Partial Strategyproofness: An Axiomatic Approach to Relaxing Strategyproofness for Assignment Mechanisms*

Timo Mennle  Sven Seuken
University of Zurich  University of Zurich

First version: January 13, 2014
This version: October 1, 2015

Abstract
We present partial strategyproofness, a new, relaxed notion of strategyproofness, to study the incentive properties of non-strategyproof assignment mechanisms. Under partially strategyproof mechanisms, truthful reporting is a dominant strategy for agents who have sufficiently different valuations for different objects. A single numerical parameter, the degree of strategyproofness, controls the extent to which agents valuations must differ. We show that the partial strategyproofness concept is axiomatically motivated, the underlying domain restriction is maximal, it can alternatively be defined in terms of partial dominance, it is intermediate between strategyproofness and many relaxed incentive concepts, and it unifies two prior local sufficiency results. Partial strategyproofness enables the comparison of important non-strategyproof mechanisms such as Probabilistic Serial, two variants of the Boston mechanism, the HBS Draft mechanism, and new hybrid mechanisms.

Keywords: Assignment, Matching, Strategyproofness, Partial Strategyproofness, Local Sufficiency, Stochastic Dominance, Probabilistic Serial, Boston Mechanism, HBS Draft

JEL: C78, D47, D82

*Department of Informatics, University of Zurich, Switzerland, Email: {mennle, seuken}@ifi.uzh.ch. We would like to thank (in alphabetical order) Atila Abdulkadiroğlu, Haris Aziz, Ivan Balbuzanov, Craig Boutilier, 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 Schuldenzucker, Arunava Sen, Tayfun Sonmez, 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, COST COMSOC Summer School ’13, UECE Lisbon Meetings ’13, CIREQ Matching ’14, Microeconomics Seminar (Uni Cologne), EC’14, SSCW’14, 25th Jerusalem School on Economic Theory, IWGTS’14, ESWC’15. Any errors remain our own. A 2-page abstract based on this paper has been published in the conference proceedings of EC’14 (Mennle and Seuken, 2014a). Part of this research was supported by research grants from the Hasler Foundation and from the Swiss National Science Foundation.
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. Since the seminal paper of Hylland and Zeckhauser (1979), the assignment problem has attracted much attention from mechanism designers (e.g., Abdulkadirouğlu and Sönmez (1998); Bogomolnaia and Moulin (2001); Abdulkadirouğlu and Sönmez (2003)). In practice, such problems often arise in situations that are of great importance to peoples’ lives. For example, we must assign seats at public schools, positions in training programs, or accommodation in subsidized housing.

As mechanism designers, we care specifically about efficiency, fairness, and strategyproofness. Strategyproofness is the “gold standard” among incentive concepts. However, it is also often in conflict with other design objectives: Zhou (1990) showed that, unfortunately, it is impossible to achieve the optimum on all three dimensions simultaneously, which makes the assignment problem an interesting mechanism design challenge. The Random Serial Dictatorship mechanism is strategyproof and anonymous, but only ex-post efficient. In fact, it is conjectured to be the unique mechanism that satisfies all three properties (Lee and Sethuraman, 2011; Bade, 2014). The more demanding ordinal efficiency is achieved by the Probabilistic Serial mechanism, but any mechanism that guarantees ordinal efficiency and symmetry will not be strategyproof (Bogomolnaia and Moulin, 2001). Finally, rank efficiency, an even stronger efficiency concept, can be achieved via Rank Value mechanisms (Featherstone, 2011), but it is incompatible with strategyproofness, even without additional fairness requirements. The fact that strategyproofness is in conflict with many desirable design objectives explains why market designers are interested in studying non-strategyproof mechanisms: to understand how to make useful trade-offs between different design objectives.

In practice, non-strategyproof mechanisms are ubiquitous, such as the Boston mechanism for the assignment of seats at public schools. It is frequently an explicit objective of administrators to assign as many students as possible to their top-1 or top-3 choices (Basteck, Huesmann and Nax, 2015). 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 by strategic students. 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. While this mechanism was manipulable, the organizers were confident that the majority of preferences were reported truthfully, because participants lacked the information that is necessary to determine beneficial misreports (Featherstone, 2011). For the allocation of courses at Harvard Business School, Budish and Cantillon (2012) demonstrated that the strategyproof Random Serial Dictatorship mechanism would lead to very unbalanced outcomes, and that the non-strategyproof HBS Draft mechanism yields preferable results, despite its manipulability.

The incompatibility of strategyproofness with other design objectives, such as ordinal or rank efficiency, and the fact that the mechanisms used in practice are frequently not strategyproof, shows the need to study non-strategyproof mechanisms; and researchers have been calling for useful relaxed notions of strategyproofness for this purpose (e.g., Azevedo and Budish (2015); Budish (2012)). In this paper, we take an axiomatic approach to this research question and present partial strategyproofness, a relaxed notion of strategyproofness that exploits the structure of the assignment problem.

1.1. A Motivating Example

To obtain an intuition about the partial strategyproofness concept, consider a setting in which agents 1, 2, 3 compete for objects a, b, c, and their preferences are

\[
\begin{align*}
P_1 & : a > b > c, \\
P_2 & : b > a > c, \\
P_3 & : b > c > a.
\end{align*}
\]

Suppose that the non-strategyproof Probabilistic Serial mechanism (Bogomolnaia and Moulin, 2001) is used to assign the objects and that agents 2 and 3 report truthfully. By reporting \(P_1\) truthfully, agent 1 receives \(a, b, c\) with probabilities \(3/4, 0, 1/4\), respectively. If instead agent 1 reports \(P'_1: b > a > c\), these probabilities change to \(1/2, 1/3, 1/6\). Now suppose that agent 1 has value 0 for its last choice \(c\) and higher values for the objects \(a\) and \(b\). Whether or not the misreport \(P'_1\) increases agent 1’s expected utility depends on

---

1Probabilistic Serial uses the simultaneous eating algorithm, see Appendix A for details.
its relative value for \( a \) over \( b \): if \( u_1(a) \) is close to \( u_1(b) \), then agent 1 will find it beneficial to report \( P'_1 \). If \( u_1(a) \) is significantly larger than \( u_1(b) \), then agent 1 will want to report truthfully. Precisely, the manipulation is not beneficial if \( \left( \frac{3}{4} - \frac{1}{2} \right) u_1(a) \geq \left( \frac{1}{3} - 0 \right) u_1(b) \), or equivalently, if \( \frac{3}{4} u_1(a) \geq u_1(b) \). We observe that agent 1’s incentive to manipulate hinges on its “degree of indifference” between objects \( a \) and \( b \): the closer agent 1 is to being indifferent, the higher the incentive to misreport.

Partial strategyproofness captures this intuition by providing a parameter \( r \). This value controls how close to indifferent agents may between different objects, but still have a dominant strategy to report their preferences truthfully. In the above example, \( r = 3/4 \) is the pivotal degree of indifference between \( a \) and \( b \) for agent 1. In fact, we will later see that PS is 3/4-

1.2. One Familiar and Two New Axioms

In this paper, we first provide a decomposition of strategyproofness into three axioms to then arrive at our new partial strategyproofness concept by relaxing one of the axioms. To understand the three axioms, suppose an agent considers swapping two consecutive objects in its report, e.g., as agent 1 in the above example, from \( P_1 : a > b > c \) to \( P'_1 : b > a > c \). Our axioms restrict the way in which the mechanism can react to this kind of change of report. The first axiom, swap monotonicity, requires that either the agent’s assignment remains unchanged, or its probability for \( b \) must strictly increase and its probability for \( a \) must strictly decrease. This means that the mechanism is responsive to the agent’s ranking of \( a \) and \( b \) and that the swap affects at least the objects \( a \) and \( b \), if any. The second axiom, upper invariance, requires that the agent’s probabilities do not change for any object that it ranks below \( b \). Our first main result is that random assignment
mechanisms are strategyproof if and only if they satisfy all three axioms.

1.3. Bounded Indifference and Partial Strategyproofness

Arguably, lower invariance is the least important one of the three axioms. Indeed, as we will show later, many non-strategyproof mechanisms that are viewed as having “good” incentive properties violate only lower invariance, but satisfy swap monotonicity and upper invariance. To understand the incentives under mechanisms that are swap monotonic and upper invariant, we define a new relaxed notion of strategyproofness: following the intuition from the motivating example, we require mechanisms to make truthful reporting a dominant strategy, but on a restricted domain, where agents have sufficiently different values for different objects. This domain restriction can be formalized as follows: a utility function satisfies uniformly relatively bounded indifference with respect to bound \( r \in [0, 1] \) (URBI(r)) if, given \( u(a) > u(b) \), the agent’s (normalized) value for \( b \) is at least a factor \( r \) lower than its value for \( a \) (i.e., \( r \cdot u(a) \geq u(b) \)). We say that a mechanism is \( r \)-partially strategyproof if the mechanism makes truthful reporting a dominant strategy for any agent whose utility function satisfies URBI(r). Our second main result is the following equivalence: for any setting (i.e., number of agents, number of objects, and object capacities) a mechanism is swap monotonic and upper invariant if and only if it is \( r \)-partially strategyproof for some \( r > 0 \). Thus, partial strategyproofness is axiomatically motivated. Furthermore, it allows us to give honest and useful strategic advice to the agents: under any swap monotonic, upper invariant mechanism, agents are best off by reporting truthfully as long as they are not too close to indifference for different objects.

1.4. Maximality of URBI(r) and the Degree of Strategyproofness

Knowing that a given mechanism is \( r \)-partial strategyproofness yields a guarantee about the set of utility functions for which the mechanism will make truthful reporting a dominant strategy, namely all those utility functions that satisfy the URBI(r) constraint. However, no guarantee is given for utility functions that violate this constraint. For our third main result, we show that the URBI(r) domain restriction is in fact maximal for \( r \)-partially strategyproof mechanisms. Specifically, there exists no larger set of utility functions for which guarantees of this form can be given without any additional information about the agents or the mechanisms. In this sense, \( r \)-partial strategyproofness
is the strongest statement that we can make about the incentive properties of swap monotonic and upper invariant mechanisms.

By virtue of this maximality, partial strategyproofness induces a meaningful measure for the incentive properties of non-strategyproof mechanisms: we define the degree of strategyproofness $\rho$ as the largest value of the bound $r$ such that the mechanism remains $r$-partially strategyproof. We show that comparing mechanisms by their degree of strategyproofness is consistent with (but not equivalent to) the method for comparing mechanisms by their vulnerability to manipulation (Pathak and Sönmez, 2013). However, the degree of strategyproofness measure has two advantages: it is a parametric measure for incentive guarantees, while vulnerability to manipulation is a binary comparison; and it is algorithmically computable, while no algorithm is known to perform the comparison by vulnerability to manipulation. This makes the degree of strategyproofness a compelling new method to measure and compare the incentive properties of non-strategyproof mechanisms.

1.5. Properties of Partial Strategyproofness

With partial strategyproofness, we have defined a new, relaxed notion of strategyproofness for assignment mechanisms. The concept has a number of appealing properties, which we formalize, prove, and discuss in this paper.

1.5.1. Dominance Interpretation of Partial Strategyproofness

When assignments are random, the agents’ preferences must be extended to lotteries in some way. This is typically done via dominance notions, such as stochastic dominance (SD) or lexicographic dominance (DL-dominance). Any dominance notion in turn induces a strategyproofness notion that arises by requiring the outcomes from truthful reporting to dominate the outcomes from misreporting. For stochastic dominance, the resulting SD-strategyproofness is in fact equivalent to the definition of strategyproofness we use in this paper, which requires truthful reporting to maximize any agent’s expected utility.

The question arises whether partial strategyproofness has an alternative definition in terms of a dominance relation. We present the notion of $r$-partial dominance ($r$-PD), which is similar to stochastic dominance, except that the influence of changes in the assignment of less preferred objects is discounted by the factor $r$. Our fourth main
result is that $r$-partial strategyproofness is equivalent to $r$-PD-strategyproofness, the incentive concept induced by $r$-partial dominance. This equivalence has two important consequences: first, it allows an alternative definition of partial strategyproofness that is independent of the agents' utility functions. Thus, partial strategyproofness integrates nicely with other incentive concepts that are defined via dominance notions, such as strong and weak SD-strategyproofness, and DL-strategyproofness. Second, the dominance interpretation provides an equivalent condition in terms of finitely many linear constraints. This makes partial strategyproofness algorithmically verifiable and enables the computation of the degree of strategyproofness measure.

1.5.2. Intermediateness of the Partial Strategyproofness Concept

We also study the relationship of partial strategyproofness and other incentive concepts, and we establish that it provides a unified view on the incentive properties of non-strategyproof assignment mechanisms: while partial strategyproofness is a weaker requirement than strategyproofness, it in turn implies many relaxed incentive concepts that have been proposed previously, namely weak, convex, and approximate strategyproofness, as well as strategyproofness in the large (when the degree of strategyproofness converges to 1). Moreover, we prove out fifth main result, the following equivalence: a mechanism is $r$-partially strategyproof for some $r > 0$ if and only if it is strategyproof for agents with lexicographic preferences (i.e., DL-strategyproof). Thus, the “upper” limit case of $r$-partial strategyproofness for $r = 1$ corresponds to strategyproofness, while the “lower” limit case for $r \rightarrow 0$ corresponds to DL-strategyproofness. In this sense, the degree of strategyproofness parametrizes the whole space of mechanisms between those that are strategyproof on the one side and those that are merely DL-strategyproof on the other side.

1.5.3. Local Sufficiency for Partial Strategyproofness

Local misreports are swaps of two consecutive objects in the reported preference order of some agent. For an incentive concept, such as strategyproofness or DL-strategyproofness, local sufficiency holds if it suffices to check only the local misreports in order to verify that a given mechanism satisfies the incentive concept. For any incentive concept, local sufficiency means that in order to establish that truthful reporting is weakly better than
any misreport, it suffices to verify that no local misreports are ever beneficial (i.e., swaps of two consecutive objects in the agents’ true preference orders). Local sufficiency is interesting from an axiomatic as well as from an algorithmic perspective: for instance, the axioms swap monotonicity, upper invariance, and lower invariance are based on swaps, which makes them simple and accessible. Furthermore, the naïve way to verify algorithmically that a given mechanism satisfies a given incentive concept would be to iterate through all incentive constraints. This is typically a large number. If local sufficiency holds, it is enough to check only those constraints that arise from swaps.

For strategyproofness and DL-strategyproofness, local sufficiency was proven by Carroll (2012) and Cho (2012), respectively. Thus, local sufficiency holds for the two limit concepts of partial strategyproofness. This raises the question whether it also holds for partial strategyproofness. Our sixth main result is that $r$-local partial strategyproofness implies $r^2$-partial strategyproofness. Furthermore, the bound “2” is tight in the sense that there exists no $\epsilon > 0$ such that $r^{2-\epsilon}$-partial strategyproofness can also be guaranteed. This insight connects the prior local sufficiency results for strategyproofness and for DL-strategyproofness as it provides a unified proof for both results.

1.5.4. Applications to Deterministic Mechanisms

Partial strategyproofness crucially depends on the randomness of mechanisms; for deterministic mechanisms, $r$-partial strategyproofness for any $r > 0$ coincides with strategyproofness. Nonetheless, partial strategyproofness can also be applied to study the incentive properties of non-strategyproof deterministic mechanisms and other mechanisms that are not “random enough.” To this end, we consider a second source of randomness, namely the agents’ uncertainty about the reports from other agents. Specifically, for our seventh main result, we give an axiomatic characterization of the mechanisms that are partially strategyproof for agents who are unsure about the preference reports from the other agents. In particular, it is true for deterministic versions of the naïve and the adaptive Boston mechanisms (where priorities are strict and fixed), as well as the HBS Draft mechanism and the Probabilistic Serial mechanism for multi-unit assignment.
1.5.5. Applications of Partial Strategyproofness

Finally, we demonstrate that our partial strategyproofness concept yields new insights about the incentive properties of many popular, non-strategyproof mechanisms. First, the Probabilistic Serial mechanism is partially strategyproof, which provides a more accurate understanding of its incentive properties than, e.g., weak SD-strategyproofness. Numerically, we show that the degree of strategyproofness of PS increases in larger markets, which is in line with prior findings by Kojima and Manea (2010). Second, the traditional “naïve” Boston mechanism\(^2\) is known to be highly manipulable. However, a variant where agents automatically skip exhausted objects in the application process is in fact partially strategyproof. The degree of strategyproofness of this “adaptive” variant is lower than that of PS, and therefore, it has intermediate incentive guarantees. Third, in (Mennle and Seuken, 2015\(^a\)) we have introduced hybrid mechanisms, which are convex combinations of two mechanisms. Under certain technical conditions, we have shown that hybrids facilitate a trade-off between strategyproofness and efficiency, where we use partial strategyproofness to quantify the incentive properties. Prior to the introduction of partial strategyproofness, no concept existed to study such trade-offs. These examples highlight that partial strategyproofness captures our intuitive understanding of what it means for a non-strategyproof mechanism to have “good” incentive properties.

In summary, our axiomatic treatment of the strategyproofness concept leads to a new way of thinking about how to relax strategyproofness for assignment mechanisms. The resulting partial strategyproofness concept is simple, tight, parametric, integrates well with existing methods, and it differentiates nicely between many popular mechanisms.

**Organization of this paper:** In Section 2, we discuss related work. In Sections 3 and 4, we introduce our formal model and the three axioms. In Section 5, we present our axiomatic decomposition of strategyproofness, and in Section 6, we derive the new partial strategyproofness concept. In Section 7, we present our maximality result and the degree of strategyproofness measure. In Section 8, we give the dominance interpretation of partial strategyproofness. In Section 9, we compare partial strategyproofness to other incentive concepts, and in Section 10, we discuss local sufficiency. In Section 11, we extend partial strategyproofness to deterministic mechanisms. In Section 12, we apply our new concept to popular assignment mechanisms, and we conclude in Section 13.

\(^2\)We consider the Boston mechanism with no priorities and single uniform tie-breaking.
2. Related Work

While the seminal paper on assignment mechanisms by Hylland and Zeckhauser (1979) proposed a mechanism that elicits agents’ cardinal utilities, this approach has proven problematic because it is difficult if not impossible to elicit cardinal utilities in settings without money. For this reason, recent work has focused on ordinal mechanisms, where agents submit preference orders over objects. In fact, it has been shown that mechanisms whose outcomes have to be independent of the agents’ levels of wealth are bound to be ordinal (Huesmann and Wambach, 2015; Ehlers et al., 2015). Throughout this paper, we only consider ordinal mechanisms.

For the case of deterministic assignment mechanisms, strategyproofness has been studied extensively. Papai (2000) showed that the only group-strategyproof, ex-post efficient, reallocation-proof mechanisms are hierarchical exchanges. Characterizations of strategyproof, efficient, and reallocation-consistent (Ehlers and Klaus, 2006) or consistent (Ehlers and Klaus, 2007) mechanisms are also available. The only deterministic, strategyproof, ex-post efficient, non-bossy, and neutral mechanisms are known to be serial dictatorships (Hatfield, 2009). Furthermore, Pycia and Ünver (2014) showed that all group-strategyproof, ex-post efficient mechanisms are trading cycles mechanisms. Barbera, Berga and Moreno (2012) gave a decomposition of strategyproofness that is similar in spirit to ours, but is restricted to deterministic social choice domains.

For random social choice rules, Gibbard (1977) gave a decomposition of strategyproofness into the two properties localized and non-perverse. This (as well as any other) decomposition of strategyproofness is by definition equivalent to our decomposition. Here our contribution lies in the definition of simple and intuitive axioms that make the conditions accessible and straightforward to interpret. For random assignment mechanisms, Abdulkadiroğlu and Sönmez (1998) showed that Random Serial Dictatorship (RSD) is equivalent to the Core from Random Endowments mechanism for house allocation. Bade (2014) extended their result by showing that taking any ex-post efficient, strategyproof, non-bossy, deterministic mechanism and assigning agents to roles in the mechanism uniformly at random is equivalent to RSD. However, it is still an open conjecture whether RSD is the unique mechanism that is strategyproof, ex-post efficient, and anonymous (Lee and Sethuraman, 2011; Bade, 2014).

Besides the baseline requirement of ex-post efficiency, the research community has
also introduced stronger efficiency concepts, such as ordinal efficiency, which is achieved by the Probabilistic Serial (PS) mechanism (Bogomolnaia and Moulin, 2001). The PS mechanism has received considerable attention from researchers: Hashimoto et al. (2014) showed that PS with uniform eating speeds is in fact the unique mechanism that is ordinally fair and non-wasteful. Bogomolnaia and Moulin (2001) had already shown that PS is not strategyproof, but Ekici and Kesten (2012) found that its Nash equilibria can lead to ordinally dominated outcomes. While incentive concerns for PS may be severe for small settings, they get less problematic for larger settings: Kojima and Manea (2010) showed that for a fixed number of object types and a fixed agent, PS makes it a dominant strategy for that agent to be truthful if the number of copies of each object is sufficiently large.

