P'_i \in \mathcal{P}$ and all $P_{-i} \in \mathcal{P}^{N \setminus \{i\}}$.

Furthermore, $\hat{\varphi}$ can be chosen to satisfy anonymity.

Proof. We construct $\hat{\varphi}$. If $\hat{u}_i$ violates URBI($r$), then for some $a, b \in M$, we have
\[
\hat{r} := \frac{\hat{u}_i(b) - \min_{j \in M} \hat{u}_i(j)}{\hat{u}_i(a) - \min_{j \in M} \hat{u}_i(j)} > r.
\] (12)

Observe that $b$ cannot be $i$’s last choice because $0/(\hat{u}_i(a) - \min_{j \in M} \hat{u}_i(j)) \leq r$ is trivially satisfied. Define the mechanism $\hat{\varphi}$ by setting the assignment for the distinguished agent $i$ as follows: Fix parameters $\delta_a, \delta_b \in [0, 1/m]$; then, for any $P_{-i} \in \mathcal{P}^N$, any preference order $\hat{P}_i \in \mathcal{P}$, and any object $j \in M$, let
\[
\hat{\varphi}_{i,j}(\hat{P}_i, P_{-i}) = \begin{cases} 
1/m, & \text{if } \hat{P}_i : a > b, \\
1/m + \delta_b \mathbb{1}_{j=b} - \delta_a \mathbb{1}_{j=a} + (\delta_a - \delta_b) \mathbb{1}_{j=d}, & \text{if } \hat{P}_i : b > a > \ldots > d,
\end{cases}
\]
where $d$ is the last choice under $\hat{P}_i$. For all other agents, distribute the remaining probabilities evenly. With parameters $\delta_a, \delta_b \in [0, 1/m]$, this mechanism is well defined. Next, choose $\delta_b = 1/m$ and $\delta_a$ from the interval $[\tilde{r}/m, \bar{r}/m)$. It is a straightforward, albeit tedious, exercise to verify that $\tilde{\varphi}$ is $r$-partially strategyproof but manipulable for agent $i$ with utility function $\tilde{u}_i$ (Lemma 2 in Appendix B). To get an anonymous mechanism with the same properties, randomly assign each agent to the role of agent $i$.

In words, Proposition 1 means that for any utility function that violates URBI($r$), there exists some $r$-partially strategyproof mechanism under which truthful reporting is not a dominant strategy for an agent with that utility function. Thus, unless we are given additional structural information about the mechanism (besides the fact that it is $r$-partially strategyproof and possibly anonymous), URBI($r$) is in fact the largest set of utility functions for which truthful reporting is guaranteed to be a dominant strategy.

### 5.4. The Degree of Strategyproofness

The partial strategyproofness concept induces a new, intuitive measure for the incentive properties of swap monotonic and upper invariant mechanisms.

**Definition 4 (Degree of Strategyproofness).** Given a setting $(N, M, q)$ and a mechanism $\varphi$, the degree of strategyproofness of $\varphi$ (in the setting $(N, M, q)$) is defined as

$$\rho_{(N,M,q)}(\varphi) = \max \{ r \in [0, 1] \mid \varphi \text{ is } r\text{-partially strategyproof in } (N, M, q) \}.$$  

Observe that for $0 < r' < r \leq 1$ we have URBI($r'$) $\subset$ URBI($r$) by construction. Thus, a higher degree of strategyproofness corresponds to a stronger guarantee. By maximality of the URBI($r$) domain restriction (Proposition 1), the degree of strategyproofness constitutes a meaningful measure for incentive properties: If the only known attributes of $\varphi$ are that it is swap monotonic and upper invariant, then no single-parameter measure conveys strictly more information about the incentive properties of $\varphi$.

The degree of strategyproofness also allows for the comparison of two mechanisms: If $\rho_{(N,M,q)}(\varphi) > \rho_{(N,M,q)}(\psi)$, then $\varphi$ makes truthful reporting a dominant strategy on a strictly larger URBI($r$) set of utility functions than $\psi$ does. From a quantitative perspective, one might ask for how many more utility functions $\varphi$ is guaranteed to have.

---

18To see that $\rho_{(N,M,q)}$ is well-defined, observe that URBI($r$) is topologically closed. Thus, a mechanism that is $r'$-partially strategyproof for all $r' < r$ must also be $r$-partially strategyproof.
good incentives, compared to \( \psi \). To answer this question, recall Remark 4, where we considered utility functions in a setting with 3 objects, \( \min_{j \in M} (u_i(j)) = 0 \), and where the utilities for the first and second choice were chosen uniformly at random from the unit square. Suppose that \( \rho_{(N,M,q)}(\varphi) = 4/5 \) and \( \rho_{(N,M,q)}(\psi) = 2/5 \). Then the set \( \text{URBI}(4/5) \) has twice the size of \( \text{URBI}(2/5) \). Thus, the guarantee for \( \varphi \) extends over twice as many utility functions as the guarantee for \( \psi \).

A different approach to comparing mechanisms by their vulnerability to manipulation was proposed by Pathak and Sönmez (2013), and an extension of their concept to the cardinal case is straightforward: We say that \( \psi \) is strongly as manipulable as \( \varphi \) if, whenever an agent with utility function \( u_i \) finds a beneficial misreport under \( \varphi \), the same agent in the same situation also finds a beneficial misreport under \( \psi \). Interestingly, this comparison is consistent with the comparison of mechanisms by their degrees of strategyproofness in the sense that the two comparisons never yield contradictory results.\(^{19}\) However, despite this consistency, neither concept is always better at strictly differentiating between two mechanisms. A comparison by vulnerability to manipulation may be inconclusive when the degree of strategyproofness yields a strict winner. Conversely, the degrees of strategyproofness may be equal, even though one of the mechanisms is in fact more strongly manipulable than the other.

An important difference is that the comparison by vulnerability to manipulation considers each preference profile separately, while the partial strategyproofness constraints must hold uniformly for all preference profiles. Vulnerability to manipulation thus yields a best response notion of incentives, while the degree of strategyproofness yields a dominant strategy notion of incentives. However, the degree of strategyproofness has two important advantages. First, it is computable (see Section 7), whereas Pathak and Sönmez (2013) do not present a method to perform their comparison algorithmically, and the definition of such a method is not trivial. Second, and more importantly, the degree of strategyproofness allows a parametric comparison, whereas the comparison by vulnerability to manipulation is binary (i.e., it does not indicate how much more robust one mechanism is to strategic manipulation than another). Moreover, it is simple for mechanism designers to express a minimal acceptable degree of strategyproofness and identify mechanisms within this restricted class that have other desirable properties (and this search can be performed algorithmically).

\(^{19}\)Proposition 7 in Section 13 of the Online Appendix establishes this consistency formally.
6. Partial Strategyproofness and Other Concepts

In this section, we show that partial strategyproofness takes an intermediate position between strategyproofness and other incentive concepts that have been discussed previously in the context of the assignment problem or more broadly in domains with no monetary transfers. Figure 2 illustrates these relations.

6.1. Relation to Weak SD-Strategyproofness

Weak SD-strategyproofness requires that no agent can obtain an assignment that it strictly prefers, independent of its underlying utility function, by misreporting.

Definition 5 (Weak SD-Strategyproofness). A mechanism $\varphi$ is weakly SD-strategyproof if, for all agents $i \in N$, all preference profiles $(P_i, P_{-i}) \in \mathcal{P}^N$, and all misreports $P'_i \in \mathcal{P}$, $\varphi_i(P_i, P_{-i})$ is not strictly stochastically dominated by $\varphi_i(P'_i, P_{-i})$ at $P_i$.

Bogomolnaia and Moulin (2001) employed weak SD-strategyproofness to describe the incentive properties of the Probabilistic Serial mechanism. The next proposition establishes its relation with partial strategyproofness.