While ex-post efficiency and ordinal efficiency are the most well-studied efficiency concepts for assignment mechanisms, some mechanisms used in practice aim to achieve rank efficiency, which is a further refinement of ordinal efficiency (Featherstone, 2011). However, no rank efficient mechanism can be strategyproof in general. Other popular mechanisms, like the Boston Mechanism (Ergin and Sönmez, 2006; Miralles, 2008), are highly manipulable but nevertheless in frequent use. Budish and Cantillon (2012) found practical evidence from combinatorial course allocation, suggesting that using a non-strategyproof mechanism may lead to higher social welfare than using an ex-post efficient and strategyproof mechanism, such as RSD. The fact that strategyproofness is in conflict with many other design objectives challenges whether it should be taken as an indispensable requirement in mechanism design.

Given that strategyproofness is such a strong restriction, many researchers have tried to relax it. Bogomolnaia and Moulin (2001) used weak SD-strategyproofness to describe the incentive properties under PS, and Balbuzanov (2014) showed that PS in fact satisfies the slightly more demanding convex strategyproofness. Carroll (2013) adapted approximate strategyproofness for bounded utilities to quantify agents’ incentives to manipulate in the voting domain. Azevedo and Budish (2015) proposed a desideratum called strategyproof in the large (SP-L), which formalizes the intuition that as the number of agents in the market gets large, the incentives for an individual agent to misreport its preference order should vanish in the limit. Finally, Cho (2012) considered strategyproofness for agents with lexicographic preferences (DL-strategyproofness). We show that partial strategyproofness unifies these relaxations of strategyproofness: on the one hand, many
non-strategyproof mechanisms that are generally viewed as “having better incentive properties,” because they satisfy these various notions of strategyproofness, turn out to satisfy partial strategyproofness as well, such as Probabilistic Serial, a variant of the Boston Mechanism, and newly defined hybrid mechanisms. On the other hand, partial strategyproofness implies the other notions. Pathak and Sönmez (2013) introduced a general method to compare different mechanisms by their vulnerability to manipulation. The degree of strategyproofness measure we propose in this paper is consisten with (but not equivalent to) this method. However, our concept has two advantages: it yields a parametric relaxation of strategyproofness and it is computable. We discuss the connection in more detail in Section 7.2.2.

Local sufficiency is a property of an incentive concept, which holds when the absence of local manipulation opportunities implies the absence of manipulation opportunities “globally.” Carroll (2012) and Cho (2012) showed that local sufficiency holds for strategyproofness and DL-strategyproofness, respectively. We prove a local sufficiency result for partial strategyproofness, which bridges the gap between (and provides a unified proof for) both of these prior results.

3. Model

A setting \((N, M, q)\) consists of a set \(N\) of \(n\) agents, a set \(M\) of \(m\) objects, and a vector \(q = (q_1, \ldots, q_m)\) of capacities (i.e., there are \(q_j\) units of object \(j\) available). We assume that supply satisfies demand (i.e., \(n \leq \sum_{j \in M} q_j\)), since we can always add a dummy object with capacity \(n\). Agents \(i \in N\) have strict preference orders \(P_i\) over objects, where \(P_i: a > b\) means that agent \(i\) prefers object \(a\) to object \(b\). The set of all preference orders is denoted by \(\mathcal{P}\). A preference profile \(\mathbf{P} = (P_1, \ldots, P_n) \in \mathcal{P}^N\) is a collection of preference orders of all agents, and we denote by \(P_{-i} \in \mathcal{P}^{N\setminus i}\) the collection of preference orders of all agents, except \(i\). Agents’ preferences are extended to lotteries via von Neumann-Morgenstern (vNM) utilities \(u_i : M \rightarrow \mathbb{P}^+\). The utility function \(u_i\) is consistent with preference order \(P_i\) (denoted \(u_i \sim P_i\)) if \(P_i: a > b\) whenever \(u_i(a) > u_i(b)\). We denote by \(U_{P_i} = \{u_i : u_i \sim P_i\}\) the set of all utility functions that are consistent with \(P_i\).

A (random) assignment is a matching of objects to agents. It is represented by an \(n \times m\)-matrix \(x = (x_{i,j})_{i \in N, j \in M}\) satisfying the fulfillment constraint \(\sum_{j \in M} x_{i,j} = 1\), the capacity constraint \(\sum_{i \in N} x_{i,j} \leq q_j\), and the probability constraint \(x_{i,j} \in [0, 1]\) for all
\[
\mathbf{x}_{i,j} \in \mathcal{P} \times \mathcal{M}.
\]
The entry \(x_{i,j}\) of the matrix \(x\) is the probability that agent \(i\) gets object \(j\). An assignment is deterministic if all agents get exactly one full object (i.e., \(x_{i,j} \in \{0, 1\}\) for all \(i \in N, j \in M\)). For any agent \(i\), the \(i\)th row \(x_i = (x_{i,j})_{j \in M}\) of the matrix \(x\) is called the assignment vector of \(i\) (or \(i\)'s assignment for short). The Birkhoff-von Neumann Theorem and its extensions (Budish et al., 2013) ensure that, given any random assignment, we can always find a lottery over deterministic assignments that implements these marginal probabilities. Finally, let \(X\) and \(\Delta(X)\) denote the spaces of all deterministic and random assignments, respectively.

A (random) mechanism is a mapping \(\varphi : \mathcal{P}^N \rightarrow \Delta(X)\) that chooses an assignment based on a profile of reported preference orders. \(\varphi_i(P_i, P_{-i})\) is the assignment vector that agent \(i\) receives if it reports \(P_i\) and the other agents report \(P_{-i}\). A mechanism is deterministic if it only selects deterministic assignments (i.e., \(\varphi : \mathcal{P}^N \rightarrow X\)). Note that mechanisms only receive preference profiles as input (i.e., only agents’ preference orders), but no additional cardinal information. Thus, we consider ordinal mechanisms, where the assignment is independent of the actual vNM utilities. The expected utility for \(i\) is given by the dot product
\[
\langle u_i, \varphi_i(P_i, P_{-i}) \rangle = \mathbb{E}_{\varphi_i(P_i, P_{-i})}[u_i] = \sum_{j \in M} u_i(j) \cdot \varphi_{i,j}(P_i, P_{-i}).
\]

4. The Axioms

In this section, we introduce the axioms that we use to decompose and relax strategyproofness. To do so formally, we require the auxiliary concepts of neighborhoods and contour sets.

**Definition 1** (Neighborhood). For any two preference orders \(P, P' \in \mathcal{P}\) we say that \(P\) and \(P'\) are adjacent if they differ by just a swap of two consecutive objects; formally,
\[
\begin{align*}
P & : a_1 > \ldots > a_k > a_{k+1} > \ldots > a_m, \\
P' & : a_1 > \ldots > a_{k+1} > a_k > \ldots > a_m.
\end{align*}
\]
The set of all preference orders adjacent to \(P\) is the neighborhood of \(P\), denoted \(N_P\).
Definition 2 (Upper and Lower Contour Sets). For a preference order $P \in \mathcal{P}$ with $P : a_1 > \ldots > a_k > \ldots > a_m$, the upper contour set $U(a_k, P)$ and the lower contour set $L(a_k, P)$ of $a_k$ at $P$ are the sets of objects that an agent with preference order $P$ strictly prefers or likes strictly less than $a_k$, respectively; formally,

$$U(a_k, P) = \{a_1, \ldots, a_{k-1}\} = \{j \in M \mid P : j > a_k\}, \quad \text{(2)}$$

$$L(a_k, P) = \{a_{k+1}, \ldots, a_m\} = \{j \in M \mid P : a_k > j\}. \quad \text{(3)}$$

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 manipulation that an agent could consider. Our axioms limit the way in which a mechanism can change the assignment of the reporting agent under this basic manipulation.

Axiom 1 (Swap Monotonic). A mechanism $\varphi$ is swap monotonic if for any agent $i \in N$, any preference profile $(P_i, P_{-i}) \in \mathcal{P}^N$, and any misreport $P'_i \in N_{P_i}$ from the neighborhood of $P_i$ with $P_i : a_k > a_{k+1}$ and $P'_i : a_{k+1} > a_k$, one of the following holds:

- either $\varphi_i(P_i, P_{-i}) = \varphi_i(P'_i, P_{-i})$,
- or $\varphi_{i,a_k}(P_i, P_{-i}) > \varphi_{i,a_k}(P'_i, P_{-i})$ and $\varphi_{i,a_{k+1}}(P_i, P_{-i}) < \varphi_{i,a_{k+1}}(P'_i, P_{-i})$.

Swap monotonicity is an intuitive axiom, because it simply requires the mechanism to react to the swap in a direct and responsive way: the swap reveals information about the agent’s relative ranking of $a_k$ and $a_{k+1}$. Thus, if anything changes about the assignment for that agent, the probabilities for the objects $a_k$ and $a_{k+1}$ must be affected directly. In addition, the mechanism must respond to the agent’s preferences by assigning more probability for the object the agent claims to like more and less probability for the object the agent claims to like less.

Swap monotonicity 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 strictly 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 first order-stochastic dominance sense, as Example 1 shows.
Example 1. Consider a mechanism where reporting $P : a > b > c > d$ leads to an assignment vector of $(0, 1/2, 0, 1/2)$ of $a, b, c, d$, respectively, and reporting $P' : a > c > b > d$ leads to $(1/2, 0, 1/2, 0)$. This is consistent with swap monotonicity. However, the latter assignment vector first order-stochastically dominates the former at $P$.

While the swap monotonic mechanism in Example 1 is manipulable in a first order-stochastic dominance sense, the manipulations involves changes of the probabilities for other objects besides $a_k$ and $a_{k+1}$. To prevent this as well, we need an additional axiom.

Axiom 2 (Upper Invariant). A mechanism $\varphi$ is upper invariant if for any agent $i \in N$, any preference profile $(P_i, P_{-i}) \in \mathcal{P}^N$, and any misreport $P'_i \in N_{P_i}$ from the neighborhood of $P_i$ with $P_i : a_k > a_{k+1}$ and $P'_i : a_{k+1} > a_k$, we have that $i$'s assignment for objects from the upper contour set of $a_k$ does not change (i.e., $\varphi_{i,j}(P_i, P_{-i}) = \varphi_{i,j}(P'_i, P_{-i})$ for all $j \in U(a_k, P_i)$).

Intuitively, under upper invariance, an agent cannot influence its probabilities for obtaining one of its better choices by swapping two less preferred objects. Upper invariance was introduced by Hashimoto et al. (2014) as one of the central axioms to characterize the Probabilistic Serial mechanism. If a null object is present and the mechanism is individually rational, then upper invariance is equivalent to truncation robustness. Truncation robustness is a type of robustness to manipulation that is important in theory and application: it prevents that by bringing the null object up in its preference order, an agent can increase its chances of being assigned a more preferred object. Many mechanisms from the literature satisfy upper invariance, including Random Serial Dictatorship, Probabilistic Serial, the Boston mechanism, and Deferred Acceptance (for the proposing agents), and the HBS Draft mechanism for multi-unit assignment.

Axiom 3 (Lower Invariant). A mechanism $\varphi$ is lower invariant if for any agent $i \in N$, any preference profile $(P_i, P_{-i}) \in \mathcal{P}^N$, and any misreport $P'_i \in N_{P_i}$ from the neighborhood of $P_i$ with $P_i : a_k > a_{k+1}$ and $P'_i : a_{k+1} > a_k$, we have that $i$'s assignment for objects from the lower contour set of $a_{k+1}$ does not change (i.e., $\varphi_{i,j}(P_i, P_{-i}) = \varphi_{i,j}(P'_i, P_{-i})$ for all $j \in L(a_{k+1}, P_i)$).

Lower invariance complements upper invariance: it requires that an agent cannot influence its probabilities for obtaining any less preferred objects by swapping two more

\footnote{In Appendix B, we give a swap monotonic continuation of this mechanism for all preference orders.}
preferred objects. Lower invariance has a subtle effect on incentives: if agents were endowed with upward-lexicographic preferences (Cho, 2012), mechanisms that are not lower invariant will be manipulable for these agents, even if they are swap monotonic and upper invariant. Arguably, lower invariance is the least important axiom, but in Section 5, we show that it is exactly the missing link to guarantee strategyproof mechanisms. In Section 6, we drop lower invariance for our axiomatization of partially strategyproof mechanisms.

5. A New Decomposition of Strategyproofness

In this section, we review the strategyproofness concept. Then we present our first main result, an axiomatic decomposition of strategyproofness for random assignment mechanisms.

5.1. The Strategyproofness Concept

Ideally, a mechanism makes truthful reporting a dominant strategy for all agents. This requirement is formalized by strategyproofness.

**Definition 3** (Strategyproof). A mechanism \( \varphi \) is strategyproof if for any agent \( i \in N \), any preference profile \((P_i, P_{-i}) \in \mathcal{P}^N\), any misreport \( P_i' \in \mathcal{P} \), and any utility function \( u_i \in U_{P_i} \) that is consistent with \( P_i \), we have

\[
\langle u_i, \varphi_i(P, P_{-i}) - \varphi_j(P_i', P_{-i}) \rangle \geq 0
\]  

Alternatively, strategyproofness can be defined via the notion of stochastic dominance. For a preference order \( P \in \mathcal{P} \) with \( P : a_1 > \ldots > a_m \) and assignment vectors \( x \) and \( y \), we say that \( x \) (first order-)stochastically dominates \( y \) at \( P \) if for all ranks \( k \in \{1, \ldots, m\} \) we have

\[
\sum_{i=1}^{k} x_{a_i} \geq \sum_{i=1}^{k} y_{a_i}.
\]  

This means that the probability of obtaining one’s \( k \)th choice or better is weakly higher under \( x \) than under \( y \). Intuitively, an agent with preference order \( P \) would unambiguously prefer \( x \) to \( y \). The dominance is strict if in addition inequality (5) is strict for some
A mechanism $\varphi$ is \textit{stochastic dominance-strategyproof (SD-strategyproof)} if truthful reporting always yields a stochastically dominant assignment vector for the reporting agent (i.e., $\varphi_i(P, P_{-i})$ stochastically dominates $\varphi_i(P', P_{-i})$ at $P_i$). Strategyproofness in the sense of expected utilities from Definition 3 and SD-strategyproofness are equivalent (Erdil, 2014), and we will simply refer to this requirement as \textit{strategyproofness} for the rest of the paper.

For deterministic mechanisms, strategyproofness and swap monotonicity coincide.

\textbf{Proposition 1.} A deterministic mechanism $\varphi$ is strategyproof if and only if it is swap monotonic.

The proof is straightforward and given in Appendix C. Example 1 shows that this equivalence no longer holds for random mechanisms. Our following decompositions result shows which additional axioms are needed for strategyproofness of random mechanisms.

\section{5.2. Decomposition Result}

We are now ready to formulate our first main result, the decomposition of strategyproofness into the three axioms swap monotonicity, upper invariance, and lower invariance.

\textbf{Theorem 1.} A mechanism $\varphi$ is strategyproof if and only if it is swap monotonic, upper invariant, and lower invariant.

\textit{Proof Outline (formal proof in Appendix D).} Assuming strategyproofness, we consider a swap of two consecutive objects in the report of some agent. Towards contradiction, we suppose that $\varphi$ violates either upper or lower invariance. We show that this implies that the assignment for the manipulating agent from misreporting is not stochastically dominated by the assignment from truthful reporting, which contradicts strategyproofness. With upper and lower invariance established, swap monotonicity follows as well. For necessity, if $\varphi$ satisfies the axioms, we show that any swap of two consecutive objects produces an assignment that is stochastically dominated by the assignment from reporting truthfully. Using a result by Carroll (2012), this \textit{local} strategyproofness can be extended to strategyproofness.

\textbf{Theorem 1} illustrates why strategyproofness is so restrictive: if an agent swaps two consecutive objects in its preference order, the only thing that a strategyproof mechanism
can do (if anything) is to increase that agent’s probability for the object that is swapped up and decrease its probability for the object that is swapped down by the same amount.

In his seminal paper, Gibbard (1977) gave a decomposition of strategyproofness for random social choice rules that is similar in spirit to ours: he showed that any rule is strategyproof if and only if it is localized and non-perverse. We would like to point out that any characterization of the set of strategyproof mechanisms is by its very nature equivalent to any other characterization of this set. Our decomposition in Theorem 1 is appealing because of the choice of axioms, which are simple and straightforward. This makes the decomposition useful, e.g., when proving strategyproofness of new mechanisms, or when encoding strategyproofness as constraints to an optimization problem under the automated mechanism design paradigm (Sandholm, 2003).

Remark 1. Theorem 1 can be extended to the case where indifferences between different objects are possible. In the present paper we focus on a model that rules out indifferences, but we have given the extension in (Mennle and Seuken, 2014b).

6. An Axiomatic Decomposition of Partial Strategyproofness

In the previous section, we have seen that swap monotonicity, upper invariance, and lower invariance are necessary and sufficient conditions for strategyproofness. Example 1 has shown that swap monotonicity and upper invariance are essential to guarantee at least truncation robustness and the absence of manipulations in a stochastic dominance-sense. Furthermore, we observed that lower invariance is the least intuitive and the least important of the axioms. Obviously, mechanisms that violate lower invariance are not strategyproof. However, we show that such mechanisms still make truthful reporting a dominant strategy for a subset of the utility functions. This leads to a relaxed notion of strategyproofness, which we call partial strategyproofness: we show that swap monotonicity and upper invariance are equivalent to strategyproofness on utility functions that satisfy uniformly relatively bounded indifference.
6.1. Uniformly Relatively Bounded Indifference URBI(r)

Recall the motivating example from the introduction, where an agent was contemplating a misreport under the Probabilistic Serial mechanism. $r = 3/4$ was the pivotal degree of indifference, which determined whether the misreport was beneficial to the agent or not. Generalizing the idea from this example, we introduce the concept of uniformly relatively bounded indifference: loosely speaking, an agent must value any object at least a factor $r$ less than the next better object (after appropriate normalization).

**Definition 4 (URBI).** A utility function $u$ satisfies uniformly relatively bounded indifference with respect to bound $r \in [0, 1]$ (URBI($r$)) if for any objects $a, b \in M$ with $u(a) > u(b)$, we have

$$r \cdot (u(a) - \min u) \geq u(b) - \min u,$$

where we write $\min u = \min_{j \in M} (u(j))$ for the utility of the last choice.

If $\min_{j \in M} (u(j)) = 0$, uniformly relatively bounded indifference has an intuitive interpretation, because inequality (6) simplifies to $r \cdot u(a) \geq u(b)$. In words, this means that given a choice between two objects $a$ and $b$, the agents must value $b$ at least a factor $r$ less than $a$.

For a geometric interpretation of URBI($r$), consider Figure 1: the condition means that the agent’s utility function, represented by the vector $u$, cannot be arbitrarily close to the indifference hyperplane $H(P, P')$ between the sets of consistent utility functions.
$U_P$ and $U_{P'}$, but it must lie within the shaded area in $U_P$. A different utility function in $U_P$, represented by the vector $\tilde{u}$, would violate URBI$(r)$. For convenience we introduce the convention that URBI$(r)$ denotes the set of all utility functions that satisfy uniformly relatively bounded indifference with respect to $r$, so that we can write “$u \in$ URBI$(r)$” to indicate that $u$ satisfies the requirement with respect to bound $r$.

**Remark 2.** To gain some intuition about the “size” of the set URBI$(r)$, consider a setting with $m = 3$ objects. Suppose that $\min_{j \in M} (u(j)) = 0$ and that the utilities for the first and second choice are determined by drawing a vector uniformly at random from $(0,1)^2 \setminus H(P, P')$ (i.e., from the open unit square excluding the indifference hyperplane). Then the share of utilities that satisfy URBI$(r)$ is $r$. For example, if $r = 0.4$, the probability of drawing a utility function from URBI$(0.4)$ is $0.4$. In Figure 1, this corresponds to the area of the shaded triangle over the area of the larger triangle formed by the $x$-axis, the diagonal $H(P, P')$, and the vertical dashed line on the right.

### 6.2. Definition of $r$-Partial Strategyproofness

Using URBI$(r)$, we now define our new relaxed notion of strategyproofness.

**Definition 5 (Partially Strategyproof).** 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 any agent $i \in N$, any preference profile $(P_i, P_{-i}) \in \mathcal{P}_N$, any misreport $P'_i \in \mathcal{P}$, and any utility function $u_i \in U_{P_i} \cap$ URBI$(r)$ that is consistent with $P_i$ and satisfies URBI$(r)$, we have

$$\langle u_i, \varphi_i(P_i, P_{-i}) - \varphi_j(P'_i, P_{-i}) \rangle \geq 0. \quad (7)$$

When the setting is clear from the context, we simply write $r$-partially strategyproof without explicitly stating the setting. Furthermore, we say that $\varphi$ is partially strategyproof if there exists some non-trivial bound $r > 0$ for which $\varphi$ is $r$-partially strategyproof.

The definition of partial strategyproofness is very similar to Definition 3 of strategyproofness; the only difference is that inequality (7) only has to hold for utility functions $u_i$ that satisfy URBI$(r)$. In this sense, $r$-partial strategyproofness for $r < 1$ is weaker than strategyproofness, but it is equivalent to strategyproofness for $r = 1$. In Section 9, we explore the connection of partial strategyproofness to other incentive concepts in detail.
6.3. Axiomatic Decomposition Result

In this section, we show our second main result that dropping lower invariance, but requiring swap monotonicity and upper invariance, leads to the class of partially strategyproof mechanisms.

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

**Proof outline (formal proof in Appendix E).** Suppose, an agent \(i\) has true preference order \(P_i : a_1 > \ldots > a_K > a_{K+1} > \ldots > a_m\) and is considering a misreport \(P'_i\) that leaves the positions of the first \(K\) choices unchanged. We first show that under swap monotonicity and upper invariance, it suffices to consider misreports \(P'_i\) for which the assignment of \(a_{K+1}\) strictly decreases. The key insight comes from considering certain chains of swaps and their impact on the assignment (called canonical transitions, see Claim 1). Then we show that for sufficiently small \(r \in (0, 1]\), the decrease in expected utility that is caused by the decrease in the assignment of \(a_{K+1}\) is sufficient to deter manipulation by any agent whose utility function satisfies \(\text{URBI}(r)\), even though its assignment for less preferred objects \(a_{K+2}, \ldots, a_m\) may improve. Finally, we show that a strictly positive \(r\) can be chosen uniformly for all preference profiles and misreports. Thus, the bound \(r\) depends only on the mechanism and the setting.

To see necessity, we assume towards contradiction that the mechanism is not upper invariant. For any \(r \in (0, 1]\) we construct a utility function that satisfies \(\text{URBI}(r)\), but for which the mechanism would be manipulable. The key idea is to make the agent want the object from the upper contour set strongly enough to remedy any other (negative) effects of the swap. Finally, using upper invariance, swap monotonicity follows in a similar fashion.

Theorem 2 yields an axiomatic motivation for the definition of partial strategyproofness: if we wish to retain truncation robustness and prevent manipulability in a stochastic dominance-sense, the set of partially strategyproof mechanisms arises naturally. It also shows what requiring strategyproofness on top of partial strategyproofness buys, namely lower invariance.
Finally, the equivalence also teaches us what straightforward and honest strategic advice we can give to the agents: if the mechanism is swap monotonic and upper invariant, we can tell agents that they are best off reporting their preferences truthfully as long as their cardinal valuations for different objects are sufficiently different.

Remark 3. In light of the interpretation of upper invariance as robustness to manipulation by truncation, dropping lower invariance suggests itself as the most sensible approach to relaxing strategyproofness. Alternatively, one could consider mechanisms that are swap monotonic and lower invariant, but violate upper invariance. Naturally, this gives rise to a different class of non-strategyproof mechanisms, which is related to upward lexicographic-strategyproofness in a similar way in which partial strategyproofness is related to downward lexicographic-strategyproofness (see Theorem 5). We leave this approach to future research.

7. Maximality of URBI(r) and the Degree of Strategyproofness

$r$-partial strategyproofness of a mechanism $\varphi$ requires that agents with a utility function in URBI($r$) have a dominant strategy to report their preferences truthfully. However, this does not imply that the set of utility functions for which $\varphi$ makes truthful reporting a dominant strategy is exactly equal to the set URBI($r$). The following Example 2 shows that in general we cannot hope for an exact equality.

Example 2. Consider a setting with 4 agents and 4 objects $a, b, c, d$ in unit capacity. In this setting, the adaptive Boston mechanism (ABM) (Mennle and Seuken, 2015b) is $r$-partially strategyproof for any $r \leq \frac{1}{3}$, but not $r$-partially strategyproof for any $r > \frac{1}{3}$. However, an agent $i$ with preference order $P_i : a > b > c > d$ and consistent utility function $\tilde{u}_i = (6, 2, 1, 0)$ will not find a beneficial manipulation for any reports $P_{-i}$ from the other agents. Thus, ABM makes truthful reporting a dominant strategy for an agent with utility function $\tilde{u}_i$ in this setting. But $\tilde{u}_i$ violates URBI $(1/3)$, since

$$\frac{\tilde{u}_i(c) - \min_{j \in M} (\tilde{u}(j))}{\tilde{u}_i(b) - \min_{j \in M} (\tilde{u}(j))} = \frac{1 - 0}{2 - 0} = \frac{1}{2} > \frac{1}{3}.$$ (8)

This example can be verified using Algorithm 1 in Appendix H.
7.1. Maximality of URBI(r) for Partially Strategyproof Mechanisms

Despite Example 2, the URBI(r) domain restriction is maximal in the following sense: consider a mechanism \( \varphi \) that is \( r \)-partially strategyproof for some bound \( r \in (0, 1) \). We show that, unless we are given additional structural information about \( \varphi \), URBI(r) is in fact the largest set of utilities for which truthful reporting is guaranteed to be a dominant strategy. We now prove this maximality of the URBI(r) domain restriction, our third main result.

**Theorem 3.** For any setting \((N, M, q)\) with \( m \geq 3 \), any bound \( r \in (0, 1) \), and any utility function \( \tilde{u}_i \) (consistent with a preference order \( P_i \)) that violates URBI(r), there exists a mechanism \( \tilde{\varphi} \) such that

1. \( \tilde{\varphi} \) is \( r \)-partially strategyproof, but
2. there exist preferences of the other agents \( P_{-i} \) and a misreport \( P'_i \) such that

\[
\langle \tilde{u}_i, \tilde{\varphi}_i(P_i, P_{-i}) - \tilde{\varphi}_i(P'_i, P_{-i}) \rangle < 0. \tag{9}
\]

Furthermore, \( \tilde{\varphi} \) can be chosen to satisfy anonymity.

Proof outline (formal proof in Appendix F). If \( \tilde{u}_i \) violates URBI(r), there must be a pair of objects \( a, b \in M \) with \( P_i : a > b \), such that

\[
\frac{\tilde{u}_i(b) - \min_{j \in M} (\tilde{u}(j))}{\tilde{u}_i(a) - \min_{j \in M} (\tilde{u}(j))} = \tilde{r} > r. \tag{10}
\]

We construct the mechanism \( \tilde{\varphi} \) explicitly. \( \tilde{\varphi} \) assigns a constant vector to agent \( i \), except when \( i \) reports some preference order \( P'_i \) with \( P'_i : b > a \). In that case \( \tilde{\varphi} \) assigns less of \( a \), more of \( b \), and less of \( i \)'s reported last choice (say, \( c \)) to \( i \). Then \( \tilde{\varphi} \) is swap monotonic and upper invariant. The re-assignment between \( a, b \), and \( c \) must be constructed in such a way that \( i \) would want to manipulate if its utility is \( \tilde{u}_i \), but would not want to manipulate if its utility satisfied URBI(r). We show that this is possible. Finally, by keeping all other agents’ assignment vectors constant and randomizing over the roles of the agents in the mechanism, we obtain an anonymous mechanism with these properties.

If some additional constraints are imposed on the space of possible mechanisms, the mechanism \( \tilde{\varphi} \) constructed in the proof of Theorem 3 may no longer be feasible, such that
the counterexample fails. However, without any such constraints, we cannot rule out the possibility that an agent with some utility function outside URBI\((r)\) may want to manipulate an \(r\)-partially strategyproof mechanism. The following Corollary 1 makes this argument precise.

**Corollary 1.** For any setting \((N, M, q)\) with \(m \geq 3\) objects, we have

\[
URBI(r) = \bigcap_{\varphi \text{ r-partially strategyproof in } (N, M, q)} \left\{ u \left| \varphi \text{ makes truthful reporting a dominant strategy for agents with utility } u \text{ in } (N, M, q) \right. \right\}.
\]

(11)

This means that when considering the set of \(r\)-partially strategyproof mechanisms, the set of utilities for which all of them make truthful reporting a dominant strategy is exactly equal to URBI\((r)\). Thus, there is no larger domain restriction for which all these mechanisms will also guarantee good incentives.

### 7.2. A New Parametric Measure for Incentive Properties

The partial strategyproofness concept leads to a new, intuitive measure for the incentive properties of swap monotonic, upper invariant mechanisms: we consider the largest possible value \(r\) for which the mechanism is still \(r\)-partially strategyproof.

**Definition 6** (Degree of Strategyproofness). Given a setting \((N, M, q)\) and a swap monotonic, upper invariant mechanism \(\varphi\), the degree of strategyproofness of \(\varphi\) is

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

(12)

Observe that for \(0 < r' < r \leq 1\) we have URBI\((r') \subset URBI(r)\) by construction. A mechanism that is \(r\)-partially strategyproof will also be \(r'\)-partially strategyproof. Thus, a higher degree of strategyproofness corresponds to a stronger guarantee.

**Remark 4.** In (12), we use the maximum (rather than the supremum). This is possible, because the constraint (7) in Definition 5 of partial strategyproofness is a weak inequality. Thus, the set of utilities for which a mechanism makes truthful reporting a dominant strategy is topologically closed. Consequently, there exists some maximal value \(\rho > 0\), for
which the mechanism is \( \rho \)-partially strategyproof, but it is not \( r \)-partially strategyproof for any \( r > \rho \). This ensures that \( \rho_{(N,M,q)}(\varphi) \) is well-defined.

### 7.2.1. Interpretation of the Degree of Strategyproofness

Maximality of the URBI(\( r \)) domain restriction (especially, Corollary 1) implies that when measuring the degree of strategyproofness of swap monotonic and upper invariant mechanisms using \( \rho_{(N,M,q)}(\varphi) \), no utility functions are omitted for which a guarantee could also be given. Thus, without further information on the mechanism, there cannot exist another single-parameter measure that conveys strictly more information about the incentive guarantees of \( \varphi \). The degree of strategyproofness also allows for the comparison of two mechanisms: “\( \rho_{(N,M,q)}(\varphi) > \rho_{(N,M,q)}(\psi) \)” means that \( \varphi \) make truthful reporting a dominant strategy for a larger URBI(\( r \)) domain restriction than \( \psi \).

From a quantitative perspective one might ask for “how many more” utility functions \( \varphi \) is guaranteed to have good incentives, compared to \( \psi \). Recall Remark 2, where we considered URBI(0.4) in a setting with 3 objects, \( \min_{j \in M} (u(j)) = 0 \), and the remaining utilities for the first and second choice were chosen uniformly at random from the unit square. Suppose that \( \rho_{(N,M,q)}(\varphi) = 0.8 \) and \( \rho_{(N,M,q)}(\psi) = 0.4 \). The set URBI(0.8) has “twice the size” of URBI(0.4). Thus, the guarantee for \( \varphi \) extends over twice as many utility functions as the guarantee for \( \psi \). In this sense, \( \varphi \) is “twice as strategyproof” as \( \psi \).

### 7.2.2. Relation of Degree of Strategyproofness and Vulnerability to Manipulation

Pathak and Sönmez (2013) proposed an interesting method for comparing mechanisms by their vulnerability to manipulation. An extension of this concept to the case of vNM utilities is straightforward: \( \psi \) is strongly as manipulable as \( \varphi \) if whenever an agent with utility \( u \) finds a beneficial misreport under \( \varphi \), the same agent in the same situation also finds a beneficial misreport under \( \psi \). The following Proposition 2 shows that vulnerability to manipulation and the degree of strategyproofness are consistent (but not equivalent).

**Proposition 2.** For any setting \( (N,M,q) \) and mechanisms \( \varphi \) and \( \psi \), the following hold:

1. If \( \psi \) is strongly as manipulable as \( \varphi \), then \( \rho_{(N,M,q)}(\varphi) \geq \rho_{(N,M,q)}(\psi) \).
2. If \( \rho_{(N,M,q)}(\varphi) > \rho_{(N,M,q)}(\psi) \) and \( \varphi \) and \( \psi \) are comparable by their vulnerability to manipulation, then \( \psi \) is strongly as manipulable as \( \varphi \).
The proof as well as the formal definition of the strongly as manipulable as-relation for random assignment mechanisms are given in Appendix F.2. Despite the consistency result, neither concept is always better at strictly differentiating mechanisms: a comparison by vulnerability to manipulation may be inconclusive when the degree of strategyproofness yields a strict winner; conversely, the degree of strategyproofness may indicate indifference (i.e., $\rho_{(N,M,q)}(\varphi) = \rho_{(N,M,q)}(\psi)$) when one of the mechanisms is in fact strongly more manipulable than the other.

An important difference between the two concepts is that the comparison by vulnerability to manipulation considers each preference profile separately, while the partial strategyproofness constraint must hold uniformly for all preference profiles. Thus, vulnerability to manipulation 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, Pathak and Sönmez (2013) do not present a method to perform the comparison by vulnerability to manipulation algorithmically, and the definition of such a method is not straightforward. In contrast, $\rho_{(N,M,q)}$ is computable (see Remark 5 in Section 8 and Algorithm 2 in Appendix H). Second, and more importantly, the degree of strategyproofness is a parametric measure while the strongly as manipulable as-relation is not. A mechanism designer could easily express a minimal acceptable degree of strategyproofness and then consider only mechanisms satisfying this constraint. A similar design approach using vulnerability to manipulation appears much more difficult as it would require the definition of a “benchmark mechanism” $\psi$ with maximal acceptable manipulability and considering only mechanisms that are less manipulable than $\psi$.

8. A Dominance Interpretation of Partial Strategyproofness

Partial strategyproofness restricts the set of utility functions for which the mechanism must make truthful reporting a dominant strategy, but the definition is otherwise analogous to Definition 3 of strategyproofness. Furthermore, recall that strategyproofness is equivalent to SD-strategyproofness, the incentive concept induced by stochastic dominance. Our fourth main result shows that an analogous equivalence exists for
$r$-partial strategyproofness. Specifically, it is equivalent to the incentive concept induced by a certain dominance notion. In this section, we formally define $r$-partial dominance, and we show that $r$-partial strategyproofness and $r$-partial dominance-strategyproofness are in fact the same.

**Definition 7 (Partial Dominance).** For a preference order $P \in \mathcal{P}$ with $P : a_1 > \ldots > a_m$, a bound $r \in [0, 1]$, and assignment vectors $x, y$, we say that $x$ $r$-partially dominates $y$ at $P$ if for all ranks $k \in \{1, \ldots, m\}$ we have

$$\sum_{l=1}^{k} r^l \cdot x_{a_l} \geq \sum_{l=1}^{k} r^l \cdot y_{a_l}. \quad (13)$$

Observe that for $r = 1$, this definition is precisely the same as stochastic dominance, since $1^l = 1$ for any $l$. However, for $r < 1$, the impact of less preferred objects is discounted by the factor $r$. Intuitively, inequality (13) can be interpreted as incentive constraint for certain “extreme” utility functions that satisfy URBI($r$), put very high value on the first $k$ objects, and next to no value on all other objects.

Analogous to stochastic dominance for SD-strategyproofness, we can use $r$-partial dominance to define $r$-partial dominance-strategyproofness.

**Definition 8 (PD-Strategyproof).** Given a setting $(N, M, q)$ and a bound $r \in (0, 1]$, a mechanism $\varphi$ is $r$-partial dominance-strategyproof ($r$-PD-strategyproof) if for any agent $i \in N$, any preference profile $(P_i, P_{\sim i}) \in \mathcal{P}^N$, and any misreport $P'_i \in \mathcal{P}$, $\varphi_i(P_i, P_{\sim i})$ $r$-partially dominates $\varphi_i(P'_i, P_{\sim i})$ at $P_i$.

We are now ready to formally state our fourth main result, the equivalence of $r$-partial strategyproofness and $r$-PD-strategyproofness.

**Theorem 4.** Given a setting $(N, M, q)$ and a mechanism $\varphi$, the following are equivalent:

1. $\varphi$ is $r$-partially strategyproof.
2. $\varphi$ is $r$-PD-strategyproof.

**Proof Outline (formal proof in Appendix G).** The challenge is that the sets URBI($r$) $\cap$ $U_P$ are unbounded, and therefore, they cannot be represented as convex polytopes with finitely many corner points. Nonetheless, the partial sums in the definition of partial
dominance can be interpreted as the incentives to misreport that are induced by different “extreme utility functions.” We prove that these extreme utilities are the essential limit cases that determine \( r \)-partial strategyproofness: we show that \( x \) is preferred to \( y \) by an agent with some utility \( u \) in \( \text{URBI}(r) \) if and only if this is also true for at least one of the extreme utilities.

Theorem 4 means that the two requirements that (i) “a mechanism makes truthful reporting a dominant strategy for any agent with a utility function in \( \text{URBI}(r) \),” and (ii) “any assignment vector that an agent can obtain by misreporting is \( r \)-partially dominated by the assignment vector this agent can obtain by reporting truthfully,” are in fact the same. This yields an alternative way of defining \( r \)-partial strategyproofness that does not rely on the agents’ utility functions. This shows that the concept integrates nicely in the landscape of existing incentive concepts, most of which also rely on dominance notions (e.g., strong and weak SD-strategyproofness and DL-strategyproofness).

Moreover, the alternative definition of partial strategyproofness via partial dominance also unlocks the concept to algorithmic analysis. The original definition imposed inequalities that had to hold for all utility functions within the set \( \text{URBI}(r) \) (infinitely many). While this provides a good economic intuition, it 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 partial dominance are satisfied. The finite condition can also be used to encode \( r \)-partial strategyproofness as linear constraints to an optimization problem. This enables an automated search within the set of \( r \)-partially strategyproof mechanisms while optimizing for some other design objective under the automated mechanism design paradigm (Sandholm, 2003).

Remark 5. In Appendix H we give algorithms that exploit the structure of \( r \)-PD-strategyproofness to verify whether a mechanism \( \varphi \) is \( r \)-partially strategyproof in a given setting (Algorithm 1), and to compute its degree of strategyproofness \( \rho_{(N,M,q)}(\varphi) \) (Algorithm 2).
9. Intermediateness of Partial Strategyproofness

In this section, we study the relationship of partial strategyproofness to other incentive concepts that have been discussed previously in the context of assignment mechanisms or more broadly in domains with no monetary transfer. We demonstrate that our new concept takes an intermediate position between strategyproofness on one side and many other concepts on the other side: while partial strategyproofness is implied by strategyproofness, it in turn implies weak, convex, and approximate strategyproofness, as well as strategyproofness in the large (if the degree of strategyproofness converges to 1). Most importantly, while strategyproofness is the upper limit concept (for $r = 1$), we show that strategyproofness for agents with lexicographic preferences is the lower limit concept for $r \to 0$. Figure 2 gives an overview of the relationships between the different incentive concepts.

### 9.1. Relation to Strategyproofness

We have already observed that $r$-partial strategyproofness for $r = 1$ is equivalent to strategyproofness. Thus, strategyproofness can be considered an upper limit case of $r$-partial strategyproofness.

*Remark* 6. To obtain a more formal understanding of this limit case, let $\text{SP}(N, M, q)$ and $r$-$\text{PSP}(N, M, q)$ denote the sets of strategyproof and $r$-partially strategyproof mechanisms.

---

Figure 2: Relationships between incentive concepts ($SP$: strategyproofness).
in the setting \((N, M, q)\), respectively. It is straightforward to see that

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

In words, any mechanisms that are \(r\)-partially strategyproof for all \(r < 1\) must be strategyproof. In Section 9.5, we will prove a corresponding formal statement about the lower limit concept, DL-strategyproofness.

### 9.2. Relation to Weak SD-Strategyproofness

Weak SD-strategyproofness was employed by Bogomolnaia and Moulin (2001) to describe the incentive properties of the Probabilistic Serial mechanism. Recall that an assignment vector \(x\) stochastically dominates another assignment vector \(y\) at a preference order \(P\) if for any rank \(k\) an agent with preference order \(P\) is at least as likely to obtain its \(k\)th choice or better under \(x\) than under \(y\). This dominance of \(x\) over \(y\) is strict if in addition for some rank the probability is strictly greater under \(x\) than under \(y\). Under weakly SD-strategyproof mechanisms, agents cannot attain a strictly dominant assignment vector by misreporting; however, in contrast to strategyproof mechanisms, the assignment vectors do not need to be comparable by stochastic dominance.

**Definition 9** (Weakly SD-Strategyproof). A mechanism \(\varphi\) is *weakly SD-strategyproof* if for any agent \(i \in N\), any preference profile \((P_i, P_{-i}) \in \mathcal{P}^N\), and any misreport \(P'_i \in \mathcal{P}\), agent \(i\)'s assignment vector from truthful reporting is not strictly stochastically dominated by its assignment vector from reporting \(P'_i\).

Weak SD-strategyproofness is equivalent to requiring that for a given preference profile \((P_i, P_{-i})\) and a potential misreport \(P'_i\), there exists a consistent utility \(u_i \in U_P\) such that agent \(i\) would prefer reporting \(P_i\) to reporting \(P'_i\). This is an extremely weak requirement, because \(u_i\) may depend on \(P'_i\). In other words, the mechanism might still offer an opportunity to manipulate to the agent with utility \(u_i\). The only guarantee is that the particular report \(P'_i\) will not increase its expected utility. Thus, it is possible that for some preference order \(P_i\), truthful reporting is not a dominant strategy, independent of agent \(i\)'s utility functions.

We have shown that partial strategyproofness implies weak SD-strategyproofness.
**Proposition 3.** 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 Outline (formal proof in Appendix I.1).** We show that partial strategyproofness implies *convex strategyproofness* (Balbuzanov, 2014). This in turn implies weak SD-strategyproofness. Balbuzanov also gave an example of a mechanism that is weakly SD-strategyproof, but violates convex strategyproofness.

### 9.3. Relation to Approximate Strategyproofness

Approximate strategyproofness is a different relaxation of strategyproofness that has attracted interest in quasi-linear domains (Lubin and Parkes, 2012). Approximately strategyproof mechanisms may be manipulable, but there exists an upper bound on the gain that an agent can obtain by misreporting. The economic intuition behind this concept is that if the potential gain is small, agents might not be willing to collect the necessary information and deliberate about misreports, but stick with truthful reporting instead. In this section, we formalize a notion of approximate strategyproofness that is meaningful for assignment mechanisms. Then we show that partial strategyproofness implies approximate strategyproofness but that the converse may not hold.