Proposition 2. Given a setting $(N, M, q)$, if a mechanism $\varphi$ is partially strategyproof (i.e., $r$-partially strategyproof for some $r > 0$), then it is weakly SD-strategyproof. The converse may not hold.
Proof. Observe that \( r \)-partial strategyproofness implies (but is not implied by) convex strategyproofness,\(^{20}\) which implies weak SD-strategyproofness. Balbuzanov (2015) gave an example to show that weak SD-strategyproof does not imply convex strategyproofness; thus, it does not imply partial strategyproofness.

Proposition 2 is interesting because, as we show in Section 8.2, the Probabilistic Serial mechanism is actually partially strategyproof. Thus, partial strategyproofness is a strictly more informative statement about the incentive properties of this mechanism than either weak SD-strategyproofness or convex strategyproofness.

6.2. Relation to Approximate Strategyproofness

The impossibility results pertaining to strategyproofness have motivated the introduction of various notions of *approximate strategyproofness*. The general idea is that non-strategyproof mechanisms should at least be *close to* strategyproof in a quantifiable sense. In domains with money, this idea translates to the requirement that the agents’ gain (measured in units of money) from misreporting should be bounded.\(^{21}\) However, utilities in the assignment domain are not comparable across agents, which makes the definition of meaningful notions of approximate strategyproofness more challenging. Earlier work has overcome this difficulty by bounding the utility functions, e.g., between 0 and 1 (Birrell and Pass, 2011; Carroll, 2013; Lee, 2015).\(^{22}\)

**Definition 6** (Approximate Strategyproofness). Given a setting \((N, M, q)\) and a bound \(\varepsilon \in [0, 1]\), a mechanism \(\varphi\) is \(\varepsilon\)-approximately strategyproof if, for all agents \(i \in N\), all preference profiles \((P_i, P_{-i}) \in \mathcal{P}^N\), all misreports \(P'_i \in \mathcal{P}\), and all utility functions \(u_i \in U_{P_i}\) that are bounded between 0 and 1, we have

\[
E_{\varphi_i(P_i, P_{-i})}[u_i] + \varepsilon \geq E_{\varphi_i(P'_i, P_{-i})}[u_i].
\]

(14)

Obviously, \(\varepsilon\)-approximate strategyproofness implies \(\varepsilon'\)-approximate strategyproofness.

\(^{20}\)A mechanism is *convex strategyproof* if its makes truthful reporting a dominant strategy for at least one utility function (Balbuzanov, 2015).

\(^{21}\)Lubin and Parkes (2012) provided a survey of applications in quasi-linear domains.

\(^{22}\)Liu and Pycia (2013) defined the notion of \(\varepsilon\)-strategy-proofness differently by imposing a bound on the violations of the stochastic dominance constraints. Proposition 8 in Section 14 of the Online Appendix shows that their notion and our notion are in fact equivalent.
for any $\varepsilon' > \varepsilon$; 0-approximate strategyproofness is equivalent to strategyproofness; and 1-approximate strategyproofness is a void constraint. In this sense, approximate strategyproofness yields a different parametric relaxation of strategyproofness, where the constraint is stronger when the bound $\varepsilon$ is smaller. The next proposition shows the relation of approximate strategyproofness with partial strategyproofness.

**Proposition 3.** Given a setting $(N, M, q)$, if a mechanism $\varphi$ is $r$-partially strategyproof for some $r > 0$, then it is $\varepsilon$-approximately strategyproof for some $\varepsilon < 1$. The converse may not hold. Moreover, for any $\varepsilon > 0$, there exists an $r < 1$ such that any $r$-partially strategyproof mechanism is $\varepsilon$-approximately strategyproof.

We give the proof in Appendix C.

Recall that $r$-partially strategyproof mechanisms are guaranteed to make truthful reporting a dominant strategy for any agent whose utility function satisfies URBI($r$), which enables us to give honest and useful strategic advice: It is in the agent’s best interest to report truthfully if the agent’s utilities for any two objects are sufficiently different. Proposition 3 now allows a refinement of this advice. Since partial strategyproofness implies approximate strategyproofness, any agent’s potential gain from misreporting is also quantitatively bounded, even if the agent’s utility function does not satisfy URBI($r$); and this quantitative bound is tighter the closer $r$ is to 1.

### 6.3. Relation to Strategyproofness in the Large

Azvedo and Budish (2015) proposed strategyproofness in the large (SP-L) as an alternative incentive concept when strategyproofness is incompatible with other essential design objectives. This concept captures the intuition that the ability of any single agent to benefit from misreporting may vanish when more agents participate. For example, in school choice, where thousands of students compete for seats at a relatively small number of schools, this requirement may facilitate interesting design alternatives.

Azvedo and Budish (2015) defined strategyproofness in the large with respect to a finite set of utility functions $\{u^1, \ldots, u^K\}$. Loosely speaking, a mechanism satisfies SP-L if, for any $\varepsilon > 0$, there exist a number $n_0$ of agents such that in any setting with at least $n_0$ agents, no agent with a utility function from $\{u^1, \ldots, u^K\}$ can improve its expected
utility by more than $\varepsilon$ through misreporting.\textsuperscript{23} To connect this concept with partial strategyproofness, we need to specify in what sense assignment problems get large. Here, we follow Kojima and Manea (2010): We keep the number of objects constant but we let the number of agents grow, and we increase the objects’ capacities such that the total capacity is not lower than the number of agents.

**Proposition 4.** Fix any finite set of utility functions $\{u^1, \ldots, u^K\}$ and consider any sequence of settings $(N^n, M^n, q^n)_{n \geq 1}$ with $\#N^n = n$, $M^n = M$, $\sum_{j \in M} q^n_j \geq n$, and $\min_{j \in M} q^n_j \to \infty$ as $n \to \infty$. If the degree of strategyproofness of $\varphi$ converges to 1 as $n \to \infty$, then $\varphi$ is strategyproof in the large (with respect to $\{u^1, \ldots, u^K\}$).

**Proof.** Any utility function $u_i \in U_{P_i}$ satisfies URBI($r$) for some (sufficiently large) $r < 1$. Thus, we can find $\bar{r}$ such that $u^k \in$ URBI($\bar{r}$) for all $k \in \{1, \ldots, K\}$. Since the degree of strategyproofness converges to 1 as $n$ grows, $\varphi$ is $\bar{r}$-partially strategyproof in any setting with sufficiently many agents. Thus, the gain from misreporting vanishes for large $n$. \hfill $\square$

Proposition 4 shows that convergence of $r$ to 1 in large settings implies SP-L. In Section 8.2, we explain how Proposition 4 might be applied to obtain an elegant, parametric proof of strategyproofness in the large of the Probabilistic Serial mechanism.

### 6.4. Relation to Lexicographic Strategyproofness

Finally, we establish the relation of partial strategyproofness with strategyproofness for agents with lexicographic preferences (Cho, 2012). An agent with lexicographic preferences prefers any (arbitrarily small) increase in the probability for some more preferred object to any (arbitrarily large) increase in the probability for some less preferred object.

**Definition 7** (Lexicographic Dominance). For a preference order $P_i \in \mathcal{P}$ and assignment vectors $x_i, y_i$, we say that $x_i$ lexicographically dominates $y_i$ at $P_i$ if either $x_i = y_i$, or $x_{i,a} > y_{i,a}$ for some $a \in M$ and $x_{i,j} = y_{i,j}$ for all $j \in U(a, P_i)$.

Lexicographic dominance (LD) induces LD-strategyproofness in the same way in which stochastic dominance induces SD-strategyproofness.