When payments are possible, money provides a canonical unit of measure for the gain from misreporting. However, the assignment domain does not permit payments, which makes the definition and interpretation of approximate strategyproofness more challenging. Here, we follow earlier work, which defined approximate strategyproofness for ordinal social choice mechanisms via bounded utility functions (Birrell and Pass, 2011; Carroll, 2013).

**Definition 10 (Approximately Strategyproof).** Given a setting \((N, M, q)\) and a bound \(\varepsilon \in [0, 1]\), a mechanism \(\varphi\) is \(\varepsilon\)-approximately strategyproof (in the setting \((N, M, q)\)) if for any agent \(i \in N\), any preference profile \((P_i, P_{-i}) \in \mathcal{P}^N\), any misreport \(P'_i \in \mathcal{P}\), and any utility function \(u_i \in U_{P_i}\) that is consistent with \(P_i\) and bounded between 0 and 1 (i.e., \(u_i : M \to [0, 1]\)), the gain in expected utility from reporting \(P'_i\) is upper-bounded by \(\varepsilon\); formally,

\[
\langle u_i, \varphi(P'_i, P_{-i}) - \varphi(P_i, P_{-i}) \rangle \leq \varepsilon.
\] (15)
Note that if $u_i$ was not bounded, the potential gain from manipulation under any non-strategyproof mechanism would instantly become arbitrarily large (Carroll, 2013). Since $u_i$ is bounded between 0 and 1, a change of magnitude 1 in expected utility corresponds to getting one’s first choice instead of one’s last choice. Thus, “1” corresponds to the maximal gain from misreporting that any agent could obtain under an arbitrary mechanism. Relative to this value “1,” the parameter $\varepsilon$ is the share of this maximal gain by which any agent can at most improve its expected utility under an $\varepsilon$-approximately strategyproof mechanism. Furthermore, the gain will never exceed 1, which makes 1-approximate strategyproofness a void constraint that is trivially satisfied by any mechanism. Obviously, if $\varphi$ is $\varepsilon$-approximately strategyproof, then it is also $\varepsilon'$-approximately strategyproof for any $\varepsilon' \geq \varepsilon$.

Proposition 4 shows that partial strategyproofness implies approximate strategyproofness, but the converse is not true in general.

**Proposition 4.** 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.

**Proof Outline (formal proof in Appendix I.2).** Using the axiomatic decomposition of partial strategyproofness from Theorem 2, we derive an upper bound for the gain from manipulation that any agent with bounded utilities can obtain by misreporting. We show that this bound can be chosen strictly below 1. To see that the converse does not hold, we construct a simple counter-example. 

Proposition 4 yields new insights for partially strategyproof mechanisms: initially, the definition of $r$-partial strategyproofness only required good incentives for agents whose utility functions satisfy the URBI($r$) constraint. However, it imposes no restriction for agents with utilities outside this set. Proposition 4 shows that, even though these other agents may be able to benefit from misreporting, their incentive to do so is at least bounded by some $\varepsilon < 1$ in the sense of approximate strategyproofness.

### 9.4. Relation to Strategyproofness in the Large

Azevedo and Budish (2015) proposed strategyproofness in the large as an alternative when strategyproofness is incompatible with other essential design objectives. This
incentive concept captures the intuition that the ability of any single agent to improve its own interim assignment (i.e., in expectation) by misreporting may vanish as more agents participate in a mechanism. 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.

The model of Azevedo and Budish (2015) considered a finite set of vNM utility functions $\{u^1, \ldots, u^K\}$. Strategyproofness in the large requires that for any $\varepsilon > 0$ there exists a number $n_0$ of agents such that in any setting with sufficiently many agents (i.e., $n \geq n_0$), no agent can gain more than $\varepsilon$ by misreporting.\footnote{The original definition is more technical and involves probability measures over the other agents’ preferences. However, this simplified version suffices to illustrate the connection with partial strategyproofness.} To apply this concept to the random assignment problem, we need to specify in what sense settings get large. To this end, we follow (Kojima and Manea, 2010) and (Azevedo and Budish, 2015) and keep the number of objects constant, but we let the number of agents grow, and increase the objects’ capacities such that supply satisfies demand. Thus, we consider a sequence of settings $(N^n, M^n, q^n)_{n \geq 1}$ where the set of agents $\#N^n = n$ grows, the set of objects $M^n = M$ is held fixed, and capacities grow so that $n = \sum_{j \in M} q^n_j$, and $\min_{j \in M} (q^n_j) \to \infty$ for $n \to \infty$.

**Proposition 5.** Fix any finite set of utility functions $\{u^1, \ldots, u^K\} \subseteq \bigcup_{P \in \mathcal{P}} U_P$. If the degree of strategyproofness of $\varphi$ converges to 1 as the settings grow (i.e., $\rho_{(N^n,M,q^n)}(\varphi) \to 1$ for $n \to \infty$), then $\varphi$ is strategyproof in the large with respect to $\{u^1, \ldots, u^K\}$.

**Proof.** Any consistent utility function $u$ satisfies uniformly relatively bounded indifference for some (sufficiently large) $r < 1$. Let $\bar{r}$ be the largest of these values, such that $u^k \in UBRI(\bar{r})$ for all $k \in \{1, \ldots, K\}$. Since by assumption, $\varphi$ is $\bar{r}$-partially strategyproof in the settings $(N^n, M, q^n)$ for sufficiently large $n$, all agents will have a dominant strategy to report their preferences truthfully. \qed

Kojima and Manea (2010) showed that the incentives under the non-strategyproof Probabilistic Serial (PS) mechanism improve in larger settings: for any fixed utility function, PS eventually makes truthful reporting a dominant strategy for any agent with that utility function. Azevedo and Budish (2015) used this result to show that PS is in fact strategyproof in the large. In Section 12.2, we will show that PS is partially
strategyproof (in finite settings). In combination, these insights suggest the following conjecture: *as settings grow in the way defined above, the degree of strategyproofness of PS converges to 1*. A proof of this conjecture would strengthen the result of Kojima and Manea (2010), because it would specify the precise way in which the set of utility functions with good incentives grows. In combination with Proposition 5, it would also yield an elegant proof for the observation that PS is strategyproof in the large. In Section 12.2 we provide numerical evidence that supports this conjecture.

### 9.5. Relation to Lexicographic Strategyproofness

Finally, we compare our new partial strategyproofness concept to strategyproofness for agents with lexicographic preferences. In particular, we show that this is the lower limit concept of $r$-partial strategyproofness as $r \to 0$. The intuition of lexicographic preferences is that agents prefer any (arbitrarily small) increase in the probability for some object to any (arbitrarily large) increase in the probability for some less preferred object.

**Definition 11 (DL-Dominance).** For preference order $P \in \mathcal{P}$ with $P : a_1 > \ldots > a_m$ and assignment vectors $x, y$, we say that $x$ *lexicographically dominates* $(DL$-dominates) $y$ at $P$ if either $x = y$, or for some rank $k \in \{1, \ldots, m\}$ we have $x_k > y_k$ and $x_l = y_l$ for all $l \leq k - 1$.

DL-dominance induces DL-strategyproofness in the same way in which stochastic dominance induces SD-strategyproofness.

**Definition 12 (DL-Strategyproof).** A mechanism $\varphi$ is *DL-strategyproof* if for any agent $i \in N$, any preference profile $(P_i, P_{-i}) \in \mathcal{P}^N$, and any misreport $P'_i \in \mathcal{P}$, the assignment from truthfully reporting $P_i$ DL-dominates the assignment from misreporting $P'_i$.

Our fifth main result is Theorem 5, which yields an equivalence between partial strategyproofness and DL-strategyproofness.

**Theorem 5.** *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 DL-strategyproof.*

\footnote{Note that Theorem 1 from (Kojima and Manea, 2010) is precisely the statement that for any $\varepsilon > 0$, PS is $\varepsilon$-approximate strategyproofness in sufficiently large settings. However, by Proposition 4 this is strictly not sufficient for partial strategyproofness.}
Proof Outline (formal proof in Appendix I.3). The proof is analogous to the proof of Theorem 2, where we showed that partial strategyproofness is equivalent to the axioms swap monotonicity and upper invariance. The minimal change in the probability for the highest ranking object (for which there is any change) is now reflected by the strict change induced by DL-strategyproofness, if any.

Theorem 5 demonstrates that for random assignment mechanisms, DL-strategyproofness is an unnecessarily weak concept. Imposing swap monotonicity and upper invariance already yields that the mechanism must be \( r \)-partially strategyproof for some positive bound \( r \) (Theorem 2). Partial strategyproofness specifies the precise structure of the incentive guarantees via the URBI(\( r \)) domain restriction. In contrast, DL-strategyproofness is a purely binary requirement that is either satisfied by a mechanism or not, but it ignores the parametric nature of the set of utility functions for which truthful reporting is guaranteed to be a dominant strategy.

A second interesting consequence of Theorem 5 is the fact that DL-strategyproofness is the lower limit concept for partial strategyproofness. In Remark 6, we denoted by \( r\text{-PSP}(N,M,q) \) the set of all mechanisms that were \( r \)-partially strategyproof in the setting \((N,M,q)\). Similarly, let \( DL\text{-SP}(N,M,q) \) be the set of DL-strategyproof mechanisms in that setting.

**Corollary 2.** Given a setting \((N,M,q)\), we have

\[
DL\text{-SP}(N,M,q) = \bigcup_{r>0} r\text{-PSP}(N,M,q).
\]

In words, any mechanisms that is DL-strategyproof must be \( r \)-partially strategyproof for some strictly positive bound \( r > 0 \).

Corollary 2 is the formal counterpart to Remark 6, where we showed that strategyproofness is the upper limit of partial strategyproofness in the sense that

\[
SP(N,M,q) = \bigcap_{r<1} r\text{-PSP}(N,M,q).
\]
10. Local Sufficiency of Partial Strategyproofness

For some incentives concepts, it suffices to check whether no local misreports are beneficial in order to establish that no misreports are beneficial at all. In this case, we say that the incentive concept satisfies *local sufficiency*. Local sufficiency simplifies the respective incentive concept from an axiomatic as well as from an algorithmic perspective. For assignment mechanisms, the *local* misreports are those in the neighborhood of the manipulating agent’s true preference order, which arise by swapping two consecutive objects. Carroll (2012) and Cho (2012) proved local sufficiency for strategyproofness and DL-strategyproofness, respectively. In this section, we prove an analogous local sufficiency result for partial strategyproofness.

First, we formally define three notions of local strategyproofness.

**Definition 13** (Locally Strategyproof & Locally DL-Strategyproof). A mechanism $\varphi$ is *locally strategyproof* if for any agent $i \in N$, any preference profile $(P_i, P_{-i}) \in \mathcal{P}^N$, any misreport $P'_i \in N_{P_i}$ from the neighborhood of $P_i$, and any utility $u_i \in U_{P_i}$ that is consistent with $P_i$, we have

$$\langle u_i, \varphi_i(P_i, P_{-i}) - \varphi_i(P'_i, P_{-i}) \rangle \geq 0.$$  \hspace{1cm} (18)

$\varphi$ is *locally DL-strategyproof* if $\varphi_i(P_i, P_{-i})$ DL-dominates $\varphi_i(P'_i, P_{-i})$.

Analogously, we can define a local variant of partial strategyproofness.

**Definition 14** (Locally Partially Strategyproof). Given a setting $(N, M, q)$ and a bound $r \in (0, 1)$, a mechanism $\varphi$ is *$r$-locally partially strategyproof* (in the setting $(N, M, q)$) if for any agent $i \in N$, any preference profile $(P_i, P_{-i}) \in \mathcal{P}^N$, any misreport $P'_i \in N_{P_i}$ from the neighborhood of $P_i$, and any utility $u_i \in U_{P_i} \cap URBI(r)$ that is consistent with $P_i$ and satisfies URBI$(r)$, we have

$$\langle u_i, \varphi_i(P_i, P_{-i}) - \varphi_i(P'_i, P_{-i}) \rangle \geq 0.$$  \hspace{1cm} (19)

and we say that $\varphi$ is *locally partially strategyproof* if it is $r$-locally partially strategyproof for some non-trivial $r > 0$.

Facts 1 and 2 summarize the known local sufficiency results. Since local constraints are obviously necessary, local sufficiency always implies equivalence.
Fact 1 (Carroll, 2012). Local strategyproofness is sufficient for strategyproofness.


In combination with our equivalence result for partial strategyproofness and DL-strategyproofness (Theorem 5), Fact 2 immediately yields a weak notion of local sufficiency for partial strategyproofness.

Corollary 3. Given a setting \((N, M, q)\), if a mechanism \(\varphi\) is \(r\)-locally partially strategyproof for some \(r > 0\), then \(\varphi\) is \(r'\)-partially strategyproof for some \(r' > 0\).

Corollary 3 follows from the observation that local partial strategyproofness implies local DL-strategyproofness, which implies DL-strategyproofness, which in turn implies partial strategyproofness by Theorem 5. However, the bound \(r\) (locally) and the bound \(r'\) (globally) are not necessarily the same. Since \(r'\)-partial strategyproofness implies \(r'\)-local partial strategyproofness, we must have \(r' \leq r\), but \(r'\) may still be (much) smaller than \(r\). Our sixth main result establishes a precise connection between \(r\) and \(r'\).

Theorem 6. Given a setting \((N, M, q)\), if \(\varphi\) is \(r\)-locally partially strategyproof, then \(\varphi\) is \(r^2\)-partially strategyproof.

Proof Outline (formal proof in Appendix J.1). For any \(u \in U_P\) that satisfies URBI\((r^2)\) and any misreport \(P'\), we construct a line segment in the utility space that starts in \(u\) and ends in another utility function \(v \in U_P\). We then express the incentive to misreport \(P'\) (instead of \(P\)) for an agent with utility \(u\) as a telescoping sum over local incentive constraints along this line segment.\(^6\) In this representation all but the first and the last term cancel out, such that it collapses to the required inequality. Since local incentive constraints are only available for utility functions inside URBI\((r)\), we need to ensure that the line segment intersects the sets \(U_{P_k} \cap \text{URBI}(r)\) for every preference order \(P_k\) through which it passes.

Theorem 6 means that \(r\)-local partial strategyproofness is sufficient to guarantee \(r'\)-partial strategyproofness, where \(r' \leq r^2\). As a special case, we obtain that 1-local partial strategyproofness implies 1-partial strategyproofness, the local sufficiency result for strategyproofness of Carroll (2012). Furthermore, considering a sequence of bounds

\(^6\)This step is inspired by Carroll (2012)'s proof of local sufficiency for strategyproofness.
that approaches 0, we obtain the local sufficiency result for DL-strategyproofness of Cho (2012) in the limit. Thus, Theorem 6 unifies both prior results.

Local sufficiency is an intriguing concept: it can be used to greatly reduce the complexity of the incentive concepts. As an example, recall that our axioms swap monotonicity, upper invariance, and lower invariance were simple and intuitive, in part because they restricted the behavior of the mechanisms only for local misreports. From a computational perspective, local sufficiency reduces algorithmic complexity, because it reduces the number of constraints in the optimization problem that is used for searching optimal mechanisms under the automated mechanism design paradigm (Sandholm, 2003).

The question remains whether Theorem 6 is tight or whether the bound \( r' \leq r^2 \) can be improved in any way. First, note that it is straightforward to construct a counter-example showing that \( r' = r \) is out of the question, unless \( r \in \{0, 1\} \). In fact, as we show in the next Theorem 7, the bound \( r' = r^{2^\epsilon} \) is tight in the sense that \( 2^\epsilon \) is the smallest exponent for which a can be guaranteed.

**Theorem 7.** Given a setting \((N, M, q)\) with \( m \geq 4 \) objects, for any \( \epsilon > 0 \) there exists a bound \( r \in (0, 1) \) and a mechanism \( \varphi \) such that

1. \( \varphi \) is \( r \)-locally partially strategyproof,
2. \( \varphi \) is not \( r^{2^\epsilon} \)-partially strategyproof.

**Proof Outline (formal proof in Appendix J.2).** The proof is constructive by giving \( \varphi \) explicitly. To find a suitable mechanism, we initially generated special instances of \( \varphi \) for fixed \( \epsilon \) as solutions to a particular linear program. The main challenge was to subsequently infer the general structure of \( \varphi \) from the examples and to prove the required properties.  

Tightness by Theorem 7 means that \( r' = r^{2^\epsilon} \) is the best polynomial bound that allows a general statement about local sufficiency of the partial strategyproofness concept.

**Remark 7.** Note that the value \( r \) in the counter-examples in the proof of Theorem 7 may depend on \( \epsilon \). We leave the exploration of the relationship between \( r \) and \( r' \) for fixed \( r \) to future research.
11. An Extension of Partial Strategyproofness for Deterministic Mechanisms

Some assignment mechanisms do not involve randomization. For example, most results for school choice mechanisms were obtained under the assumption that priorities are fixed and strict. This makes the mechanisms deterministic. In Proposition 1, we have shown that swap monotonicity and strategyproofness coincide for deterministic mechanisms. Thus, any deterministic mechanism that is \( r \)-partially strategyproof for some \( r > 0 \) must be strategyproof. Since non-strategyproof, deterministic mechanisms play an important role, one would like to apply the partial strategyproofness concept to study the incentive properties of these mechanisms as well. To this end, we consider a second source of randomness, namely the agents uncertainty about the reports from other agents. In this section, we exploit this uncertainty to apply the partial strategyproofness concept to deterministic mechanisms (and mechanisms that are not “random enough”).

Consider a setting \((N, M, q)\) and a deterministic mechanism \( \varphi \). Suppose that an agent \( i \in N \) does not know exactly what the other agents are going to report, but it has a probabilistic belief about these reports. Formally, \( i \) believes that \( P_{-i} \in \mathcal{P}^{N \setminus \{i\}} \) is drawn from a distribution \( \mathcal{P} \). Then from \( i \)'s perspective, the relevant mechanism is random and given by \( \varphi^P \) with

\[
\varphi^P(P_i) = \sum_{P_{-i} \in \mathcal{P}^{N \setminus \{i\}}} \varphi_i(P_i, P_{-i}) \cdot \mathbb{P}[P_{-i}],
\]

where \( \varphi^P(P_i) \) is simply \( i \)'s expected assignment vector from reporting \( P_i \).

Our main result in this section is a axiomatic characterization of the mechanisms \( \varphi \), deterministic or otherwise, that admit the construction of partially strategyproof random mechanisms \( \varphi^P \). For this, we introduce two new axioms.

**Axiom 4 (Monotonic).** A mechanism \( \varphi \) is monotonic if for any agent \( i \in N \), any preference profile \( (P_i, P_{-i}) \in \mathcal{P}^N \), and any preference order \( P'_{-i} \in N_{P_i} \) from the neighborhood of \( P_i \) with \( P_i : a_k > a_{k+1} \) and \( P'_{-i} : a_k > a_{k+1} \), the misreport \( P'_{-i} \) only increases \( i \)'s chances at \( a_k \) and only decreases \( i \)'s chances at \( a_{k+1} \) (i.e., \( \varphi_{i,jk}(P_i, P_{-i}) \geq \varphi_{i,jk}(P'_{i}, P_{-i}) \) and \( \varphi_{i,jk+1}(P_i, P_{-i}) \leq \varphi_{i,jk+1}(P'_{i}, P_{-i}) \)).

Monotonicity is a very natural requirement. It simply captures the intuition that bringing an object up in the preference report should not reduce the chances of obtaining
this object. Swap monotonicity implies monotonicity, while the converse does not hold (see Example 1).

**Axiom 5 (Plain).** A mechanism $\varphi$ is plain if for any agent $i \in N$, any preference orders $P_i \in \mathcal{P}$ and $P'_i \in N_{P_i}$ from the neighborhood of $P_i$ with $P_i : a_k > a_{k+1}$ and $P'_i : a_{k+1} > a_k$, the following holds: if $\varphi_i(P_i, P_{-i}) \neq \varphi_i(P'_i, P_{-i})$ for some $P_{-i} \in \mathcal{P}^{N\setminus\{i\}}$, then there exist $P^k_{-i}, P^{k+1}_{-i} \in \mathcal{P}^{N\setminus\{i\}}$ such that $\varphi_{i,a_k}(P_i, P^k_{-i}) \neq \varphi_{i,a_k}(P'_i, P^k_{-i})$ and $\varphi_{i,a_{k+1}}(P_i, P^{k+1}_{-i}) \neq \varphi_{i,a_{k+1}}(P'_i, P^{k+1}_{-i})$.

Intuitively, a mechanism is plain if the agent’s relative preferences matter for the assignment of the respective objects: if the agent’s assignment changes at all, then there must exist situations in which the two objects $a_k$ and $a_{k+1}$ are actually affected by this change. Again, swap monotonicity implies plainness, but the converse does not hold. Finally, we say that $\mathbb{P}$ has full support if $\mathbb{P}[P_{-i}] > 0$ for all $P_{-i} \in \mathcal{P}^{N\setminus\{i\}}$.