\textsuperscript{23}The original definition weakens strategyproofness further by taking an *interim perspective*, where agents are uncertain about the reports from the other agents. The simplified definition that we use yields a more demanding variant of SP-L. Observe that this only makes our Proposition 4 stronger.
**Definition 8** (LD-Strategyproofness). A mechanism $\varphi$ is **LD-strategyproof** if, for all agents $i \in N$, all preference profiles $(P_i, P_{-i}) \in \mathcal{P}^N$, and all misreports $P'_i \in \mathcal{P}$, $\varphi_i(P_i, P_{-i})$ lexicographically dominates $\varphi_i(P'_i, P_{-i})$ at $P_i$.

Theorem 3 shows the relation of partial strategyproofness and LD-strategyproofness.

**Theorem 3.** Given a setting $(N, M, q)$, a mechanism $\varphi$ is partially strategyproof (i.e., $r$-partially strategyproof for some $r > 0$) if and only if $\varphi$ is LD-strategyproof.

**Proof.** This follows from Lemma 1 in the proof of Theorem 2.

Theorem 3 has two interesting consequences: First, by describing a mechanism as being LD-strategyproof, one ignores the parametric nature of the set of utility functions for which this mechanism is guaranteed to make truthful reporting a dominant strategy. The partial strategyproofness thus concept provides a more refined understanding of the incentive properties of non-strategyproof mechanisms.

Second, Theorem 3 shows that LD-strategyproofness is the lower bounding concept for partial strategyproofness, symmetric to the sense in which strategyproofness is the upper bounding concept. To see this, let $\text{SP}(N, M, q)$, $r$-$\text{PSP}(N, M, q)$, and $\text{LD-SP}(N, M, q)$ denote the sets of mechanisms that are strategyproof, $r$-partially strategyproof, and LD-strategyproof respectively in the setting $(N, M, q)$. Observe that

$$\text{SP}(N, M, q) = \bigcap_{r<1} r$\text{-PSP}(N, M, q).$$

(15)

In words, any mechanism that is $r$-partially strategyproof for all $r < 1$ must be strategyproof. Theorem 3 implies the complementary statement that

$$\text{LD-SP}(N, M, q) = \bigcup_{r>0} r$\text{-PSP}(N, M, q).$$

(16)

In words, any mechanisms that is LD-strategyproof must be $r$-partially strategyproof for some strictly positive $r > 0$. Thus, the degree of strategyproofness $\rho_{(N, M, q)}$ parametrizes the space of incentive concepts between strategyproofness on the one hand and the weaker LD-strategyproofness on the other hand.
7. Discounted Dominance and Partial Strategyproofness

Recall that strategyproofness (defined in terms of expected utility) is equivalent to SD-strategyproofness. In this section, we present \( r \)-discounted dominance, a new dominance notion that generalizes stochastic dominance but includes \( r \) as a discount factor. We then provide an alternative definition of partial strategyproofness by showing that it is equivalent to the incentive concept induced by \( r \)-discounted dominance.

Definition 9 (Discounted Dominance). For a bound \( r \in [0, 1] \), a preference order \( P_i \in \mathcal{P} \) with \( P_i : j_1 > \ldots > j_m \), and assignment vectors \( x_i, y_i \), we say that \( x_i \) \( r \)-discounted dominates \( y_i \) at \( P_i \) if, for all ranks \( K \in \{1, \ldots, m\} \), we have

\[
\sum_{k=1}^{K} r^k \cdot x_{i,j_k} \geq \sum_{k=1}^{K} r^k \cdot y_{i,j_k}.
\] (17)

Observe that for \( r = 1 \), this is precisely the same as stochastic dominance. However, for \( r < 1 \), the difference in the agent’s assignment for the \( k \)th choice is discounted by the factor \( r^k \). Analogous to stochastic dominance for SD-strategyproofness, we can use \( r \)-discounted dominance (\( r \)-DD) to define the corresponding incentive concept.

Definition 10 (\( r \)-DD-Strategyproofness). Given a setting \((N, M, q)\) and a bound \( r \in (0, 1] \), a mechanism \( \varphi \) is \( r \)-DD-strategyproof if, for all agents \( i \in N \), all preference profiles \((P_i, P_{-i}) \in \mathcal{P}^N \), and all misreports \( P_i' \in \mathcal{P} \), \( \varphi_i(P_i, P_{-i}) \) \( r \)-discounted dominates \( \varphi_i(P_i', P_{-i}) \) at \( P_i \).

The next theorem yields the equivalence to \( r \)-partial strategyproofness.

Theorem 4. Given a setting \((N, M, q)\) and a bound \( r \in [0, 1] \), a mechanism \( \varphi \) is \( r \)-partially strategyproof if and only if it is \( r \)-DD-strategyproof.

We give the proof in Appendix D.

Theorem 4 generalizes the equivalence between strategyproofness and SD-strategyproofness. Moreover, it yields an alternative definition of \( r \)-partial strategyproofness in terms of discounted dominance. This shows that the partial strategyproofness concept integrates nicely into the landscape of existing incentive concepts, many of which are defined using dominance notions (e.g., SD-, weak SD-, LD-, and ST\( ^{24} \)-strategyproofness).

\( ^{24} \)Sure thing dominance; see (Aziz, Brandt and Brill, 2013)
Table 1: Application of partial strategyproofness.  

| Mechanism                        | Randomness       | SM | UI | LI | PSP | SP | Section |
|----------------------------------|------------------|----|----|----|-----|----|---------|
| Random Serial Dictatorship (RSD) | Intrinsic        | ✔  | ✔  | ✔  | ✔   | ✔  | 8.1     |
| Probabilistic Serial (PS)        | Intrinsic        | ✔  | ✔  | ✗  | ✔   | ✗  | 8.2     |
| Boston (random priorities)       | Priorities       | ✗  | ✔  | ✗  | ✗   | ✗  | 8.3     |
| Adaptive Boston (random priorities) | Priorities     | ✔  | ✗  | ✔  | ✔   | ✗  | 8.4     |
| Rank Value (RV)                  | Intrinsic        | ✗  | ✗  | ✔  | ✔   | ✗  | 8.5     |
| Hybrids of RSD & PS              | Intrinsic        | ✔  | ✔  | ✗  | ✔   | ✗  |         |
| Hybrids of RSD & Boston          | Intrinsic & priorities | ✗  | ✔  | ✗  | ✗   | ✗  | 8.6     |
| Hybrids of RSD & Adaptive Boston | Intrinsic & priorities | ✔  | ✗  | ✔  | ✔   | ✗  |         |
| Hybrids of RSD & RV              | Intrinsic        | ✗  | ✗  | ✗  | ✔   | ✗  |         |
| Classical Boston (fixed priorities) | Interim perspective | ✔  | ✔  | ✗  | ✔   | ✗  | 8.7     |
| Adaptive Boston (fixed priorities) | Interim perspective | ✔  | ✔  | ✗  | ✔   | ✗  |         |
| Multi-unit PS                    | Intrinsic & interim perspective | ✔  | ✔  | ✗  | ✔   | ✗  |         |
| HBS Draft                        | Intrinsic & interim perspective | ✔  | ✔  | ✗  | ✔   | ✗  |         |

Table 1: Application of partial strategyproofness.  

SM: swap monotonic; UI: upper invariant; LI: lower invariant; PSP: partially strategyproof; SP: strategyproof

The dominance interpretation also unlocks the partial strategyproofness concept to algorithmic analysis. Recall that the definition of $r$-partial strategyproofness imposes inequalities that have to hold for all utility functions within the set $URBI(r)$. This set is infinite, which makes algorithmic verification of $r$-partial strategyproofness infeasible via its original definition. However, by the equivalence from Theorem 4, it suffices to verify that all (finitely many) constraints for $r$-discounted dominance are satisfied (i.e., the inequalities (17) from Definition 9). These inequalities can also be used to encode $r$-partial strategyproofness as linear constraints to an optimization problem. This enables an automated search in the set of $r$-partially strategyproof mechanisms while optimizing for some other design objective under the automated mechanism design paradigm (Sandholm, 2003).