Our seventh main result characterizes the mechanisms, deterministic or not, for which uncertainty over the other agents’ reports induces a partially strategyproof mechanism from the perspective of each individual agent.

**Theorem 8.** A mechanism $\varphi$ is upper invariant, monotonic, and plain if and only if $\varphi^\mathbb{P}$ is upper invariant and swap monotonic for all distributions $\mathbb{P}$ with full support.

The formal proof is given in Appendix K. The most useful consequence of Theorem 8 is an insight about the strategic situation of agents whose uncertainty about the other agents’ reports is described by $\mathbb{P}$. Corollary 4 formalizes sufficient conditions under which these agents face a partially strategyproof mechanism.

**Corollary 4.** Given a setting $(N, M, \mathbf{q})$, a distribution $\mathbb{P}$ with full support, and a mechanism $\varphi$ that is upper invariant, monotonic and plain, there exists $r > 0$ such that $\varphi^\mathbb{P}$ is $r$-partially strategyproof.

This method of injecting randomness into the mechanism naturally extends the partial strategyproofness concept to mechanisms that are upper invariant, monotonic, and plain, even if they are deterministic. In Section 12, we use this extension to obtain partial strategyproofness of four mechanisms: first, for the school choice problem, these are the deterministic versions of the naïve and the adaptive Boston mechanism. Second, for the multi-unit assignment problem, the Probabilistic Serial mechanism and the HBS
### Table 1: Application of partial strategyproofness to popular and new mechanisms

<table>
<thead>
<tr>
<th>Mechanism</th>
<th>Source of Randomness</th>
<th>UI</th>
<th>PSP</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Serial Dictatorship</td>
<td>Priorities (single, uniform lottery)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Probabilistic Serial</td>
<td>Mechanism</td>
<td>✓</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Naïve Boston (random)</td>
<td>Priorities (single, uniform lottery)</td>
<td>✓</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Adaptive Boston (random)</td>
<td>Priorities (single, uniform lottery)</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>Rank Value</td>
<td>Mechanism (&amp; Preferences)</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Hybrids of RSD &amp; PS</td>
<td>Mixing (&amp; other)</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>Hybrids of RSD &amp; ABM</td>
<td>Mixing (&amp; other)</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>Hybrids of RSD &amp; NBM</td>
<td>Mixing (&amp; other)</td>
<td>✓</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Hybrids of RSD &amp; RV</td>
<td>Mixing (&amp; other)</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Naïve Boston (deterministic)</td>
<td>Preferences</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>Adaptive Boston (deterministic)</td>
<td>Preferences</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>Multi-unit PS</td>
<td>Preferences (&amp; Mechanism)</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>HBS Draft</td>
<td>Preferences (&amp; Priorities)</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
</tr>
</tbody>
</table>

Draft mechanism are not swap monotonic, but we show that both are upper invariant, monotonic, and plain.

### 12. Applications of Partial Strategyproofness

We now apply our new partial strategyproofness concept to a number of popular and new mechanisms. Table 1 provides an overview of our results.

#### 12.1. Random Serial Dictatorship

Random Serial Dictatorship is known to be strategyproof. Thus, it satisfies all three axioms and is 1-partially strategyproof for any setting.

#### 12.2. Probabilistic Serial

Upper invariance of Probabilistic Serial (PS) mechanism follows from Theorem 2 of Hashimoto et al. (2014). Our next Proposition 6 yields swap monotonicity.

**Proposition 6.** *PS is swap monotonic.*
Proof Outline (formal proof in Appendix L.1). We consider the times at which objects are exhausted under the Simultaneous Eating algorithm. Suppose an agent swaps two objects, e.g., from $P_i : a > b$ to $P'_i : b > a$. If anything changes about that agent’s assignment, the agent will now spend strictly more time consuming $b$. We show that by the time $b$ is exhausted, there will be strictly less of $a$ available or there will be strictly more competition at $a$ (relative to reporting $P_i : a > b$).

Since PS is known to be manipulable (as evident from the motivating example in the introduction), it is not strategyproof, and hence, by Theorem 1, it cannot be lower invariant in general. However, since it is swap monotonic and upper invariant, it is partially strategyproof by Theorem 2. This is a stronger statement than weak SD-strategyproofness, and it is also stronger than recent findings by Balbuzanov (2014), who shows that PS is convex strategyproof.

Kojima and Manea (2010) have shown that for a fixed number of objects $m$ and an agent $i$ with a fixed utility function $u_i$, $i$ will not want to misreport under PS if there are sufficiently many copies of each object. Since $u_i$ is fixed, this does not mean that PS becomes strategyproof in some finite setting. However, we conjecture that the result of Kojima and Manea (2010) can be strengthened in the following sense: for $m$ constant and $\min_{j \in M} q_j \to \infty$ we have $\rho_{(N,M,q)}(PS) \to 1$. We found numerical evidence that supports this conjecture. Figure 3 shows the development of the degree of strategyproofness $\rho_{(N,M,q)}(PS)$ as the size of settings increases. We observe that in both cases, the degree of strategyproofness increases and appears to converge to 1.
12.3. “Naïve” Boston Mechanism

We consider the Boston mechanism with priorities determined by a single, uniform lottery (Mennle and Seuken, 2015b). Intuitively, this mechanism 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 previous rounds (see (Mennle and Seuken, 2015b) for a formal proof). The Boston mechanism is, however, neither swap monotonic nor lower invariant, as the following Example 3 shows. Thus, the it is not $r$-partially strategyproof for any $r > 0$.

Example 3. Consider the setting in which 4 agents compete for 4 objects with unit capacities. Let the preferences be

\[
\begin{align*}
P_1 & : a > b > c > d, \\
P_2 & : a > c > b > d, \\
P_3, P_4 & : b > c > a > d.
\end{align*}
\]

Agent 1’s assignment vector is $(1/2, 0, 0, 1/2)$ for the objects $a, b, c, d$, respectively. If agent 1 swaps $b$ and $c$ in its report, its assignment vector change to $(1/2, 0, 1/4, 1/4)$. First, note that the probability for $b$ has not changed, but the overall assignment has, which violates swap monotonicity. Second, the assignment of $d$ has changed, even though it is in the lower contour set of $c$, which violates lower invariance.

The Boston mechanism is “naïve,” since it lets agents apply at their second, third, etc. choices, even if these have already been exhausted in previous rounds, such that agents “waste” rounds. Therefore, we refer to it as the naïve Boston mechanism (NBM).

12.4. Adaptive Boston Mechanism

Obvious manipulation strategies arise from this naïve approach of NBM: an agent who knows that its second choice will already be exhausted in the first round is better off ranking its third choice second, because this will increase its chances at all remaining objects in a stochastic dominance sense without forgoing any chances at its second choice object. If instead, the agent automatically “skipped” exhausted objects in the application process, this manipulation strategy would no longer be effective.
In (Mennle and Seuken, 2015b) we have shown that such an adaptive Boston mechanism (ABM) is swap monotonic and upper invariant, and thus partially strategyproof. However, since ABM is not strategyproof, it cannot be lower invariant. Miralles (2008) used simulations to study how unsophisticated (truthful) agents are disadvantaged under the naïve Boston mechanism and found evidence that such an adaptive correction may be attractive. Indeed, in (Mennle and Seuken, 2015b) we have proven that ABM has intermediate efficiency between RSD and NBM. Since RSD is strategyproof while NBM is not even weakly strategyproof, ABM offers a trade-off between strategyproofness and efficiency. Partial strategyproofness enabled a formal understanding of this trade-off.

Since ABM is partially strategyproof, we have computed its degree of strategyproofness for various setting, and the results are shown in Figure 4. We observe that $\rho_{(N,M,q)}(ABM)$ is significantly lower than $\rho_{(N,M,q)}(PS)$ (Figure 3).\footnote{We computed the degree of strategyproofness for small settings with $m \in \{3,4\}$ objects and $n \in \{3,6,9\}$ and $n \in \{4,8\}$ agents, respectively. The limitation to smaller settings is owed to the computational cost of determining the assignment under ABM, which becomes prohibitive in larger settings.} 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. Thus, ABM has intermediate incentive guarantees, which are stronger than those of NBM, but weaker than those of PS.

### 12.5. Rank Efficient Mechanisms

Featherstone (2011) introduced rank efficiency, a strict refinement of ex-post and ordinal efficiency. Rank efficiency often closely reflects the welfare criteria used in practical applications, e.g., in school choice (Mennle and Seuken, 2015b), or in the assignment...
of teachers to schools (Featherstone, 2011). However, no rank efficient mechanism is even weakly 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 Appendix L.2). Thus, they will not be partially strategyproof. Consequently, the attractive efficiency properties come at a price as such mechanisms will fail all of the axioms.

12.6. Hybrid Mechanisms

In (Mennle and Seuken, 2015a), we have shown how hybrid mechanisms can facilitate the trade-off between strategyproofness and efficiency for assignment mechanisms. The main idea of hybrid mechanisms is to consider convex combinations of two different mechanisms, one of which has good incentives while the other brings good efficiency properties. Under certain technical conditions, the resulting hybrid mechanisms are partially strategyproof, but can also improve efficiency beyond the ex-post efficiency of Random Serial Dictatorship. Furthermore, the trade-off is scalable in the sense that the mechanism designer can accept a lower degree of strategyproofness in exchange for more efficiency. Note that prior to the introduction of partial strategyproofness, no measure existed to evaluate the incentive properties of such hybrid mechanisms.

12.7. Deterministic Boston Mechanisms

Even though random tie-breaking plays an important role in many school choice mechanisms, many insights arise from the study of deterministic variants of these mechanisms. This is particularly true for the two variants of the Boston mechanism. We have already observed that when priorities are determined by a single, uniform lottery, then the adaptive Boston mechanism is partially strategyproof while the naïve Boston mechanism is not. However, if priorities are fixed and strict, both mechanisms are deterministic and manipulable. Consequently, both mechanisms have a degree of strategyproofness of 0.

In order to compare the two deterministic Boston mechanisms by their incentive properties, we must resort to a second source of randomness. In (Mennle and Seuken, 2015b), we have shown that both mechanisms are in fact upper invariant, monotonic, and plain. Thus, we can apply Theorem 8 and get that $\text{NBM}^P$ and $\text{ABM}^P$ are partially strategyproof for any distribution $P$ on the preference reports $P^{N\setminus\{i\}}$ with full support.
This allows a comparison of their incentive properties by the degree of strategyproofness, even if priorities are not random (or not “random enough”).

12.8. Multi-unit Assignment Mechanisms

The multi-unit assignment problem is an important extension of the assignment problem, where each agent receives a bundle of $K$ objects. For example, university students typically take several different courses each semester. If the number of participants in each course is limited, it may not be possible to accommodate every student’s favorite schedule. Instead, the university may use a mechanism to elicit preference over courses from the students and assign course schedules.

Under the *HBS Draft* mechanism, agents take turns to draw objects that they like. After all agents have taken one object, the order in which they draw their second object is reversed. A second mechanism for the multi-unit assignment problem is a straightforward extension of the *Probabilistic Serial* mechanism (Heo, 2014): as in the case of single-unit demand, agents collect probability shares of objects, and they move to their next preferred object when an object is exhausted. However, in the multi-unit version, they also move to their next preferred object when they have collected probability of 1 of some object.

Neither the HBS Draft nor the Probabilistic Serial mechanism are strategyproof for the multi-unit assignment problem. Furthermore, while both mechanisms are upper invariant and may involve randomization, they are not swap monotonic. However, they satisfy monotonicity and plainness.

**Proposition 7.** Given a setting $(N, M, q)$ with $\sum_{j \in M} q_j = n \cdot K$, the Probabilistic Serial mechanism and the HBS Draft mechanism for the $K$-unit assignment problem are upper invariant, monotonic, and plain. However, neither of them is swap monotonic.

The formal proof is given in Appendix L.3.

By Theorem 8, we can apply the partial strategyproofness concept to both mechanisms when the preference reports from the other agents are drawn from a distribution $P$ with full support. In future work, it will be interesting to study the relation of their respective degrees of strategyproofness numerically and analytically.

---

8 This mechanism is used for the assignment of courses at Harvard Business School (Budish and Cantillon, 2012).


13. Conclusion

In this paper, we have presented a new, axiomatic approach to studying the incentive properties of non-strategyproof assignment mechanisms.

First, we have shown that a mechanism is strategyproof if and only if it satisfies the three axioms swap monotonicity, upper invariance, and lower invariance. This illustrates why strategyproofness is such a strong requirement: if an agent swaps two consecutive objects, e.g., from $a > b$ to $b > a$, in its reported preference order, the only thing that a strategyproof mechanism can do (if anything) is to increase that agent’s assignment of $b$ and decrease its assignment of $a$ by the same amount.

Towards relaxing strategyproofness, we have shown that by dropping the least important axiom, lower invariance, the class of partially strategyproof mechanisms emerges: a mechanism is $r$-partially strategyproof if it makes truthful reporting a dominant strategy for agents who have sufficiently different values for different objects. These are precisely the agents whose vNM utility functions satisfy the $\text{URBI}(r)$ constraint. The set of partially strategyproof mechanisms is characterized by the axioms swap monotonicity and upper invariance. Consequently, under a swap monotonic, upper invariant mechanism, agents are best off when reporting truthfully if they are not too indifferent between different object. This provides an economic intuition for the partial strategyproofness concept and allows us to give honest and useful strategic advice to agents.

We have proven that the $\text{URBI}(r)$ domain restriction is maximal: for $r$-partially strategyproof mechanisms, $\text{URBI}(r)$ is the largest set of utility functions for which one can guarantee that truthful reporting is a dominant strategy without knowledge of further properties of the mechanism. This maximality result has allowed us to define the degree of strategyproofness, a meaningful measure for the incentive properties of non-strategyproof mechanisms. This measure is parametric, computable, and it is consistent with the method of comparing mechanisms by their vulnerability to manipulation (Pathak and Sönmez, 2013).

Next, we have established three appealing properties for partial strategyproofness: first, many important incentive concepts are defined via dominance notions, including strategyproofness, which is equivalent to stochastic dominance-strategyproofness. We have defined the notion of $r$-partial dominance, which is similar to stochastic dominance, but the influence of less preferred choices is discounted by the factor $r$. We have shown that
$r$-partial strategyproofness can alternatively be defined via this dominance notion. With an alternative definition that is independent of utility functions, partial strategyproofness integrates nicely with other dominance-based incentive concepts. Furthermore, this definition unlocks partial strategyproofness to algorithmic analysis. We have provided algorithms (Appendix H) that exploit the equivalence to verify partial strategyproofness and to compute the degree of strategyproofness of a mechanism (in any fixed setting).

Second, comparing partial strategyproofness to existing incentive concepts, we have established it as an intermediate concept: it is implied by strategyproofness, but in turn it implies many other incentive concepts, such as weak, convex, and approximate strategyproofness, as well as strategyproofness in the large if the degree of strategyproofness approaches 1 as the settings grow. Moreover, we have shown that strategyproofness and DL-strategyproofness are the upper and lower limit concepts, respectively. Thus, partial strategyproofness parametrizes the whole space between strategyproof and SL-strategyproof mechanisms.

Third, we have considered local sufficiency, a property of incentive concepts that is appealing from an axiomatic as well as from an algorithmic perspective. We have shown that $r$-local partial strategyproofness implies $r^2$-partial strategyproofness, and that the bound “2” in this implication is tight. Put differently, there exists no $\varepsilon > 0$ such that $r^{2-\varepsilon}$-partial strategyproofness can also be guaranteed. Prior work has shown that local sufficiency holds for both limit concepts, strategyproofness and DL-strategyproofness; our local sufficiency result for partial strategyproofness yields a unified proof for both statements.

Finally, we have applied the partial strategyproofness concept to gain a better understanding of the incentive properties of popular, non-strategyproof mechanisms. We have shown that the Probabilistic Serial mechanism is partially strategyproof, which is a significantly better description of its incentive properties than weak SD-strategyproofness. While the Boston mechanism in its naïve form (NBM) is not even weakly strategyproof, an adaptive variant (ABM) is in fact partially strategyproof, and we have presented numerical evidence that ABM has intermediate incentive guarantees, which are stronger than those of NBM, but weaker than those of PS. Rank Value mechanisms violate all three axioms and are therefore not partially strategyproof. These examples demonstrate that partial strategyproofness reflects our intuitive understanding of what it means for a non-strategyproof mechanism to have “good” incentive properties. We have argued
that partial strategyproofness can also be used to measure the incentive properties of new hybrid mechanisms, which enable a parametric trade-off between strategyproofness and efficiency of random assignment mechanisms. Last, we have demonstrated that the partial strategyproofness concept can be extended to deterministic mechanisms, such as the deterministic variants of the two Boston mechanisms, as well as the HBS Draft mechanism and the Probabilistic Serial mechanism for the multi-unit assignment problem.

Our new partial strategyproofness concept has an axiomatic motivation and, as we have shown in this paper, it is in multiple ways a powerful addition to the toolbox of the mechanism designer. We believe this will lead to new insights in the analysis of existing non-strategyproof mechanisms and facilitate the design of new ones.

References


Appendix

A. Probabilistic Serial Mechanism

The Probabilistic Serial mechanism (Bogomolnaia and Moulin, 2001) first collects the preference reports from all agents. Then it treats the objects as if they were divisible and uses the following simultaneous eating algorithm to determine a random assignment.

- At time 0, all agents begin consuming probability shares of their respective first choice objects at equal speeds.
- At some time, $t_1$ say, some object is exhausted. At this time, all the agents who were consuming this object before $t_1$, move on to their respective second choice.
- This process continues, i.e., every time an object is exhausted, the agents consuming this object move on to their respective next choices.
- If some agent’s next choice is also exhausted, it immediately continues to its next best available choice.
- At time 1 all agents will have collected a total of 1 in probability shares, from different objects.
- The entry $x_{i,j}$ in the resulting assignment matrix is the amount of $j$ that $i$ was able to consume during the process.
B. Continuation of Mechanism from Example 1

<table>
<thead>
<tr>
<th>Preference order</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P : a &gt; b &gt; \ldots )</td>
<td>0 1/2 0 1/2</td>
</tr>
<tr>
<td>( P : a &gt; c &gt; \ldots )</td>
<td>1/2 0 1/2 0</td>
</tr>
<tr>
<td>( P : a &gt; d &gt; \ldots )</td>
<td>1/2 0 1/2 0</td>
</tr>
<tr>
<td>( P : b &gt; \ldots )</td>
<td>0 1 0 0</td>
</tr>
<tr>
<td>( P : c &gt; \ldots )</td>
<td>0 0 1 0</td>
</tr>
<tr>
<td>( P : d &gt; \ldots )</td>
<td>0 0 0 1</td>
</tr>
</tbody>
</table>

Table 2: A mechanism (from the perspective of a single agent) that is swap monotonic, but manipulable in a first order-stochastic dominance sense.

C. Proof of Proposition 1

Proof of Proposition 1. A deterministic mechanism \( \varphi \) is strategyproof if and only if it is swap monotonic.

Let \( \varphi \) be strategyproof, and consider an agent \( i \in N \), a preference profile \( (P_i, P_{-i}) \in P^N \), and a misreport \( P'_i \in N_{P_i} \) from the neighborhood of \( P_i \) with \( P_i : a_k > a_{k+1} \), but \( P'_i : a_{k+1} > a_k \), such that \( \varphi_i(P_i, P_{-i}) \neq \varphi_i(P'_i, P_{-i}) \). Let \( j \in M \) be the object that \( i \) obtains with truthful reporting, and let \( j' \neq j \) be the object \( i \) obtains by reporting \( P'_i \). There are four cases:

Case \( j = a_k, j' = a_{k+1} \): This is consistent with swap monotonicity.

Case \( j = a_k, j' \neq a_{k+1} \): If \( P_i : j' > a_{k+1} \), then \( P_i : j > a_k \). Thus, \( i \) can obtain an object that it strictly prefers to \( a_k \) by reporting \( P'_i \), a contradiction to strategyproofness.

If \( P_i : a_{k+1} > j' \), then an agent with preference order \( P'_i \) could obtain an object (namely \( a_k \)) which it strictly prefers to \( j' \) by reporting \( P_i \) instead of reporting \( P'_i \) truthfully, again a contradiction.

Case \( j \neq a_k, j' = a_{k+1} \): This is symmetric to the previous case.
Case $j \neq a_k, j' \neq a_{k+1}$: If $P_i: j > j'$, then $P'_i: j > j'$ as well. Thus, an agent with preference order $P'_i$ could manipulate by reporting $P_i$. Conversely, if $P_i: j' > j$, then $i$ with preference order $P_i$ could manipulate. Thus, any strategyproof mechanism will be swap monotonic.

To see necessity, observe that swap monotonicity requires the assignment of $a_k$ and $a_{k+1}$ to change strictly (if there is any change at all under the swap). But if their assignment changes strictly, this must be a change from 0 to 1 or from 1 to 0. Thus, upon any swap, a swap monotonic deterministic mechanism can only change the assignment by assigning the object that has been brought up in the ranking instead of the object that has been brought down. Since any misreport can be decomposed into a sequence of swaps, the mechanism must be strategyproof.

\[ \square \]

D. Proof of Theorem 1

Proof of Theorem 1. A mechanism $\varphi$ is strategyproof if and only if it is swap monotonic, upper invariant, and lower invariant.

$\varphi$ strategyproof $\iff\varphi$ swap monotonic, upper invariant, and lower invariant: First, assume towards contradiction that $\varphi$ is not upper invariant. Then there exists some agent $i \in N$, some preference profile $P = (P_i, P_{-i}) \in \mathcal{P}^N$, and a misreport $P'_i \in N_{P_i}$ in the neighborhood of agent $i$’s true preference order such that

\[
P_i : a_1 > \ldots > a_{k-1} > a_k > a_{k+1} > \ldots > a_m, \quad P'_i : a_1 > \ldots > a_{k-1} > a_k > a_{k+1} > \ldots > a_m,
\]

and for some $l < k$ we have $\varphi_{i,a_l}(P_i, P_{-i}) \neq \varphi_{i,a_l}(P'_i, P_{-i})$. Without loss of generality, we can assume $\varphi_{i,a_l}(P_i, P_{-i}) < \varphi_{i,a_l}(P'_i, P_{-i})$ (otherwise, we invert the roles of $P_i$ and $P'_i$), and we let $l$ be the minimal rank for which this inequality is strict. This implies that $\varphi_i(P_i, P_{-i})$ does not even weakly stochastically dominate $\varphi_i(P'_i, P_{-i})$ at $P_i$, since

\[
\sum_{P_i:j > a_l} \varphi_{i,j}(P'_i, P_{-i}) > \sum_{P_i:j > a_l} \varphi_{i,j}(P_i, P_{-i}), \quad (21)
\]

a contradiction to strategyproofness of $\varphi$. 

55
Second, a similar argument yields lower invariance of \( \varphi \): again we find \( i, P = (P_i, P_{-i}), P'_i \in N_{P_i}, \) and \( l > k+1 \) such that without loss of generality \( \varphi_{i,a_l}(P_i, P_{-i}) > \varphi_{i,a_l}(P'_i, P_{-i}) \), and \( l \) is the largest rank for which this inequality is strict. Then \( \varphi_i(P_i, P_{-i}) \) does not even weakly stochastically dominate \( \varphi_i(P'_i, P_{-i}) \) at \( P_i \), since

\[
\sum_{P_{i,j} > a_l} \varphi_{i,j}(P'_i, P_{-i}) > \sum_{P_{i,j} > a_l} \varphi_{i,j}(P_i, P_{-i}). \tag{22}
\]

Third, we observe that upper and lower invariance of \( \varphi \) imply that for any swap the mechanism may only change agent \( i \)'s assignment for \( a_k \) and \( a_{k+1} \), and therefore

\[
\varphi_{i,a_k}(P_i, P_{-i}) - \varphi_{i,a_k}(P'_i, P_{-i}) = \varphi_{i,a_{k+1}}(P'_i, P_{-i}) - \varphi_{i,a_{k+1}}(P_i, P_{-i}). \tag{23}
\]

If the change in probability \( \varphi_{i,a_k}(P_i, P_{-i}) - \varphi_{i,a_k}(P'_i, P_{-i}) \) was negative, then a swap of \( a_k \) and \( a_{k+1} \) (from \( P_i \) to \( P'_i \)) would simply give agent \( i \) more probability of \( a_k \) and less probability for \( a_{k+1} \). But in this case, \( \varphi_i(P'_i, P_{-i}) \) would strictly stochastically dominate \( \varphi_i(P_i, P_{-i}) \) at \( P_i \), again a contradiction.

**\( \varphi \) swap monotonic, upper invariant, and lower invariant ⇒ \( \varphi \) strategyproof:** First, consider any local misreport, i.e., an agent \( i \in N \), a preference profile \( P = (P_i, P_{-i}) \in \mathcal{P}^N \), and a misreport \( P'_i \in N_{P_i} \) in the neighborhood of agent \( i \)'s true preference, such that \( P_i : a > b \), but \( P'_i : b > a \). Since \( \varphi \) satisfies all three axioms, we get that

- \( \varphi_{i,a}(P_i, P_{-i}) \geq \varphi_{i,a}(P'_i, P_{-i}) \),
- \( \varphi_{i,b}(P_i, P_{-i}) \leq \varphi_{i,b}(P'_i, P_{-i}) \),
- \( \varphi_{i,a}(P_i, P_{-i}) - \varphi_{i,a}(P'_i, P_{-i}) = \varphi_{i,b}(P'_i, P_{-i}) - \varphi_{i,b}(P_i, P_{-i}) \), and
- \( \varphi_{i,j}(P_i, P_{-i}) = \varphi_{i,j}(P'_i, P_{-i}) \) for all \( j \neq a, b \).

Consequently, \( \varphi_i(P_i, P_{-i}) \) stochastically dominates \( \varphi_i(P'_i, P_{-i}) \) at \( P_i \). This implies local strategyproofness of \( \varphi \), which in turn implies strategyproofness of \( \varphi \) (Carroll, 2012).

\[
\Box
\]
E. Proof of Theorem 2

Proof of 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 \(\varphi\) is swap monotonic and upper invariant.

Throughout the proof, we fix a setting \((N, M, q)\). First, we define

\[
\delta = \min \left\{ |\varphi_j(P_i, P_{-i}) - \varphi_j(P'_i, P_{-i})| : \forall i \in N, (P_i, P_{-i}) \in \mathcal{P}^N, P'_i \in \mathcal{P}, j \in M : |\varphi_j(P_i, P_{-i}) - \varphi_j(P'_i, P_{-i})| > 0 \right\}.
\] (24)

This is the smallest non-vanishing value by which the allocation of any object changes between two different types that any agent could report. Since \(N, M\), and \(\mathcal{P}\) are finite, \(\delta\) must be strictly positive (otherwise \(\varphi\) is constant).

\(\varphi\) upper invariant and swap monotonic \(\Rightarrow\) \(\varphi\) \(r\)-partially strategyproof for some \(r > 0\):

We must show that there exists \(r \in (0, 1]\) such that no agent with utility in \(\text{URBI}(r)\) can benefit from submitting a false report. Suppose, agent \(i\) with preference order

\[P_i : a_1 > \ldots > a_K > b > c_1 > \ldots c_L\]

is considering the misreport \(P'_i\), and without loss or generality let \(b\) be the most preferred object for which the allocation changes, i.e., for all \(k = 1, \ldots, K\)

\[
\varphi_{i,a_k}(P_i, P_{-i}) = \varphi_{i,a_k}(P'_i, P_{-i}),
\] (25)

\[
\varphi_{i,b}(P_i, P_{-i}) \neq \varphi_{i,b}(P'_i, P_{-i}),
\] (26)

Such an object must exist, because otherwise the allocations would be equal under both reports and \(P'_i\) would not be a beneficial misreport.

Claim 1. The allocation for \(b\) weakly decreases, i.e., \(\varphi_{i,b}(P_i, P_{-i}) \geq \varphi_{i,b}(P'_i, P_{-i})\).

Since the allocation for \(b\) must change by assumption, a weak decrease implies a strict decrease. Thus, reporting \(P'_i\) instead of \(P_i\) will necessarily decrease the probability that agent \(i\) gets object \(b\) by at least \(\delta\). None of the probabilities for the
objects $a_1, \ldots, a_K$ are affected. Hence, in the best case (for agent $i$), all remaining probability is concentrated on $c_1$, i.e., the maximum utility gain for agent $i$ is upper-bounded by

$$-\delta u_i(b) + u_i(c_1) - (1 - \delta) \min u_i. \quad (27)$$

The misreport $P'_i$ is guaranteed not to be beneficial if the value in (27) is less than or equal to 0, or equivalently,

$$u_i(c_1) - \min u_i < \delta(u_i(b) - \min(u_i)). \quad (28)$$

This sufficient condition is satisfied by all utilities in $\text{URBI}(r)$ with the choice of $r \leq \delta$. Consequently, the mechanism $\varphi$ is $r$-partially strategyproof for any $r \leq \delta$.

It remains to be proven that Claim 1 holds.

Proof of Claim 1: Consider a preference order

$$P : a_1 > \ldots > a_m.$$ 

A transition from $P$ to another preference order $P'$ is a finite sequence of preference orders that starts with $P$ and terminates with $P'$, and in each step the relative ranking of exactly two objects is inverted. Formally,

$$P_0, P_1, \ldots, P_{K-1}, P_K, \quad (29)$$

is a transition from $P$ to $P'$ if

- $P_0 = P$ and $P_K = P'$,
- for all $k \in \{0, \ldots, K - 1\}$ we have $P_k \in N_{P_{k+1}}$ and $P_{k+1} \in N_{P_k}$.

The canonical transition is a particular transition between two preference orders that is inspired the bubble-sort algorithm:

$P_0$: Set $P_0 = P$

$P_k$: Determine $P_k$ based on $P_{k-1}$ as follows:
• Let \( r \) be the rank where \( P_{k-1} \) and \( P' \) differ for the first time, i.e.,

\[
P_{k-1} : j_1 > \ldots > j_{r-1} > j_r > \ldots, \quad (30)
\]

\[
P' : j_1 > \ldots > j_{r-1} > c > \ldots, \quad (31)
\]

such that \( j_r \neq c \), and let \( c \) be the \( r \)th choice object under \( P' \).

• Find \( c \) in the ranking \( P_{k-1} \)

• Construct \( P_k \) by swapping \( c \) up one rank, i.e., if

\[
P_{k-1} : j_1 > \ldots > a > b > c > \ldots, \quad (32)
\]

then let

\[
P_k : j_1 > \ldots > a > c > b > \ldots. \quad (33)
\]

\( P_K \): Stop if \( P_k = P' \) for some \( k \), then set \( K = k \)

To prove the Claim, consider the first part of canonical transition from \( P'_i \) to \( P_i \):

\( a_1 \) is swapped with its predecessors (under \( P'_i \)) until it reaches its final position
at the front of the ranking (under \( P_i \)). With each swap the share of \( a_1 \) that the
agent receives from the respective misreport can only increase or stay constant,
because the mechanism is swap monotonic. On the other hand, once \( a_1 \) is at the
front of the ranking, the allocation of \( a_1 \) will remain unchanged during the rest
of the transition. This is because \( \varphi \) is upper invariant, i.e., no change of order
below the first position can affect the allocation of the first ranking object. Since
by assumption the allocation for \( a_1 \) did not change between \( P_i \) and \( P'_i \), none of
the swaps involving \( a_1 \) will have any effect on the allocation of \( a_1 \). But by swap
monotonicity this means that none of these swaps will have any effect on the
allocation at all.

Next consider the second part of transition, where \( a_2 \) is brought into second position
by swapping it upwards. The same argument can be applied to show that the
overall allocation must remain unchanged under any of the swaps involving \( a_2 \).
The same is true for \( a_3, \ldots, a_K \). Thus, we arrive at a preference order

\[
P''_i : a_1 > \ldots > a_K > c'_1 > \ldots c'_{L'} > b > c'_{L'+1} > \ldots > c'_{L}.
\]
Under $P_i''$ all of the objects $a_k$ are in the same positions as under $P_i$, $b$ holds some position below its rank in $P_i$, and some of the $c_l$ are ranking above $b$ (possibly in a different order). From the previous argument we know that the overall allocation is the same between $P_i'$ and $P_i''$. The next steps of the canonical transition will swap $b$ with its predecessors until it reaches its final position just below $a_K$ (as under $P_i$). During any of these swaps, the allocation for $b$ has to increase weakly by swap monotonicity of $\varphi$.

Finally, any subsequent swaps in the transition occur strictly below $b$, and therefore, the allocation for cannot change any more until $P_i$ is reached. Therefore, the allocation for $b$ weakly increases between $P_i'$ and $P_i$.

$\varphi$ r-partially strategyproof for some $r > 0 \Rightarrow \varphi$ upper invariant and swap monotonic:

**Upper invariance:** Suppose $\varphi$ is r-partially strategyproof for some fixed $r > 0$, i.e., no agent whose utility function satisfies URBI($r$) can benefit from misreporting.

We need to show that $\varphi$ is upper invariant. Suppose, the agent has preference order

$$P : \ldots > a > b > \ldots > c > d > \ldots$$

Assume towards contradiction that a swap of $c$ and $d$ changes the allocation of some object ranked before $c$, and let $a$ be the most preferred such object. Define $\delta$ as in (73), then without loss of generality the allocation of $a$ increases by at least $\delta$ due to this swap (if it decreases, consider the reverse swap). This means that by swapping $c$ and $d$, an agent with preference order $P$ can gain at least probability $\delta$ for object $a$. Because $a$ was the highest ranking object for which the allocation changed, the worst thing that can happen from the agent’s perspective is that it looses all of its chances to get $b$ and gets its last choice instead. Hence,

$$\delta u(a) - u(b) + (1 - \delta) \min u$$

is a lower bound for the benefit that the agent can have from swapping $c$ and
\(d\) in its report. This misreport is guaranteed to be strictly beneficial if the value in (27) is strictly positive, or equivalently, if

\[ u(b) - \min u < \delta(u(a) - \min u). \tag{35} \]

But for any \(r \in (0, 1]\), the set \(\text{URBI}(r)\) will contain a utility function satisfying this condition. This is a contradiction to the assumption that no agent with a utility function in \(\text{URBI}(r)\) will have a strictly beneficial manipulation. Consequently, \(\varphi\) must be upper invariant.

**Swap monotonicity** Suppose \(\varphi\) is \(r\)-partially strategyproof for some fixed \(r \in (0, 1]\). We know already that \(\varphi\) must be upper invariant. Towards contradiction, assume that upon a swap of two consecutive objects by some agent, the mechanism violates swap monotonicity, i.e., consider the preference order

\[ P : \ldots > a > b > c > \ldots > d > d' > \ldots, \]

and let \(a\) and \(b\) be the objects that change position under the swap. For \(\varphi\) to violate swap monotonicity, one of the following must hold:

1. the allocation of \(a\) increases,
2. the allocation of \(a\) remains constant, and the allocation of \(b\) increases,
3. the allocation of \(a\) remains constant, and the allocation of \(b\) decreases,
4. the allocations of \(a\) and \(b\) remain constant, but the allocation changes for some object \(d \neq a, b\),
5. the allocations of both \(a\) and \(b\) decrease.

We now consider each case separately and show that they all lead to contradictions.

Because of upper invariance, we know that the allocation of objects ranking above \(a\) cannot be affected. Therefore, in case 1, the agent can gain at least \(\delta\) probability of getting \(a\), with \(\delta\) defined as in (73). Then the worst thing (for the agent) that could happen is that it looses all its chances of getting
anything but its least preferred object. Hence,

\[ \delta u(a) - u(b) + (1 - \delta) \min u \]  

(36)
is a lower bound for the benefit the agent can have from swapping \(a\) and \(b\). But as in the proof of upper invariance, this leads to a contradiction.

In case 2, the agent gains at least \(\delta\) probability for \(b\), but may lose shares in the next lower ranking object \(c\). Again, the lower bound for the benefit is

\[ \delta u(b) - u(c) + (1 - \delta) \min u \]  

(37)

which leads to a contradiction. Note that if \(b\) is the lowest ranking object, this case is impossible.

Case 3 is symmetric to case 2, and we can consider the reverse swap instead.

In case 4, let \(d\) be the highest ranking object for which the allocation changes, which must lie after \(b\) because of upper invariance. Then without loss of generality, the agent can increase its chances of getting \(d\) by at least \(\delta\), but potentially loses all chances for the next lower ranking object \(d'\). This again leads to a contradiction.

For case 5, we consider the reverse swap, which is covered by case 1.

In conclusion, we have shown that none of the cases 1 through 5 can occur under a mechanism that is \(r\)-partially strategyproof. Therefore, the mechanism must satisfy strict swap monotonicity.

This concludes the proof of Theorem 2. \(\square\)
F. Omitted Proofs from Section 7

F.1. Proof of Theorem 3

Proof of Theorem 3. For any setting $(N, M, q)$ with $m \geq 3$, any bound $r \in (0, 1)$, and any utility function $\bar{u}_i$ (consistent with a preference order $P_i$) that violates $URBI(r)$, there exists a mechanism $\bar{\phi}$ such that

1. $\bar{\phi}$ is $r$-partially strategyproof, but
2. there exist preferences of the other agents $P_{-i}$ and a misreport $P'_i$ such that

$$\langle \bar{u}_i, \bar{\phi}_i(P_i, P_{-i}) - \bar{\phi}_i(P'_i, P_{-i}) \rangle < 0. \quad (38)$$

Furthermore, $\bar{\phi}$ can be chosen to satisfy anonymity.

By assumption, $\bar{u}$ violates $URBI(r)$. Thus, for some pair $a, b$ of consecutive objects in the preference order $P_i$ corresponding to $\bar{u}$ we have

$$\frac{\bar{u}(b) - \min \bar{u}}{\bar{u}(a) - \min \bar{u}} = \bar{r} > r. \quad (39)$$

Additionally, $b$ is not the last choice of $i$, since the constraint $\frac{0}{\bar{u}(a) - \min \bar{u}} \leq r$ is trivially satisfied. We now need to define the mechanism $\bar{\phi}$ that offers a manipulation to agent $i$ if its utility function is $\bar{u} \in U_{P_i}$, but would not offer any manipulation to agent $i$ if it has any utility satisfies $URBI(r)$ (and possibly a different preference order). For partial strategyproofness, an agent should not have a beneficial manipulation for any set of reports from the other agents. Thus, it suffices to specify $\bar{\phi}$ for a single set of reports $P_{-i}$, where only agent $i$ can vary its report. The allocation for $i$ must then be specified for any possible report $\hat{P}_i$ from $i$.

We define $\bar{\phi}_i(\cdot, P_{-i})$ as follows:

- For a report $\hat{P}_i$ with $a > b$,

$$\bar{\phi}_i(\hat{P}_i, P_{-i}) = \left( \frac{1}{m}, \ldots, \frac{1}{m} \right). \quad (40)$$
• For a report $\hat{P}_i$ with $b > a$, we adjust the original allocation by

$$
\tilde{\varphi}_i(\hat{P}_i, P_{\neg i}) (a) = \frac{1}{m} + \delta_a, \quad (41) \\
\tilde{\varphi}_i(\hat{P}_i, P_{\neg i}) (b) = \frac{1}{m} + \delta_b, \quad (42) \\
\tilde{\varphi}_i(\hat{P}_i, P_{\neg i}) (d) = \frac{1}{m} + \delta_d, \quad (43)
$$

where $\delta_a < 0$, $\delta_b > -\delta_a$, $\delta_d = -\delta_a - \delta_b < 0$. Here $d$ denotes the last choice. In case $a = d$, both $\delta_a$ and $\delta_d$ are added. Note that if the last object changes, the allocation for the new last object is decreased (by adding $\delta_d$), and the allocation of the previous last object is increased (by adding $\delta_d$).

This mechanism is upper invariant: swapping the order of $a$ and $b$ induces a change in the allocation of $a, b$, and the last object $d$. Therefore no higher ranking object is affected. Swapping the last and the second to last object also only changes the allocation for these two objects.

This mechanism is also swap monotonic: swapping $a$ and $b$ changes the allocation for both objects in the correct way, since $\delta_a < 0$, $\delta_b > 0$. Swapping the last to objects also changes the allocation appropriately, since $\delta_d < 0$. No other change of report changes the allocation.

Now we analyze the incentives for the different possible utility functions $i$ could have:

Case $u_i = \tilde{u} \in U_P$: In this case, the true preference order is $P_i : a > b$. Swapping $a$ and $b$ in its order is beneficial for $i$ if

$$
\delta_a \tilde{u}(a) + \delta_b \tilde{u}(b) + \delta_d \tilde{u}(d) = \delta_a (\tilde{u}(a) - \min \tilde{u}) + \delta_b (\tilde{u}(b) - \min \tilde{u}) > 0 \quad (44)
$$

$$
\Leftrightarrow \delta_a > -\delta_b \frac{\tilde{u}(b) - \min \tilde{u}}{\tilde{u}(a) - \min \tilde{u}}. \quad (45)
$$

(45) is satisfied if

$$
\delta_a > -\delta_b \cdot \tilde{r}, \quad (46)
$$

since $\frac{\tilde{u}(b) - \min \tilde{u}}{\tilde{u}(a) - \min \tilde{u}} = \tilde{r}$ by construction.

Case $u_i \in URBI(r), P_i : a > b$: Swapping $a$ and $b$ should no longer be beneficial for $i$. 

64
This is the case if
\[ \delta_a u_i(a) + \delta_b u_i(b) + \delta_d u_i(d) = \delta_a (u_i(a) - \min u_i) + \delta_b (u_i(b) - \min u_i) \leq 0 \quad (47) \]

\[ \iff \delta_a \leq -\frac{\delta_b u_i(b) - \min u_i}{u_i(a) - \min u_i}. \quad (48) \]

(48) is satisfied if
\[ \delta_a \leq - \cdot r, \quad (49) \]

since \( \frac{u_i(b) - \min u_i}{u_i(a) - \min u_i} \leq r \) by construction.

**Case \( u_i \in URBI(r), P_i : b > a \):** Swapping \( b \) and \( a \) to \( a > b \) should not be beneficial for \( i \). This is the case if
\[ \delta_b u_i(b) + \delta_a u_i(a) + \delta_d u_i(d) = \delta_a (u_i(a) - \min u_i) + \delta_b (u_i(b) - \min u_i) \geq 0 \quad (50) \]

\[ \iff \delta_a \geq -\frac{\delta_b u_i(b) - \min u_i}{u_i(a) - \min u_i}. \quad (51) \]

(51) is satisfied if
\[ \delta_a \geq - \frac{\delta_b}{r}, \quad (52) \]

since \( \frac{u_i(a) - \min u_i}{u_i(b) - \min u_i} \leq r \) for agents with \( b > a \) by construction.