### 8. Applications

In this section, we apply the new partial strategyproofness concept to analyze the incentive properties of popular and new non-strategyproof mechanisms. Table 1 provides an overview of our results.
8.1. Random Serial Dictatorship

Under the *Random Serial Dictatorship* mechanism (RSD) agents line up in a random order and then take turns to pick their most preferred object that is still available. RSD is strategyproof (Abdulkadiroğlu and Sönmez, 1998), and therefore it satisfies all three axioms and is 1-partially strategyproof in any setting.

8.2. Probabilistic Serial Mechanism

The *Probabilistic Serial* mechanism (PS), introduced by Bogomolnaia and Moulin (2001), is one of the most well-studied mechanism for the random assignment problem. Our next proposition shows that PS is partially strategyproof.

**Proposition 5.** Given a setting $(N, M, q)$, PS is $r$-partially strategyproof for some $r > 0$.

**Proof.** PS is LD-strategyproof by Theorem 4.7 in (Cho, 2012). Therefore, it is partially strategyproof by our Theorem 3.

Regarding the incentive properties of PS, Bogomolnaia and Moulin (2001) showed that it is weakly SD-strategyproof (but not strategyproof), and Balbuzanov (2015) showed that it is convex strategyproof. Since partial strategyproofness implies both of these properties, Proposition 5 establishes the most demanding description of the incentive properties of PS in finite settings known to date. The fact that PS is partially strategyproof also justifies our choice of this mechanism for the introductory example.

Figure 3 shows the degrees of strategyproofness of PS in various settings. Observe that, for a fixed number of objects, growing capacities, and a growing number of agents, these values increase. This suggests the conjecture that the degree of strategyproofness of PS converges to 1 in large settings.

**Remark 6.** Kojima and Manea (2010) proved that, for an arbitrary, fixed utility function, PS makes truthful reporting a dominant strategy for agents with that utility function if there are sufficiently many copies of each object; and Azevedo and Budish (2015) used this to conclude that PS is SP-L. A proof of our conjecture that the degree of strategyproofness of PS converges to 1 in large settings would strengthen the result of...
Figure 3: Plot of $\rho_{(N,M,q)}(PS)$ for $m = 3$ (left) and $m = 4$ (right) objects, for varying numbers of agents $n$, and evenly distributed capacities $q_j = n/m$.

Kojima and Manea (2010); and, in combination with our Proposition 4, it would yield an elegant, parametric proof that PS is SP-L.$^{26}$

8.3. Boston Mechanism

Under the Boston mechanism (BM), agents apply to their first choice in the first round and objects are assigned to applicants according to some priority order. Agents who have not received their first choice in the first round apply to their second choice in the second round, and the process repeats. This mechanism is frequently used to assign students to seats at public schools, but it has been heavily criticized for its susceptibility to manipulation (Abdulkadiroğlu and Sönmez, 2003).

We consider this mechanism with priorities determined by a single, uniform lottery. Intuitively, it is upper invariant because the object to which an agent applies in the $k^{th}$ round has no effect on the applications (or assignments) in any previous rounds (see (Mennle and Seuken, 2017d) for a formal proof of upper invariance). However, BM fails swap monotonicity, as the following example shows, and thus it violates $r$-partial strategyproofness for any $r > 0$.

Example 3. Consider a setting with four agents, four objects in unit capacity, and preferences $P_1 : a > b > c > d$, $P_2 : a > c > b > d$, and $P_3, P_4 : b > c > a > d$. Agent 1’s assignment is $BM_1(P_1, P_{-1}) = (1/2, 0, 0, 1/2)$ for the objects $a, b, c, d$ respectively. If agent

$^{26}$A proof of the conjecture is presently not available. Observe, however, that it is not implied by Theorem 1 of Kojima and Manea (2010), which states precisely that for any $\varepsilon > 0$, PS is $\varepsilon$-approximately strategyproof in sufficiently large settings; this is insufficient to establish partial strategyproofness.
1 swaps $b$ and $c$ in its report, its assignment changes to $BM_1(P'_1, P_{-1}) = (1/2, 0, 1/4, 1/4)$. This violates swap monotonicity as agent 1’s assignment for $b$ has not changed, but the overall assignment has. It also violates lower invariance as agent 1’s assignment for $d$ has changed, even though $d$ is in the lower contour set of $c$.

This insight adds an axiomatic argument to the criticism of the manipulability of BM.

### 8.4. Adaptive Boston Mechanism

Some obvious opportunities for manipulation arise under BM: An agent who knows that their second choice will be exhausted in the first round can manipulate by ranking their third choice in the second position. This misreport improves that agent’s chance of obtaining their third choice without reducing their chance of obtaining their first or second choice. If instead the agent automatically skipped exhausted objects in the application process, then this manipulation heuristic would no longer be effective. In (Mennle and Seuken, 2017d) we have shown that such an adaptive Boston mechanism (ABM) with priorities determined by a single, uniform lottery is swap monotonic and upper invariant, and thus partially strategyproof. However, ABM is not fully strategyproof.

Interestingly, a comparison of BM and ABM by their vulnerability to manipulation is not differentiating for fixed priorities or inconclusive for random priorities (Dur, Mennle and Seuken, 2016). In contrast, partial strategyproofness provides a formal understanding of the nuanced differences between BM and ABM in terms of their incentive properties.

As for PS, we have computed the degree of strategyproofness of ABM for various settings.\(^{27}\) The results are shown in Figure 4. Observe that $\rho_{(N,M,q)}(ABM)$ is significantly lower than $\rho_{(N,M,q)}(PS)$ (see Figure 3). Furthermore, it does not grow for larger settings, but it appears to remain constant (i.e., $\rho_{(N,M,q)}(ABM) = 1/2$ for $m = 3$ objects and $n = 3, 6, 9$ agents; and $\rho_{(N,M,q)}(ABM) = 1/3$ for $m = 4$ objects and $n = 4, 8$ agents). Partial strategyproofness thus yields a formal argument that ABM has intermediate incentive properties, which are stronger than those of BM but weaker than those of PS.

### 8.5. Rank Efficient Mechanisms

Featherstone (2011) introduced rank efficiency, a strict refinement of ordinal efficiency.

\(^{27}\)The challenges with computing the degree of strategyproofness of ABM in larger settings are compounded by the prohibitive computational cost of determining each random assignment.
Figure 4: Plot of $\rho_{(N,M,q)}(ABM)$ for $m = 3$ (left) and $m = 4$ (right) objects, for varying numbers of agents $n$, and for $q_j = n/m$.

Rank efficiency often closely reflects the welfare criteria used in practical applications, e.g., in school choice (Mennle and Seuken, 2017d), or in the assignment of teachers to schools (Featherstone, 2011). However, no rank efficient mechanism is even weakly SD-strategyproof (Theorem 3 in (Featherstone, 2011)). Furthermore, any rank efficient mechanism will be neither swap monotonic, nor upper invariant, nor lower invariant (see Examples 4 and 5 in Section 15 of the Online Appendix); thus they will not be partially strategyproof. This highlights that the attractive efficiency properties of rank efficient mechanisms come at a price, as these mechanisms fail all of the axioms that constitute strategyproofness.