This means that if \( \delta_a \) and \( \delta_b \) satisfy (46), (49), and (52), the mechanism \( \tilde{\phi} \) is in fact what we are looking for. Given some \( \delta_b > 0 \), we can choose \( \delta_a \) appropriately, since \( r < 1, r < \tilde{r} \), and
\[ -\delta_b \cdot \tilde{r} < -\delta_b \cdot r \iff r < \tilde{r}, \quad -\delta_b r > -\frac{\delta_b}{r} \iff r^2 < r. \quad (53) \]

Fixing the allocation for agent \( i \) in this manner, we can distribute all remaining probability for the objects evenly across all other agents, independent of their reports \( P_{-i} \). Then no other agent but \( i \) has any influence on their own allocation, i.e., the mechanism is constant from the perceptive from all other agents but \( i \). Finally, we can select one of the agents uniformly at random to take the role of \( i \). The resulting mechanism will be anonymous, \( r \)-partially strategyproof for all agents, but manipulable for any agent with utility function \( \tilde{u} \). \( \Box \)
F.2. Proof of Proposition 2

Definition 15 (Vulnerability to Manipulation of Random Assignment Mechanisms). For a given setting \((N, M, q)\) and two mechanisms \(\varphi, \psi\), \(\psi\) is as strongly manipulable as \(\varphi\) if for any agent \(i \in N\), any preference profile \((P_i, P_{-i}) \in \mathcal{P}^N\), and any utility function \(u_i \in U_P\), if there exists a misreport \(P'_i \in \mathcal{P}\) such that
\[
\langle u_i, \varphi_i(P_i, P_{-i}) \rangle < \langle u_i, \varphi_i(P'_i, P_{-i}) \rangle,
\]
then there exists a (possibly different) misreport \(P''_i \in \mathcal{P}\) such that
\[
\langle u_i, \psi_i(P_i, P_{-i}) \rangle < \langle u_i, \psi_i(P''_i, P_{-i}) \rangle.
\]
In words, any agent that has an incentive to manipulate \(\varphi\) would in the same situation also want to manipulate \(\psi\).

Proof of Proposition 2. For any setting \((N, M, q)\) and mechanisms \(\varphi\) and \(\psi\), the following hold:

1. If \(\psi\) is strongly as manipulable as \(\varphi\), then \(\rho_{(N, M, q)}(\varphi) \geq \rho_{(N, M, q)}(\psi)\).
2. If \(\rho_{(N, M, q)}(\varphi) > \rho_{(N, M, q)}(\psi)\) and \(\varphi\) and \(\psi\) are comparable by their vulnerability to manipulation, then \(\psi\) is strongly as manipulable as \(\varphi\).

To see 1., note that if \(\varphi\) is as strongly manipulable as \(\psi\) (in the sense of Definition 15.), then any agent who can manipulate \(\psi\) also finds a manipulation to \(\varphi\). Thus, the set of utilities on which \(\psi\) makes truthful reporting a dominant strategy cannot be larger than the set of utilities on which \(\varphi\) makes does the same. This in turn implies \(\rho_{(N, M, q)}(\varphi) \geq \rho_{(N, M, q)}(\psi)\).

For 2., observe that if \(\rho_{(N, M, q)}(\varphi) > \rho_{(N, M, q)}(\psi)\), then there exists a utility function \(\tilde{u}\) in URBI \((\rho_{(N, M, q)}(\varphi))\), which is not in URBI \((\rho_{(N, M, q)}(\psi))\), and for which \(\psi\) is manipulable, but \(\varphi\) is not. Thus, \(\varphi\) cannot be as strongly manipulable as \(\psi\), but the reverse is possible.

G. Proof of Theorem 4

Theorem 4 is a direct consequence of the following Lemma.
Lemma 1. Given a setting \((N, M, q)\), for any preference order \(P \in \mathcal{P}\), any assignment vectors \(x, y\), and any \(r \in [0, 1]\), the following are equivalent:

1. For all utility functions \(u \in U_P\) that satisfy \(\text{URBI}(r)\) we have \(\langle u, x - y \rangle \geq 0\),
2. \(x\) \(r\)-partially dominates \(y\) at \(P\).

Proof. Let \(P : a_1 > \ldots > a_m\).

2. \(\Rightarrow\) 1.: Assume towards contradiction that 2. holds, but for some utility \(u \in U_P\) satisfying \(\text{URBI}(r)\), we have

\[
\langle u, x - y \rangle = \sum_{i=1}^{m} u(a_i) \cdot (x_{a_i} - y_{a_i}) < 0. \tag{56}
\]

Without loss of generality, we can assume \(\min u = 0\). Let \(\delta_i = x_{a_i} - y_{a_i}\) and denote by

\[
S(k) = \sum_{i=1}^{k} u(a_i) \cdot (x_{a_i} - y_{a_i}) = \sum_{i=1}^{k} u(a_i) \delta_i. \tag{57}
\]

By assumption, \(S(m) < 0\), so there exists a smallest value \(K \in \{1, \ldots, m\}\) 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

\[
S(K) = \sum_{i=1}^{K} u(a_i) \delta_i = \sum_{i=1}^{K} u(a_i) u(a_K) \delta_i \tag{58}
\]

\[
= \left( \frac{S(K-1)}{u(a_K)} + \delta_K \right) u(a_K) \tag{59}
\]

\[
= \left( \frac{S(K-2)}{u(a_{K-1})} + \delta_{K-1} \right) \frac{u(a_{K-1})}{u(a_K)} u(a_K) + \delta_K \tag{60}
\]

\[
= \left( \left( \ldots \left( \delta_1 \frac{u(a_1)}{u(a_2)} + \delta_2 \right) \frac{u(a_2)}{u(a_3)} + \ldots \right) \frac{u(a_{K-1})}{u(a_K)} + \delta_K \right) u(a_K). \tag{61}
\]

Since \(u\) satisfies \(\text{URBI}(r)\), the fraction \(\frac{u(a_{K-1})}{u(a_K)}\) is lower-bounded by \(r^{-1}\). But since \(u(a_{K-1}) > 0\) and \(S(K-1) \geq 0\), we must have that

\[
\left( \frac{S(K-2)}{u(a_{K-1})} + \delta_{K-1} \right) \geq 0, \tag{62}
\]

67
and therefore, when replacing \[\frac{u(a_{K-1})}{u(a_K)}\] by \(r^{-1}\) in (60) we only make the term smaller. By the same argument, we can replace all the terms \(\frac{u(a_{k-1})}{u(a_k)}\) and obtain

\[
0 > S(K) \geq \left( \left( \frac{\delta_1}{r} + \frac{\delta_2}{r} + \ldots \right) \frac{1}{r} + \delta_K \right) u(a_K) = \frac{u(a_K)}{r^K} \sum_{l=1}^{K} r^l \cdot \delta_l. \tag{63}
\]

This is a contradiction to \(r\)-partial dominance, since by 2.,

\[
\sum_{l=1}^{K} r^l (x_{a_l} - y_{a_l}) = \sum_{l=1}^{K} r^l \delta_l \geq 0. \tag{64}
\]

1. \(\Rightarrow\) 2.: Assume towards contradiction that 1. holds, but \(x\) does not \(r\)-partially dominate \(y\) at \(P\), i.e., for some \(k \in \{1, \ldots, m\}\), we have

\[
\sum_{l=1}^{k} r^l \cdot \varphi_{i,a_l}(P_i, P_{-i}) < \sum_{l=1}^{k} r^l \cdot \varphi_{i,a_l}(P'_{i}, P_{-i}), \tag{65}
\]

and let \(k\) is the smallest rank for which inequality (65) is strict. Then the value

\[
\delta = \sum_{l=1}^{k} r^l \cdot (x_{i,a_l} - y_{i,a_l}), \tag{66}
\]

is strictly positive. Let \(u\) be a utility function that is consistent with \(P\) and has values

\[
u(a_l) = \begin{cases} Dr^l, & \text{if } l \leq k, \\ dr^l, & \text{if } k + 1 \leq l \leq m - 1, \\ 0, & l = m. \end{cases} \tag{67}
\]

This utility function satisfies URBI\((r)\) as long as \(D \geq d\). Furthermore, the difference in utility between \(x\) and \(y\) is

\[
\langle u, x_i - y_i \rangle = \sum_{l=1}^{m} u(a_l) \cdot (x_{i,a_l} - y_{i,a_l})
\]

\[
= D \sum_{l=1}^{k} r^l \cdot (x_{i,a_l} - y_{i,a_l}) + d \sum_{l=k+1}^{m-1} r^l \cdot (x_{i,a_l} - y_{i,a_l}) + D \delta - d. \tag{69}
\]

\[
\geq D \delta - d. \tag{70}
\]
Since $\delta > 0$, we can choose $D > \frac{\delta}{\delta}$ such that this change is strictly positive, a contradiction.

\[\square\]

**H. Algorithms for Verifying Partial Strategyproofness and Computing the Degree of Strategyproofness**

**ALGORITHM 1:** Verify $r$-partial strategyproofness

**Input:** setting $(N, M, \mathbf{q})$, mechanism $\varphi$, bound $r$

**Variables:** agent $i$, preference profile $(P_i, P_{-i})$, misreport $P_i'$, vector $\Delta$, counter $k$, choice function $ch : \{1, \ldots, m\} \to M$

begin
  for $i \in N, (P_i, P_{-i}) \in \mathcal{P}^N, P_i' \in \mathcal{P}$ do
    $\forall j \in M : \Delta_j \leftarrow \varphi_{i,j}(P_i, P_{-i}) - \varphi_{i,j}(P_i', P_{-i})$
    for $k \in \{1, \ldots, m\}$ do
      if $\sum_{l=1}^{k} r^l \cdot \Delta_{ch(l)} < 0$ then
        return false
      end
    end
  end
return true
end

**Remark 8.** Algorithm 1 is straightforward: it verifies all (finitely many) $r$-partial dominance constraints. Algorithm 2 optimistically sets $\rho$ to 1, then iterates through all partial dominance constraints. Each constraint is understood as a polynomial $f(r)$ in $r$ with $y$-intersect $f(0) = \Delta_{k_{\text{min}}} > 0$. If the current guess of $\rho$ is too high, one of the polynomials will have a negative value $f(\rho) < 0$ at $\rho$. In this case, the guess of $\rho$ is updated to the smallest positive real root of $f$. Note that since $f(0)$ is always strictly positive, this root exists, is positive, and can be found in polynomial time, e.g., using the LLL-algorithm (Lenstra, Lenstra and Lovász, 1982)
**Algorithm 2:** Compute $\rho_{(N,M,q)}(\varphi)$

**Input:** setting $(N,M,q)$, mechanism $\varphi$

**Variables:** agent $i$, preference profile $(P_i, P_{-i})$, misreport $P'_i$, vector $\Delta$, counters $k,k_{\text{min}}, l$, bound $\rho$, polynomial $f$, choice function $\text{ch} : \{1, \ldots, m\} \rightarrow M$

**begin**
\[ \rho \leftarrow 1 \]
for $i \in N, (P_i, P_{-i}) \in \mathcal{P}^N, P'_i \in \mathcal{P}$ do
\[ \forall j \in M : \Delta_j \leftarrow \varphi_{i,j}(P_i, P_{-i}) - \varphi_{i,j}(P'_i, P_{-i}) \]
\[ k_{\text{min}} \leftarrow \min\{k : \Delta_{\text{ch}(k)} \neq 0\} \]
for $k \in \{k_{\text{min}}, \ldots, m\}$ do
\[ f(r) \leftarrow \sum_{l=1}^{k} r^{l-k_{\text{min}}} \cdot \Delta_{\text{ch}(l)} \]
\[ \rho \leftarrow \min(\rho, \min\{r | r > 0 \text{ with } f(r) = 0\}) \]
end
return $\rho$
**end**

---

1. **Omitted Proofs from Section 9**

1.1. **Proof of Proposition 3**

*Proof of Proposition 3.* 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.

Convex strategyproofness (Balbuzanov, 2014) requires that for any agent $i$ and any preference order $P_i \in \mathcal{P}$, the set of utility functions $u_i \in U_{P_i}$ which make truthfully reporting $P_i$ is a dominant strategy for $I$ (independent of the other agents’ reports $P_{-i}$) is a non-empty (convex) subset of $U_{P_i}$. For any $r > 0$, this is precisely the set of utilities $\text{URBI}(r) \cap U_{P_i} \neq \emptyset$. Thus, partial strategyproofness implies convex strategyproofness in the sense that the set of utilities for which truthful reporting must be a dominant strategy is specified.

Balbuzanov (2014) gives an example of a mechanism that is weakly SD-strategyproof, but not convex strategyproof, and therefore, it cannot be partially strategyproof either. \(\square\)
I.2. Proof of Proposition 4

Proof of Proposition 4. 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.

Consider a fixed setting $(N, M, q)$ and fixed $r > 0$, and let $\varphi$ be a mechanism that is $r$-partially strategyproof in this setting. Let

$$\delta = \min \left\{ \left| \varphi_j(P_i, P_{-i}) - \varphi_j(P'_{i}, P_{-i}) \right| : \forall i \in N, (P_i, P_{-i}) \in \mathcal{P}^N, P'_{i} \in \mathcal{P}, j \in M : \left| \varphi_j(P_i, P_{-i}) - \varphi_j(P'_{i}, P_{-i}) \right| > 0 \right\}$$

(71)

be the smallest amount by which the probability for any object changes for any agent under any misreport. As in the proof of Theorem 2, this value must be strictly positive. From Theorem 2 we also know that $\varphi$ must be upper invariant and swap monotonic. Thus, under any manipulation, there is some highest ranking object $a$ for which the manipulating agent’s probability decreases strictly. The magnitude of this decrease is at least $\delta$, and for all more preferred objects than $a$ the probabilities do not change.

In the worst case (from the mechanism designer’s perspective), the manipulating agent can lose $\delta$ probability for $a$, but at the same time, it will convert probability for its last choice object ($d$, say) to probability for its next choice below $a$ ($b$, say). Setting $u(a) = 1$, $u(b)$ close to 1, and $u(c) = 0$, the gain from any manipulation is bounded by

$$u(b) - (1 - \delta)u(d) - \delta u(a) \leq 1 - \delta.$$  

(72)

Thus, the agent can improve its utility by at most $1 - \delta < 1$, i.e., $\varphi$ is $\varepsilon$-approximately strategyproof for $\varepsilon = 1 - \delta$.

To see that the converse may not hold we construct a mechanism that is $\varepsilon$-approximately strategyproof, but not partially strategyproof. Suppose, there are only 2 objects, $a$ and $b$. If the agent reports $P : a > b$, then the mechanism assigns $(1/2, 1/2)$ for $a$ and $b$, respectively. If instead the agent reports $P' : b > a$, the mechanism assigns $(1/2 + \varepsilon, 1/2 - \varepsilon)$. For $\varepsilon > 0$, this mechanism is manipulable in a stochastic dominance sense, and therefore not partially strategyproof. However, the maximal gain from manipulation is $\varepsilon$ if $u(a) = 1$ and $u(b) = 0$. Therefore, the mechanism $\varepsilon$-approximately strategyproof.
1.3. Proof of Theorem 5

Proof of Theorem 5. 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 DL-strategyproof.

Consider a fixed setting \((N, M, q)\) and fixed \(r > 0\), and let \(\varphi\) be a mechanism that is \(r\)-partially strategyproof in this setting. Let

\[
\delta = \min \left\{ |\varphi_j(P_i, P_{-i}) - \varphi_j(P'_i, P_{-i})| \mid \forall i \in N, (P_i, P_{-i}) \in \mathcal{P}^N, P'_i \in \mathcal{P}, j \in M : |\varphi_j(P_i, P_{-i}) - \varphi_j(P'_i, P_{-i})| > 0 \right\}
\]

be the smallest amount by which the probability for any object changes for any agent under any misreport. As in the proof of Theorem 2, this value must be strictly positive. From the Theorem 2 we also know that \(\varphi\) must be upper invariant and swap monotonic. Thus, under any manipulation, there is some highest ranking object \(a\) for which the manipulating agent’s probability decreases strictly. The magnitude of this decrease is at least \(\delta\), and for all objects that the agent prefers strictly to \(a\) the probabilities do not change. This immediately implies DL-strategyproofness of \(\varphi\).

To see that the other direction also holds, let \(\delta\) be defined as above. By DL-strategyproofness, the highest-ranking object for which there is any change in probability under any misreport must be assigned with lower probability under the misreport. Thus, we can proceed analogously to the proof of necessity in Theorem 2. \(\square\)

J. Omitted Proofs from Section 10

J.1. Proof of Theorem 6

Proof of Theorem 6. Given a setting \((N, M, q)\), if \(\varphi\) is \(r\)-locally partially strategyproof, then \(\varphi\) is \(r^2\)-partially strategyproof.

To prove this Theorem, we must verify that an \(r\)-locally partially strategyproof mechanism \(\varphi\) satisfies the conditions for \(r^2\)-partial strategyproofness, i.e., for any agent \(i \in N\), any preference profile \(P = (P_i, P_{-i}) \in \mathcal{P}^N\), any misreport \(P'_i \in \mathcal{P}\), and any utility function \(u_i \in U_P\) with \(u_i \in \text{URBI}(r^2)\) the inequality

\[
\langle u_i, \varphi_i(P_i, P_{-i}) - \varphi_i(P'_i, P_{-i}) \rangle \geq 0
\]
holds. Without loss of generality, we can assume that \( \min u = 0 \), since the manipulation incentives are exactly the same for an agent with utility function \( \tilde{u} = u - \min u \).

To simplify notation, we fix an arbitrary combination of agent, preference profile, misreport, and utility to satisfy these preconditions. We drop the index \( i \) on the preference orders, utility functions, and mechanism, and we omit the preferences of the other agents. With this simplification, inequality (74) becomes

\[
\langle u, \varphi(P) - \varphi(P') \rangle \geq 0. \tag{75}
\]

Recall that \( U_P \) denotes the set of utility functions that are consistent with \( P \), i.e.,

\[
U_P = \{ w : M \to \mathbb{R}_{\geq 0}^M \mid w \sim P \}, \tag{76}
\]

and \( U \) denotes the utility space, i.e., the union of all consistent utility functions

\[
U = \bigcup_{P \in \mathcal{P}} U_P. \tag{77}
\]

We say that a (utility) function \( w : M \to \mathbb{R}_{\geq 0}^M \) implies indifference between two different objects \( a, b \in M \) if \( w(a) = w(b) \), and we denote by \( W = \{ w : M \to \mathbb{R}_{\geq 0}^M \} \) the extended utility space, i.e., the set of all possible (utility) functions, including those that imply indifferences.

For the proof, we have fixed a preference order \( P \) and a consistent utility function \( u \in U_P \). Let \( v \) be a utility function that is consistent with the misreport \( P' \), and let

\[
\text{co}(u, v) = \{ u_\alpha = (1 - \alpha)u + \alpha v \mid \alpha \in [0, 1] \} \tag{78}
\]

be the convex line segment in \( W \) that connects \( u \) and \( v \). This line segment “starts” in \( U_P \), then (for increasing \( \alpha \)) traverses the extended utility space \( W \) and eventually “ends” at \( v \) in \( U_{P'} \). \( \text{co}(u, v) \) is said to pass a preference order \( P' \) if for some value \( \alpha \in [0, 1] \) we have that \( u_\alpha \) is consistent with \( P' \), or equivalently, if \( u_\alpha \in U_{P'} \). By construction, \( \text{co}(u, v) \) passes a sequence of preference orders \( P = P_0, P_1, \ldots, P_{K-1}, P_K = P' \) in this order, i.e., as \( \alpha \) increases, \( u_\alpha \) is first consistent with \( P_0 \), then with \( P_1 \), etc. until it is consistent with \( P_K = P' \). Note that intermittently, it is possible that \( u_\alpha \) is not consistent with any preference order as it may imply indifferences. By linearity we have that for any
two objects \( a, b \in M \) with \( u(a) > u(b) \) but \( v(a) < v(b) \), there exists a unique \( \alpha \in (0, 1) \) for which \( u_\alpha \) implies indifference between \( a \) and \( b \), for any smaller \( \alpha^- < \alpha \) we have \( u_{\alpha^-}(a) > u_{\alpha^-}(b) \), and for any larger \( \alpha^+ > \alpha \) we have \( u_{\alpha^+}(a) < u_{\alpha^+}(b) \).

We are now ready to formally define two important requirements:

- We say that \( \text{co}(u, v) \) makes no simultaneous transitions if for any three different objects \( a, b, c \in M \) we have
  \[
  \text{co}(u, v) \cap \{ w \in U \mid w(a) = w(b) = w(c) \} = \emptyset, \tag{79}
  \]
  i.e., for no value of the parameter \( \alpha \) does \( u_\alpha \) imply indifference between all three objects \( a, b, c \). Intuitively, this means that two consecutive preference orders \( P_k, P_{k+1} \) in the sequence \( (P_0, \ldots, P_K) \) differ by exactly one swap of two consecutive objects.