8.6. Hybrid Mechanisms

In (Mennle and Seuken, 2017b), we have investigated a natural way to trade off strategyproofness and efficiency in the random assignment problem. The main idea of hybrid mechanisms is to mix two different mechanisms, one with good incentive properties and the other with good efficiency properties. We have shown that, under certain technical conditions, the resulting mechanisms are partially strategyproof, but can also improve efficiency beyond the ex-post efficiency of Random Serial Dictatorship. Furthermore, the trade-offs are scalable in the sense that the mechanism designer can accept a lower degree of strategyproofness in exchange for more efficiency. Partial strategyproofness enables this analysis because it allows a meaningful parametric evaluation of incentive properties. Moreover, its equivalent formulation in terms of discounted dominance enables the computation of optimal mixing factors between the two mechanisms.
8.7. Deterministic and Multi-Unit Assignment Mechanisms

In the examples discussed so far, randomness was either an intrinsic feature of the mechanism (e.g., with PS) or it was induced by random priorities (e.g., with BM and ABM). A third source of randomness may arise from the fact that agents are not certain about the preference reports of the other agents. By considering the effect of this uncertainty on the agents’ strategic situation, one takes what Azevedo and Budish (2015) call the interim perspective. In Section 16 of the Online Appendix, we show formally how the notion of partial strategyproofness can be applied from this interim perspective. In particular, we show that mechanisms (random or deterministic) are partially strategyproof from the interim perspective precisely if they satisfy three axioms: upper invariance, weak monotonicity, and sensitivity. Since the combination of weak monotonicity and sensitivity is a strictly weaker requirement than swap monotonicity, this allows the application of partial strategyproofness to a larger class of mechanisms, including deterministic ones.

In Sections 8.3 and 8.4, we have already observed that partial strategyproofness captures the subtle differences in terms of incentives between BM and ABM when priorities are random. However, priorities in school choice may be known, e.g., if they are based on walk-zones, siblings, or exam scores. Nonetheless, partial strategyproofness can be applied to study the incentive properties of BM and ABM even with non-random priorities: Both mechanisms satisfy upper invariance, weak monotonicity, and sensitivity (Mennle and Seuken, 2017); they can therefore be analyzed from the interim perspective.

Multi-unit assignment is an important extension of the assignment problem, where each agent receives $K$ objects (instead of just one). For example, the HBS Draft mechanism is used to assign schedules of courses to students at Harvard Business School (Budish and Cantillon, 2012). Another example of a multi-unit assignment mechanism is the straightforward extension of the Probabilistic Serial mechanism (Heo, 2014), where agents consume a total of $K$ (instead of 1.0) probability shares. Neither the HBS Draft mechanism nor the Probabilistic Serial mechanism is strategyproof for the multi-unit assignment problem, even if agents have additive utilities. Furthermore, while both

---

28 Under the HBS Draft mechanism, students report preference orders over courses. They then take turns in some random order to pick a single course each, according to their reported preferences. For the second course, the picking order is reversed. Then the third courses are again picked in the original order, and so on.
mechanisms are upper invariant and involve randomization, neither is swap monotonic. However, they satisfy weak monotonicity and sensitivity (see Proposition 9 in Section 16.3 of the Online Appendix). Thus, partial strategyproofness can be applied to study their incentive properties from the interim perspective.

9. Conclusion

We have introduced partial strategyproofness, a new concept to understand the incentive properties of non-strategyproof assignment mechanisms that involve some form of randomness; and we have argued why this concept is an important addition to the mechanism designer’s toolbox. The definition of partial strategyproofness is motivated by the observation that incentives for misreporting can be attributed to small differences in the agents’ preference intensities. Interestingly, this definition also has an appealing axiomatic motivation, and yields a meaningful parametric measure to study and compare the incentive properties of non-strategyproof mechanisms. In addition, partial strategyproofness comes with an array of useful technical results: First, locality of the axioms swap monotonicity and upper invariance makes them intuitive and easily accessible. Second, its dominance interpretation enables the algorithmic verification of partial strategyproofness and the computation of the degree of strategyproofness (for any mechanism in any finite setting). Third, partial strategyproofness provides a unified view on many other relaxed incentive requirements. Fourth, it allows market designers to give honest and useful strategic advice: If an agent differentiates sufficiently between any two objects, then the agent is best off by reporting their preferences truthfully; otherwise, the agent’s incentive to misreport may be positive, but it is still quantitatively bounded. These facts distinguish the partial strategyproofness concept as a natural and powerful relaxed notion of strategyproofness for the assignment problem.

When applied to concrete mechanisms, the concept has provided new and useful insights, such as a more refined understanding of the incentive properties of the Probabilistic Serial mechanism and the meaningful comparison of the two variants of the Boston mechanism. In addition, it is applicable across a broad spectrum of other mechanisms. In the future, we expect the partial strategyproofness concept to be useful in research on assignment mechanisms. For example, from a theoretical perspective, it enables the development of new mechanisms with appealing properties which are incompatible with
full strategyproofness; and for practical market design, it opens up new ways of reasoning about the trade-offs between incentives for truth-telling and other desiderata in school choice settings.

References

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 1998. “Random Serial Dictatorship and the Core from Random Endowments in House Allocation Problems.” *Econometrica*, 66(3): 689–702.

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 2003. “School Choice: A Mechanism Design Approach.” *American Economic Review*, 93(3): 729–747.

Abdulkadiroğlu, Atila, Parag Pathak, and Alvin Roth. 2009. “Strategy-Proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match.” *American Economic Review*, 99(5): 1954–78.

Azevedo, Eduardo, and Eric Budish. 2015. “Strategy-proofness in the Large.” Working Paper.

Aziz, Haris, Felix Brandt, and Markus Brill. 2013. “On the Tradeoff Between Economic Efficiency and Strategy Proofness in Randomized Social Choice.” In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*.

Bade, Sophie. 2016. “Random Serial Dictatorship: The One and Only.” Working Paper.

Balbuzanov, Ivan. 2015. “Convex Strategyproofness with an Application to the Probabilistic Serial Mechanism.” *Social Choice and Welfare*, 1–10.

Basteck, Christian, Katharina Huesmann, and Heinrich Nax. 2015. “Matching Practices for Secondary Schools - Germany.” matching-in-practice.eu, accessed May 16, 2015.

Birrell, Eleanor, and Rafael Pass. 2011. “Approximately Strategy-Proof Voting.” In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*.

Bogomolnaia, Anna, and Hervé Moulin. 2001. “A New Solution to the Random Assignment Problem.” *Journal of Economic Theory*, 100(2): 295–328.
Budish, Eric. 2012. “Matching “versus” Mechanism Design.” *ACM SIGecom Exchanges*, 11(2): 4–15.

Budish, Eric, and Estelle Cantillon. 2012. “The Multi-Unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard.” *American Economic Review*, 102(5): 2237–2271.

Budish, Eric, Yeon-Koo Che, Fuhito Kojima, and Paul Milgrom. 2013. “Designing Random Allocation Mechanisms: Theory and Applications.” *American Economic Review*, 103(2): 585–623.

Calsamiglia, Caterina, and Maia Güell. 2014. “The Illusion of School Choice: Empirical Evidence from Barcelona.” Working Paper.

Carroll, Gabriel. 2012. “When Are Local Incentive Constraints Sufficient?” *Econometrica*, 80(2): 661–686.

Carroll, Garbiel. 2013. “A Quantitative Approach to Incentives: Application to Voting Rules.” Working Paper.

Cho, Wonki. 2012. “Probabilistic Assignment: A Two-fold Axiomatic Approach.” Working Paper.

Dur, Umut, Timo Mennle, and Sven Seuken. 2016. “First Choice Maximizing School Choice Mechanisms.” Working Paper.

Erdil, Aytek. 2014. “Strategy-proof Stochastic Assignment.” *Journal of Economic Theory*, 151: 146–162.

Erdil, Aytek, and Haluk Ergin. 2008. “What’s the Matter with Tie-Breaking? Improving Efficiency in School Choice.” *American Economic Review*, 98(3): 669–89.

Featherstone, Clayton. 2011. “A Rank-based Refinement of Ordinal Efficiency and a new (but Familiar) Class of Ordinal Assignment Mechanisms.” Working Paper.

Gibbard, Allan. 1977. “Manipulation of Schemes That Mix Voting with Chance.” *Econometrica*, 45(3): 665–81.

Hashimoto, Tadashi, Daisuke Hirata, Onur Kesten, Morimitsu Kurino, and Utku Ünver. 2014. “Two Axiomatic Approaches to the Probabilistic Serial Mechanism.” *Theoretical Economics*, 9(1): 253–277.
Heo, Eun. 2014. “Probabilistic Assignment Problem with Multi-Unit Demands: A Generalization of the Serial Rule and its Characterization.” *Journal of Mathematical Economics*, 54: 40–47.

Hylland, Aanund, and Richard Zeckhauser. 1979. “The Efficient Allocation of Individuals to Positions.” *The Journal of Political Economy*, 87(2): 293–314.

Kojima, Fuhito, and Mihai Manea. 2010. “Incentives in the Probabilistic Serial Mechanism.” *Journal of Economic Theory*, 145(1): 106–123.

Kojima, Fuhito, and Utku Ünver. 2014. “The Boston School-Choice Mechanism: an Axiomatic Approach.” *Economic Theory*, 55(3): 515–544.

Lee, David. 2015. “Efficient, Private, and epsilon-Strategyproof Elicitation of Tournament Voting Rules.” In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*.

Lee, Thiam, and Jay Sethuraman. 2011. “Equivalence Results in the Allocation of Indivisible Objects: A Unified View.” Working Paper.

Li, Shengwu. 2016. “Obviously Strategy-Proof Mechanisms.” Working Paper.

Liu, Qingmin, and Marek Pycia. 2013. “Ordinal Efficiency, Fairness, and Incentives in Large Markets.” Working Paper.

Lubin, Benjamin, and David Parkes. 2012. “Approximate Strategyproofness.” *Current Science*, 103(9): 1021–1032.

Mennle, Timo, and Sven Seuken. 2014. “An Axiomatic Approach to Characterizing and Relaxing Strategyproofness of One-sided Matching Mechanisms [Extended Abstract].” In *Proceedings of the 15th ACM Conference on Economics and Computation (EC)*.

Mennle, Timo, and Sven Seuken. 2017a. “An Axiomatic Decomposition of Strategyproofness for Ordinal Mechanism with Indifferences.” Research Note.

Mennle, Timo, and Sven Seuken. 2017b. “Hybrid Mechanisms: Trading Off Strategyproofness and Efficiency of Random Assignment Mechanisms.” Working Paper.

Mennle, Timo, and Sven Seuken. 2017c. “Local Sufficiency for Partial Strategyproofness.” Research Note.
Mennle, Timo, and Sven Seuken. 2017d. “Trade-offs in School Choice: Comparing Deferred Acceptance, the Naïve and the Adaptive Boston Mechanism.” Working Paper.

Mennle, Timo, and Sven Seuken. 2017e. “Two New Impossibility Results for the Random Assignment Problem.” Research Note.

Pathak, Parag, and Tayfun Sönmez. 2008. “Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism.” American Economic Review, 98(4): 1636–52.

Pathak, Parag, and Tayfun Sönmez. 2013. “School Admissions Reform in Chicago and England: Comparing Mechanisms by Their Vulnerability to Manipulation.” American Economic Review, 103(1): 80–106.

Roth, Alvin. 1984. “The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory.” Journal of Political Economy, 92(6): 991–1016.

Sandholm, Tuomas. 2003. “Automated Mechanism Design: A New Application Area for Search Algorithms.” In Proceedings of the International Conference on Principles and Practice of Constraint Programming (CP).

Zhou, Lin. 1990. “On a Conjecture by Gale about One-sided Matching Problems.” Journal of Economic Theory, 52(1): 123–135.

Appendix

A. Proof of Lemma 1

Proof of Lemma 1. Given a setting \((N, M, q)\) and a mechanism \(\varphi\), the following are equivalent:

a. For all agents \(i \in N\), all preference profiles \((P_i, P_{-i}) \in \mathcal{P}^N\), and all misreports \(P'_i \in \mathcal{P}\) with \(\varphi_i(P_i, P_{-i}) \neq \varphi_i(P'_i, P_{-i})\), there exists an object \(a \in M\) such that \(\varphi_{i,a}(P_i, P_{-i}) > \varphi_{i,a}(P'_i, P_{-i})\) and \(\varphi_{i,j}(P_i, P_{-i}) = \varphi_{i,j}(P'_i, P_{-i})\) for all \(j \in U(a, P_i)\),

b. \(\varphi\) is swap monotonic and upper invariant.
\textbf{a} \implies \textbf{b}. First, assume that \( \varphi \) satisfies \textbf{a} but violates upper invariance. Then there exist an agent \( i \in N \), a preference profile \((P_i, P_{-i}) \in \mathcal{P}^N \), and a misreport \( P'_i \in N_{P_i} \) with \( P_i : a > b \) but \( P'_i : b > a \), such that \( i \)'s assignment for some \( j \in U(a, P_i) \) changes. Let \( j \) be the highest-ranking such object, then \textbf{a} implies \( \varphi_{ij}(P_i, P_{-i}) > \varphi_{ij}(P'_i, P_{-i}) \). Moreover, \textbf{a} imposes restrictions in the opposite direction when \( i \) changes its report from \( P'_i \) to \( P_i \).

Second, assume towards contradiction that \( \varphi \) satisfies \textbf{a}, but violates swap monotonicity. Again, there exist \( i \), \((P_i, P_{-i}) \), and \( P'_i \) as above such that the change in \( i \)'s assignment violates swap monotonicity. This violation can occur in four ways:

1. \( i \)'s assignment may not change for \( a \) and \( b \), but for some other object. Then, by upper invariance, this object must be some \( j \in L(b, P_i) \). We can choose \( j \) to be the highest-ranking such object and derive a contradiction by a two-fold application of \textbf{a} as in the proof of upper invariance.

2. \( i \)'s assignment may change for \( a \) but not for \( b \). Then, by upper invariance, \( a \) is the highest-ranking object for which \( i \)'s assignment changes. The contradiction again arises by a two-fold application of \textbf{a}.

3. \( i \)'s assignment may change for \( b \) but not for \( a \). This case is analogous to case 2, where the roles of \( P_i \) and \( P'_i \) are reversed.

4. \( i \)'s assignment may change for both \( a \) and \( b \), but in the wrong direction. If \( i \)'s assignment for \( a \) increases, then \( a \) is the highest-ranking object for which \( i \)'s assignment changes. This leads to a direct contradiction with \textbf{a}. If \( i \)'s assignment for \( b \) decreases, we obtain a contradiction by reversing the roles of \( P_i \) and \( P'_i \).

\textbf{b} \implies \textbf{a}. For this proof we introduce the notion of transitions between preference profiles. For any two preference orders \( P_i, P'_i \in \mathcal{P} \) a transition from \( P'_i \) to \( P_i \) is a finite sequence of preference orders \( P^0, \ldots, P^K \) such that

- \( P^0 = P'_i \) and \( P^K = P_i \),
- for all \( k \in \{0, \ldots, K - 1\} \) we have \( P^k \in N_{P_{k+1}} \).

Intuitively, such a transition resembles a series of consecutive swaps that transform the preference order \( P'_i \) into the preference order \( P_i \). The canonical transition (from \( P'_i \) to \( P_i \))
is a particular transition that is inspired by the bubble-sort algorithm: Initially, we set $P^0 = P_i'$. The preference orders $P^1, \ldots, P^K$ are constructed in phases. In the first phase, we identify the highest ranking object under $P_i$ that is not ranked in the same position under $P_i'$, say $j$. Then we construct the preference orders $P^1, P^2, \ldots$ by swapping $j$ with the respective next-more preferred objects. When $j$ has reached the same position under $P_k$ as under $P_i$, the first phase ends. Likewise, at the beginning of the second phase, we identify the object that is ranked highest under $P_i$ of those objects that are not ranked in the same positions under $P_k$; and new preference orders are constructed by swapping this object up to its final position under $P_i$. Subsequent phases are analogous. The construction ends when the last created preference order $P^K$ coincides with $P_i$.

Consider a fixed setting $(N, M, q)$ and a mechanism $\varphi$ that is swap monotonic and upper invariant. Given an agent $i \in N$, a preference profile $(P_i, P_i') \in \mathcal{P}^N$, and a misreport $P_i' \in \mathcal{P}$, we must verify statement a in Lemma 1: If $j$ is the highest-ranking object under $P_i$ for which $i$’s assignment changes, then $i$’s assignment for $j$ should decrease strictly if $i$ reports $P_i'$ instead of $P_i$. Let

$$P_i : j_1 > \ldots > j_k > j_{k+1} > \ldots > j_m, \quad (18)$$
$$P_i' : j_1 > \ldots > j_k > j_{k+1}' > \ldots > j_m', \quad (19)$$

where $P_i$ and $P_i'$ coincide on the top-$k$ choices but $j_{k+1} \neq j_{k+1}'$, and consider the canonical transition from $P_i'$ to $P_i$. In the first phase, the object $j_{k+1}$ is swapped upward in $P_i'$ until it reaches the $(k + 1)^{st}$ position. By upper invariance, $i$’s assignment for the top-$k$ choices $j_1, \ldots, j_k$ does not change during this phase. Moreover, by swap monotonicity, $i$’s assignment for $j_{k+1}$ cannot decrease, and, again by upper invariance, it does not change for the remainder of the canonical transition.

Suppose that $i$’s assignment for $j_{k+1}$ changes strictly during the first phase. Then a change of report from $P_i$ to $P_i'$ must lead to a strict decrease in $i$’s assignment for $j_{k+1}$, and $i$’s assignment for any of the more preferred objects $j_1, \ldots, j_k$ does not change. But this is precisely the statement a for the combination $i, (P_i, P_i'), P_i'$. Alternatively, suppose that $i$’s assignment for $j_{k+1}$ does not change during the first phase. Then swap monotonicity implies that $i$’s assignment cannot change at all during this phase. In this case, we proceed to the second phase, where the arguments are analogous. Finally, if $i$’s assignment does not change in any phase, then it does not change at all, which would also be consistent with statement a.

\[\square\]
B. Proof of Lemma 2 for Proposition 1

Lemma 2. The mechanism $\tilde{\varphi}$ constructed in the proof of Proposition 1 is $r$-partially strategyproof and manipulable for agent $i$ with utility function $\tilde{u}_i$.

Proof. Observe that under $\tilde{\varphi}$, $i$’s assignment is independent of the other agents’ preference reports, so the mechanism is strategyproof for all other agents except $i$. Thus, we can guarantee $r$-partial strategyproofness of $\tilde{\varphi}$ by verifying that truthful reporting maximizes $i$’s expected utility for any preference order $P_i$ and any utility function $u_i \in U_{P_i}$ that satisfies URBI($r$).

If $P_i : a > b$, then $i$’s assignment remains unchanged, unless $i$ claims to prefer $b$ to $a$ (i.e., $\hat{P}_i : b > a$). If $a$ is $i$’s reported last choice under $\hat{P}_i$, then the change in $i$’s expected utility from the misreport is $-\delta_a \cdot u_i(a) + \delta_b \cdot u_i(b) \leq 0$. If $d \neq a$ is $i$’s reported last choice, then this change is $-\delta_a \cdot u_i(a) + \delta_b \cdot u_i(b) + (\delta_a - \delta_b) \cdot u_i(d)$. Since $\delta_a < \delta_b$ by assumption, this value strictly increases if $i$ ranks its true last choice $d$ last. Thus, $i$’s gain from misreporting is upper bounded by

$$-\delta_a \cdot u_i(a) + \delta_b \cdot u_i(b) + (\delta_a - \delta_b) \cdot u_i(d)$$

$$= -\delta_a \cdot \left( u_i(a) - \min_{j \in M} u_i(j) \right) + \delta_b \cdot \left( u_i(b) - \min_{j \in M} u_i(j) \right).$$

This bound is guaranteed to be weakly negative if $\delta_a \geq r \cdot \delta_b$, where we use the fact that $u_i$ satisfies URBI($r$).

Next, if $P_i : b > a$, then $i$’s assignment from truthful reporting stochastically dominates any assignment that $i$ can obtain by misreporting.

Finally, we need to verify that $i$ finds a beneficial misreport if its utility function is $\tilde{u}_i$. Let $P'_i$ be the same preference order as $\hat{P}_i$, except that $a$ and $b$ trade positions. The change in $i$’s expected utility from reporting $P'_i$ is

$$-\delta_a \cdot \tilde{u}_i(a) + \delta_b \cdot \tilde{u}_i(b) + (\delta_a - \delta_b) \cdot \tilde{u}_i(d)$$

$$= -\delta_a \cdot \left( \tilde{u}_i(a) - \min_{j \in M} \tilde{u}_i(j) \right) + \delta_b \cdot \left( \tilde{u}_i(b) - \min_{j \in M} \tilde{u}_i(j) \right),$$

where $d \neq b$ is $i$’s true last choice. This change is strictly positive if $\delta_a < \tilde{r} \cdot \delta_b$. $\square$
C. Proof of Proposition 3

Proof of Proposition 3. Given a setting \((N, M, q)\), if a mechanism \(\varphi\) is \(r\)-partially strategyproof for some \(r > 0\), then it is \(\varepsilon\)-approximately strategyproof for some \(\varepsilon < 1\). The converse may not hold. Moreover, for any \(\varepsilon > 0\), there exists an \(r < 1\) such that any \(r\)-partially strategyproof mechanism is \(\varepsilon\)-approximately strategyproof.

Let \(\varphi\) be \(r\)-partially strategyproof for some \(r > 0\) and let \(\delta\) be the smallest non-zero variation in the assignment from any change of report by any agent (see (6) in the proof of Theorem 2). We need to find \(\varepsilon < 1\) such that for all agents \(i \in N\) with a utility function \(u_i\) that is bounded between 0 and 1, all preference profiles \((P_i, P_{-i}) \in \mathcal{P}^N\), and all misreports \(P_i' \in \mathcal{P}\), \(i\)'s gain from misreporting \(P_i'\) (instead of reporting \(P_i\) truthfully) is at most \(\varepsilon\). Let \(a, b, d \in M\) be objects such that \(P_i : \ldots > a > b > \ldots > d\), where \(a \in M\) is the most-preferred object (under \(P_i\)) for which \(i\)'s assignment changes, \(b\) is the object that \(i\) ranks just below \(a\), and \(d\) is \(i\)'s least-preferred object. Lemma 1 yields \(\varphi_{i,a}(P_i, P_{-i}) - \varphi_{i,a}(P_i', P_{-i}) > \delta\). Thus, \(i\)'s gain is upper bounded by

\[
- u_i(a) \cdot \delta + u_i(b) - u_i(d) \cdot (1 - \delta) \leq 1 - \delta. \tag{24}
\]

Since \(\delta > 0\), \(\varphi\) is \(\varepsilon\)-approximately strategyproof for all \(\varepsilon \geq 1 - \delta\).

It is simple to see that the converse does not hold: Any mechanism that is almost constant (i.e., only responds to misreports by small changes to the assignments) but manipulable in a stochastic dominance sense can serve as a counter-example.

To see the second statement, fix an arbitrary \(\varepsilon > 0\). Theorem 4 and \(r\)-partial strategyproofness of \(\varphi\) imply that for all \(i \in N, (P_i, P_{-i}) \in \mathcal{P}^N\) (where \(P_i : j_1 > \ldots > j_m\), \(P_i' \in \mathcal{P}\), and \(K \in \{1, \ldots, m\}\), we have

\[
\sum_{k=1}^{K} r^k \cdot \varphi_{i,j_k}(P_i, P_{-i}) \geq \sum_{k=1}^{K} r^k \cdot \varphi_{i,j_k}(P_i', P_{-i}). \tag{25}
\]
With \( r > 0 \), we obtain

\[
\sum_{k=1}^{K} \varphi_{i,j,k}(P'_i, P_{-i}) - \varphi_{i,j,k}(P_i, P_{-i}) \leq \sum_{k=1}^{K} (1 - r^k) \cdot (\varphi_{i,j,k}(P'_i, P_{-i}) - \varphi_{i,j,k}(P_i, P_{-i})) \\
\leq \sum_{k=1}^{K} (1 - r^k) \leq \sum_{k=1}^{m} (1 - r^k). \tag{26}
\]

The last term in (26) is arbitrarily small for \( r \) sufficiently close to 1. With this, Proposition 8 in Section 14 of the Online Appendix implies \( \varepsilon \)-approximate strategyproofness of \( \varphi \). \( \square \)

D. Proof of Theorem 4

**Proof of Theorem 4.** Given a setting \((N, M, q)\) and a bound \( r \in [0, 1] \), a mechanism \( \varphi \) is \( r \)-partially strategyproof if and only if it is \( r \)-DD-strategyproof.

Given the setting \((N, M, q)\), we fix an agent \( i \in N \), a preference profile \((P_i, P_{-i}) \in \mathcal{P}^N\), and a misreports \( P'_i \in \mathcal{P} \). The following claim establishes equivalence of the \( r \)-partial strategyproofness constraints and the \( r \)-DD-strategyproofness constraints for any such combination \((i, (P_i, P_{-i}), P'_i)\) with \( x = \varphi_i(P_i, P_{-i}) \) and \( y = \varphi_i(P'_i, P_{-i}) \).

**Claim 1.** Given a setting \((N, M, q)\), a preference order \( P_i \in \mathcal{P} \), assignment vectors \( x, y \), and a bound \( r \in [0, 1] \), the following are equivalent:

a. For all utility functions \( u_i \in U_{P_i} \cap \text{URBI}(r) \) we have \( \sum_{j \in M} u_i(j) \cdot x_j \geq \sum_{j \in M} u_i(j) \cdot y_j \).

b. \( x_i \) \( r \)-discounted dominates \( y_i \) at \( P_i \).

**Proof of Claim 1.** \textbf{b} \( \Rightarrow \) \textbf{a}. Let \( P_i : j_1 > \ldots > j_m \). Assume towards contradiction that \textbf{b} holds, but that for some utility function \( u_i \in U_{P_i} \cap \text{URBI}(r) \), we have

\[
\sum_{l=1}^{m} u_i(j_l) \cdot (x_{j_l} - y_{j_l}) < 0. \tag{27}
\]

Without loss of generality, we can assume \( \min_{j \in M} u_i(j) = 0 \). Let \( \delta_k = x_{j_k} - y_{j_k} \) and let

\[
S(K) = \sum_{k=1}^{K} u_i(j_k) \cdot (x_{j_k} - y_{j_k}) = \sum_{k=1}^{K} u_i(j_k) \cdot \delta_k. \tag{28}
\]
By assumption, $S(m) = S(m - 1) < 0$ (see (27)), so there exists a smallest value $K' \in \{1, \ldots, m - 1\}$ such that $S(K') < 0$, but $S(k) \geq 0$ for all $k < K'$. Using Horner’s method, we rewrite the partial sum and get

$$S(K') = \sum_{k=1}^{K'} u_i(j_k) \cdot \delta_k = \left( \frac{S(K' - 1)}{u_i(j_{K'})} + \delta_{K'} \right) \cdot u_i(j_{K'}) \tag{29}$$

$$= \left( \left( \frac{S(K' - 2)}{u_i(j_{K' - 1})} + \delta_{K' - 1} \right) \cdot \frac{u_i(j_{K' - 1})}{u_i(j_{K'})} + \delta_{K'} \right) \cdot u_i(j_{K'}) \tag{30}$$

$$= \left( \left( \frac{\delta_1}{r} \cdot \frac{u_i(j_1)}{u_i(j_2)} + \delta_2 \right) \cdot \frac{u_i(j_2)}{u_i(j_3)} + \ldots \right) \cdot \frac{u_i(j_{K' - 1})}{u_i(j_{K'})} + \delta_{K'} \right) \cdot u_i(j_{K'}). \tag{31}$$

Since $u_i$ satisfies URBI($r$), the fraction $\frac{u_i(a_{K' - 1})}{u_i(j_{K'})}$ is bounded from below by $1/r$. But since $S(K' - 1) \geq 0$ and $u_i(j_{K' - 1}) > 0$, we must have that

$$\left( \frac{S(K' - 2)}{u_i(j_{K' - 1})} + \delta_{K' - 1} \right) = \frac{S(K' - 1)}{u_i(j_{K' - 1})} \geq 0. \tag{32}$$

Therefore, when replacing $\frac{u_i(j_{K' - 1})}{u_i(j_{K'})}$ by $1/r$ in (31) we only make the term smaller. By the same argument, we can successively replace all the terms $\frac{u_i(j_{K' - 1})}{u_i(j_{K'})}$ and obtain

$$0 > S(K') \geq \left( \left( \frac{\delta_1}{r} + \frac{1}{r} + \ldots \right) \cdot \frac{1}{r} + \delta_{K'} \right) \cdot u_i(j_{K'}) = \frac{u_i(j_{K'})}{r^{K'}} \cdot \sum_{k=1}^{K'} r^k \cdot \delta_k. \tag{33}$$

This contradicts $r$-discounted dominance of $x$ over $y$ at $P_i$, since

$$\sum_{k=1}^{K'} r^k \cdot \delta_k = \sum_{k=1}^{K'} r^k \cdot (x_{j_k} - y_{j_k}) \geq 0. \tag{34}$$

$a \Rightarrow b$. Let $P_i : j_1 > \ldots > j_m$. Assume towards contradiction that $a$ holds, but $x$ does not $r$-discounted dominate $y$ at $P_i$, i.e., for some $K \in \{1, \ldots, m\}$, we have

$$\sum_{k=1}^{K} r^k \cdot x_{j_k} < \sum_{k=1}^{K} r^k \cdot y_{j_k}, \tag{35}$$

43
and let $K$ be the smallest rank for which inequality (35) is strict. Then the value

$$
\Delta = \sum_{k=1}^{K} r^k \cdot (y_{jk} - x_{jk}),
$$

(36)
is strictly positive. Let $D \geq d > 0$, and let $u_i$ be the utility function defined by

$$
u_i(j_k) = \begin{cases} Dr^k, & \text{if } k \leq K, \\ dr^k, & \text{if } K + 1 \leq k \leq m - 1, \\ 0, & k = m. \end{cases}
$$

(37)

This utility function satisfies URBI($r$). Furthermore, we have

$$
\sum_{j \in M} u_i(j) \cdot x_j - \sum_{j \in M} u_i(j) \cdot y_j = \sum_{l=1}^{m} u(a_l) \cdot (x_{i,a_l} - y_{i,a_l})
$$

(38)

$$
= D \cdot \sum_{k=1}^{K} r^k \cdot (x_{jk} - y_{jk}) + d \cdot \sum_{k=K+1}^{m-1} r^k \cdot (x_{jk} - y_{jk})
\leq -D \cdot \Delta + d.
$$

(39)

For $d < D \cdot \Delta$, $\sum_{j \in M} u_i(j) \cdot x_j - \sum_{j \in M} u_i(j) \cdot y_j$ is strictly negative, a contradiction. $\square$

This concludes the proof of Theorem 4. $\square$