- We say that \( \text{co}(u, v) \) passes \( P_k \) in \( \text{URBI}(r) \) if it passes \( P_k \) and there exists some \( \alpha_k \in [0, 1] \) such that \( u_{\alpha_k} \in U_{P_k} \cap \text{URBI}(r) \). This means that the line segment contains at least one utility function that is consistent with \( P_k \) and in addition satisfies the \( \text{URBI}(r) \)-constraint.

Recall that \( P \) is a preference order, \( u \) a utility consistent with \( P \) that satisfies \( \text{URBI}(r^2) \), and the mechanism \( \varphi \) is \( r \)-locally partially strategyproof.

**Claim 2.** There exists \( v \in U_P \cap \text{URBI}(r) \) such that

1. \( \text{co}(u, v) \) makes no simultaneous transitions,
2. if \( \text{co}(u, v) \) passes a preference order \( P'' \), then it passes \( P'' \) in \( \text{URBI}(r) \).

Using Claim 2, we can now show the inequality

\[
\langle u, \varphi(P) - \varphi(P') \rangle \geq 0. \tag{80}
\]

We will show this by writing the left side as a telescoping sum over local incentive constraints, where all but the first and the last term cancel out, such that it collapses to yield the inequality. This idea is inspired by the proof of local sufficiency for strategyproofness in (Carroll, 2012).
Consider the utility function $v$ constructed in Claim 2 and the convex line segment $co(u,v)$. Let $\alpha_0 = 0$, $\alpha_K = 1$, and for each $k \in \{0, \ldots, K\}$ let $\alpha_k$ be the parameters for which $u_{\alpha_k} \in U_{P_k} \cap \text{URBI}(r)$, which exist by Claim 2.2. For any $k \in \{0, \ldots, K - 1\}$ we know that the preference order $P_k$ and $P_{k+1}$ are neighbors of each other, i.e., $P_k \in N_{P_{k+1}}$ and $P_{k+1} \in N_{P_k}$ (by Claim 2.1). Thus, by $r$-local partial strategyproofness of $\varphi$ we obtain

$$\langle u_{\alpha_k}, \varphi(P_k) - \varphi(P_{k+1}) \rangle \geq 0$$

(81)

and

$$\langle u_{\alpha_{k+1}}, \varphi(P_{k+1}) - \varphi(P_k) \rangle \geq 0.$$  

(82)

Multiplication by $\alpha_k$ and $-\alpha_{k+1}$, respectively, and then adding both inequalities yields

$$\langle \alpha_k u_{\alpha_{k+1}} - \alpha_{k+1} u_{\alpha_k}, \varphi(P_{k+1}) - \varphi(P_k) \rangle \geq 0.$$ 

(83)

Now, observe that $\alpha_k u_{\alpha_{k+1}} - \alpha_{k+1} u_{\alpha_k} = (\alpha_k - \alpha_{k+1}) u$, and therefore

$$\langle u, \varphi(P_{k+1}) - \varphi(P_k) \rangle \geq 0$$ 

(84)

for all $k \in \{0, \ldots, K - 1\}$. Summing over all $k$, we get

$$\langle u, \varphi(P) - \varphi(P') \rangle = \sum_{k=0}^{K-1} \langle u, \varphi(P_{k+1}) - \varphi(P_k) \rangle \geq 0.$$ 

(85)

We now proceed to prove Claim 2.

Proof of Claim 2. The proof for the existence of $v$ is constructive. Recall that for a preference order $P$ the rank of an object $j$ under $P'$ is the position that $a$ holds in the ranking, i.e.,

$$\text{rank}_{P'}(j) = \# \{ j \in M \mid P' : j > a \} + 1.$$ 

(86)

Define $v : M \to \mathbb{R}_{\geq 0}^M$ by setting

$$v(j) = C^{m-\text{rank}_{P'}(j)}$$ 

(87)

for any $j \in M$. If $C > 1$, then $v \in U_{P^*}$. Furthermore, for sufficiently large $C$, $v \in \text{URBI}(r)$,
since for any $a, b \in M$ with $P' : a > b$

$$\frac{v(b) - \min v}{v(a) - \min v} = \frac{C^{\text{rank}_{P'}(b)} - 1}{C^{\text{rank}_{P}(a)} - 1} = o(1/C).$$

(88)

It remains to be shown that for sufficiently large $C$ 1 and 2 hold.

To prove both statements, we require the concept of the canonical transitions (same as in Claim 1 in the proof of Theorem 2). A transition is a finite sequence of preference orders that starts and terminates with given preference orders and in each step the relative ranking of exactly two objects is inverted. Formally,

$$P_0, P_1, \ldots, P_{K-1}, P_K,$$

(89)

is a transition from $P$ to $P'$ if

- $P_0 = P$ and $P_K = P'$,
- for all $k \in \{0, \ldots, K - 1\}$ we have $P_k \in N_{P_{k+1}}$ and $P_{k+1} \in N_{P_k}$.

The canonical transition is a particular transition between two preference orders that is inspired the bubble-sort algorithm:

$P_0$: Set $P_0 = P$

$P_k$: Determine $P_k$ based on $P_{k-1}$ as follows:

- Let $r$ be the rank where $P_{k-1}$ and $P'$ differ for the first time, i.e.,

$$P_{k-1} : \; j_1 > \ldots > j_{r-1} > j_r > \ldots,$$

(90)

$$P' : \; j_1 > \ldots > j_{r-1} > c > \ldots,$$

(91)

such that $j_r \neq c$, and let $c$ be the $r$th choice object under $P'$.

- Find $c$ in the ranking $P_{k-1}$

- Construct $P_k$ by swapping $c$ up one rank, i.e., if

$$P_{k-1} : \; j_1 > \ldots > a > b > c > \ldots,$$

(92)

then let $P_k : \; j_1 > \ldots > a > c > b > \ldots$.
$P_K$: Stop if $P_k = P'$ for some $k$, then set $K = k$

Besides the canonical transition, we formalize *transition times*. Suppose that for two objects $a, b \in M$, $P : a > b$, but $P' : b > a$, such that $u(a) > u(b)$, but $v(a) < v(b)$. Recall that in this case, there exists a unique parameter $\alpha$ for which $u_{\alpha}(a) = u_{\alpha}(b)$, and for any smaller $\alpha^- < \alpha$ we have $u_{\alpha^-}(a) > u_{\alpha^-}(b)$, and for any larger $\alpha^+ > \alpha$ we have $u_{\alpha^+}(a) < u_{\alpha^+}(b)$. The line segment $co(u, v)$ “pierces” the hyperplane of indifference between $a$ and $b$ at the point $u_{\alpha}$, i.e., it transitions from preference orders that prefer $a$ to preference orders that prefer $b$ to $a$. Formally, the *transition time* $\alpha_{a, b, 1}$ is the parameter for which $u_{\alpha_{a,b,1}}(a) = u_{\alpha_{a,b,1}}(b)$. Extending this notation, we define $\alpha(a, b, r)$ as the first time when $u_{\alpha}$ violates the URBI($r$) constraint for $a > b$, i.e.,

$$\alpha(a, b, r) = \inf \left\{ \alpha \in [0, 1] \mid \frac{u_{\alpha}(b) - \min u_{\alpha}}{u_{\alpha}(a) - \min u_{\alpha}} > r \right\},$$

and $\alpha(b, a, r)$ as the last time when $u_{\alpha}$ violates the URBI($r$) constraint for $b > a$, i.e.,

$$\alpha(a, b, r) = \sup \left\{ \alpha \in [0, 1] \mid \frac{u_{\alpha}(a) - \min u_{\alpha}}{u_{\alpha}(b) - \min u_{\alpha}} > r \right\}.$$  

Obviously,

$$\alpha(a, b, r) < \alpha(a, b, 1) < \alpha(b, a, r),$$

i.e., as $\alpha$ increases, $u_{\alpha}$ violates URBI($r$) for $a > b$ at some time, then subsequently it transitions from $a > b$ to $b > a$, and finally it no longer violates the URBI($r$) constraint for $b > a$.

We are now ready to formulate Claims 3, 4, and 5, which are needed to establish 1 (no simultaneous transitions) and 2 (passing all preference orders in URBI($r$)), respectively, and the fact that only pairs of objects are relevant that rank differently under $P$ and $P'$.

**Claim 3.** For sufficiently large $C$, $co(u, v)$ induces the canonical transition

$$P_0 = P, P_1, \ldots, P_{K-1}, P_K = P'. $$

**Claim 4.** For sufficiently large $C$, if $\alpha(a, b, 1) < \alpha(c, d, 1)$, then

$$\alpha(a, b, r) \leq \alpha(c, d, r).$$
Claim 5. If $P : a > b$ and $P' : a > b$ and $u, v \in URBI(r)$, then for all $\alpha \in [0,1]$,
\[
\frac{u_\alpha(b) - \min u_\alpha}{u_\alpha(a) - \min u_\alpha} \leq r. \tag{99}
\]

Since $co(u, v)$ induces a transition by Claim 3, we already know that for all pairs
$(a, b) \neq (c, d)$ we have $\alpha(a, b, 1) \neq \alpha(c, d, 1)$. Thus, $co(u, v)$ makes no simultaneous transitions.

If $a$ is preferred to $b$ under both $P$ and $P'$, then by Claim 5 the $URBI(r)$ constraint
for $a$ over $b$ is satisfied for any $\alpha$. Suppose now that $P : a > b$, $P : c > d$, $P' : b > a$, $P' : d > c$, and $\alpha(a, b, 1) < \alpha(c, d, 1)$. Then $co(u, v)$ “enters” a new set of consistent
utility functions $U_{P_k}$ at time $\alpha(a, b, 1)$, where $P_k$ differs from $P_{k-1}$ by a swap of $a$ and $b$, and it “leaves” $U_{P_k}$ at time $\alpha(c, d, 1)$, where $P_k$ differs from $P_{k+1}$ by a swap of $c$ and $d$. In
this case the $URBI(r)$ constraint for $b$ over $a$ is satisfied after time $\alpha(a, b, r) > \alpha(a, b, 1)$, and the $URBI(r)$ constraint for $c$ over $d$ is satisfied before time $\alpha(c, d, r) < \alpha(c, d, 1)$. Claim 4 yields the constraint for $c$ over $d$ holds “long enough” for the constraint for $b$ over $a$ to be restored. Thus, at any time $\alpha_k \in [\alpha(a, b, r), \alpha(c, d, r)] \neq \emptyset$, both constraints
are satisfied. Iterated application of this argument yields that for any $k \in \{0, \ldots, K\}$, there exists some $\alpha_k$ for which $u_{\alpha_k}$ satisfies $URBI(r)$ with respect to preference order $P_k$.

This concludes the proof of Claim 2.

We now provide the proofs of Claims 3 and 4.

Proof of Claim 3. First, we formulate an equivalent condition for $co(u, v)$ to induce the
canonical condition in terms of transition times $\alpha(a, b, 1)$.

Claim 6. The following are equivalent:

1. $co(u, v)$ induces the canonical transition

\[
P_0 = P, P_1, \ldots, P_{K-1}, P_K = P'. \tag{100}
\]

2. For any $a, b, c, d \in M$ with $P : a > b$, $P : c > d$, $P' : b > a$, $P' : d > c$,
   i. if $P' : b > d$, then $\alpha(a, b, 1) < \alpha(c, d, 1)$,
   ii. if $b = d$ and $P : c > a$, then $\alpha(a, b, 1) < \alpha(c, d, 1)$. 

78
Proof of Claim 6. First, we show sufficiency (“⇒”). To see that 2i holds, observe that since $P' : b > d$, $b$ will be “brought up” by bubble sort before $d$ is ever swapped up against another object. Since $P : c > d$, the swap of $c \leftrightarrow d$ is such a swap, and therefore, it has to occur after the swap $a \leftrightarrow b$. 2ii follows by observing that from $b = d$ and $P : c > a$ we get that $P : c > a > b$, but ultimately $P' : b > (a, c)$. The bubble sort algorithm will bring $b$ up by swapping it with $a$ before it swaps $b$ and $c$.

To see necessity (“⇐”), let $(a \leftrightarrow b)$ and $(c \leftrightarrow d)$ be two swaps that occur at $\alpha(a, b, 1)$ and $\alpha(c, d, 1)$, respectively. If $P' : b > d$, then 2i implies that $(a \leftrightarrow b)$ occurs before $(c \leftrightarrow d)$, which is consistent with the canonical transition. By symmetry, the case $P' : d > b$ also follows. Next, observe that any case not covered by this argument involves identity of $b$ and $d$, i.e., $b = d$. If $a = c$ as well, then there is nothing to show, so assume $P : a > c$, where 2ii implies the correct behavior. The last remaining case where $b = d$ and $P : c > a$ follows again by symmetry.

We now verify that the sequence of types through which $co(u, v)$ passes is indeed a canonical transition. Let $a, b, c, d \in M$ be such that $P : a > b$, $P : c > d$, $P' : b > a$, $P' : d > c$, and either $P' : b > d$ (as in 2i of Claim 6) or $b = d$ and $P : c > a$ (as in 2ii of Claim 6). We can write

$$\alpha(a, b, 1) = \frac{u(a) - u(b)}{u(a) - u(b) + v(b) - v(a)}$$
and

$$\alpha(c, d, 1) = \frac{u(c) - u(d)}{u(c) - u(d) + v(d) - v(c)},$$

and we need to show that

$$\alpha(a, b, 1) < \alpha(c, d, 1)$$  \(102\)

$$\Leftrightarrow (u(a) - u(b)) (u(c) - u(d) + v(d) - v(c))$$  \(103\)

$$< (u(c) - u(d)) (u(a) - u(b) + v(b) - v(a))$$  \(104\)

$$\Leftrightarrow (u(a) - u(b)) (v(d) - v(c)) < (u(c) - u(d)) (v(b) - v(a))$$  \(105\)

$$\Leftrightarrow \frac{u(a) - u(b)}{u(c) - u(d)} < \frac{v(d) - v(c)}{v(b) - v(a)}.$$  \(106\)

If $P : b > d$, the left side of (106) grows faster than $C$, i.e.,

$$\frac{v(d) - v(c)}{v(b) - v(a)} = \frac{C_{\text{m-rank}^P(d)} - C_{\text{m-rank}^P(c)}}{C_{\text{m-rank}^P(b)} - C_{\text{m-rank}^P(a)}} = \omega(C),$$  \(107\)
since $\text{rank}_{P'}(b) < \text{rank}_{P'}(d), \text{rank}_{P'}(b) < \text{rank}_{P'}(a),$ and $\text{rank}_{P'}(d) < \text{rank}_{P'}(c)$. Similarly, if $b = d$ and $P : c > a$, we obtain that

$$\frac{v(d) - v(c)}{v(b) - v(a)} = \frac{C^{m - \text{rank}_{P'}(b)} - C^{m - \text{rank}_{P'}(c)}}{C^{m - \text{rank}_{P'}(b)} - C^{m - \text{rank}_{P'}(a)}} = \omega(C).$$

(108)

Since for sufficiently large $C$, the right side in (106) is not small, we can ensure that $\alpha(a, b, 1) < \alpha(c, d, 1)$ whenever the conditions of 2i or 2ii of Claim 6 are satisfied.

Proof of Claim 4. First we define a “conservative estimate” for the violation times $\alpha(a, b, r)$ and $\alpha(c, d, r)$. Let

$$s(a, b, \alpha) = \frac{u_{\alpha}(b)}{u_{\alpha}(a)}$$

(109)

and observe that $s(a, b, \alpha)$ is continuous and strictly monotone in $\alpha$ and $s(a, b, \alpha(a, b, 1)) = 1$. Thus, we can define the inverse $\alpha(a, b, s)$ for which $s(a, b, \alpha(a, b, s)) = s$ for any value of $s$ that is attained by $s(a, b, \alpha)$. In particular for $\alpha = 0$, $s(a, b, 0) = \frac{u(b)}{u(a)} \leq r$ and for $\alpha = 1$, $s(a, b, 1) = \frac{v(b)}{v(a)} > \frac{1}{r}$, so $\alpha(a, b, s)$ is well-defined for all values $s \in [r, \frac{1}{r}]$. In fact, we can solve

$$\frac{u_{\alpha(a, b, s)}(b)}{u_{\alpha(a, b, s)}(a)} = s$$

(110)

for $\alpha(a, b, s)$ and obtain the expression

$$\alpha(a, b, s) = \frac{su(a) - u(b)}{su(a) - u(b) + v(b) - sv(a)}.$$  

(111)

Using $\min u_{\alpha} \geq 0$,

$$s(a, b, \alpha) = \frac{u_{\alpha}(b)}{u_{\alpha}(a)} \leq r$$

(112)

implies

$$\frac{u_{\alpha}(b) - \min u_{\alpha}}{u_{\alpha}(b) - \min u_{\alpha}} \leq r,$$

(113)

and therefore,

$$\overline{\alpha}(a, b, r) \leq \alpha(b, a, r) \text{ and } \alpha(c, d, r) \leq \alpha(c, d, r).$$

(114)

We now show that for sufficiently large $C$, $\alpha(b, a, r) \leq \alpha(c, d, r)$ holds. Recall that we are considering objects $a, b, c, d \in M$, where $P : a > b$, $P : c > d$, $P' : b > a$, and $P' : d > c,$
so that the required inequality can be rewritten equivalently as

$$\alpha(b, a, r) \leq \alpha(c, d, r) \iff \frac{u(a) - ru(b)}{ru(c) - u(d)} \leq \frac{rv(b) - v(a)}{v(d) - rv(c)}. \quad (115)$$

By Claim 3, $co(u, v)$ induces the canonical transition for sufficiently large $C$. Thus, by Claim 6, $\alpha(a, b, 1) < \alpha(c, d, 1)$ holds if

i. either $P^r : b > d$,

ii. or $b = d$ and $P : c > a$.

In case i, we observe that the left side of (115) is constant, but the right side grows for growing $C$, i.e., it is in $\omega(C)$. Therefore, (115) is ultimately satisfied for sufficiently large $C$.

In case ii, the right side converges to $r$ (from below) as $C$ becomes large. Thus, it suffices to verify

$$\frac{u(a) - ru(b)}{ru(c) - u(d)} \leq r \quad (116)$$

$$\iff u(a) - ru(b) \leq r^2 u(c) - ru(b) \quad (117)$$

$$\iff 0 \leq r^2 u(c) - u(a). \quad (118)$$

Using the assumption that $u$ satisfies URBI($r^2$), min $u = 0$, and $P : c > a$ we get that

$$\frac{u(c)}{u(a)} \leq r^2 \iff r^2 u(c) - u(a) \geq 0. \quad (119)$$

\[\square\]

\[\square\]

**J.2. Proof of Theorem 7**

**Proof of Theorem 7.** Given a setting $(N, M, q)$ with $m \geq 4$ objects, for any $\varepsilon > 0$ there exists a bound $r \in (0, 1)$ and a mechanism $\varphi$ such that

1. $\varphi$ is $r$-locally partially strategyproof,

2. $\varphi$ is not $r^{2-\varepsilon}$-partially strategyproof.
Consider a mechanism $\varphi$ that selects the following assignments:

\[
\begin{align*}
\varphi(a > \ldots) &= (\alpha, 0, 0, 1 - \alpha), \\
\varphi(b > \ldots) &= (0, \beta, 0, 1 - \beta), \\
\varphi(d > \ldots) &= (0, 0, 0, 1), \\
\varphi(c > d > \ldots) &= (0, 0, \gamma_c, 1 - \gamma_c), \\
\varphi(c > a > d > b) &= (1 - \gamma_c - \gamma_d, 0, \gamma_c, \gamma_d), \\
\varphi(c > b > \ldots) = \varphi(c > a > b > d) &= (1 - \gamma_c - \gamma_d, \gamma_d, \gamma_c, 0)
\end{align*}
\]

for the objects $a, b, c, d$, respectively, where

\[
\begin{align*}
\alpha, \beta, \gamma_c, \gamma_d &\in [0, 1], \\
s &= \frac{1}{r}, \\
\beta &= s\alpha, \\
\gamma_c &= \frac{(1 - \alpha)}{(s - 1) (s (s + 1) - 1)}, \\
\gamma_d &= \frac{s (s + 1) (1 - \alpha)}{s (s + 1) - 1}.
\end{align*}
\]

Observe that $\varphi$ is entirely specified by the values of $r$ and $\alpha$. We will now show that for sufficiently small $r > 0$ we can chose $\alpha$ such that

1. $\varphi$ is feasible,
2. $\varphi$ is $r$-locally partially strategyproof,
3. but not $r^{2-\varepsilon}$-partially strategyproof.

First, we verify 1. that $\varphi$ is feasible.

**Claim 7.** For $s > 1$, $\varphi$ is feasible if and only if $\alpha \in \left[\frac{s}{s^2 - s + 1}, \frac{1}{s}\right]$.

**Proof of Claim 7.** Note that for $s > 1$ and $\alpha < 1$, $\gamma_c$ and $\gamma_d$ are positive. We must ensure that $\beta = s\alpha \leq 1$, which is the case if and only if $\alpha \leq \frac{1}{s}$. Next, we give a condition for $\gamma_c + \gamma_d \leq 1$, which in turn implies feasibility of the mechanism. This inequality holds if and only if $\alpha \geq \frac{s}{s^2 - s + 1}$. Observing that $\frac{1}{s} > \frac{s}{s^2 - s + 1}$ for $s > 1$, we have that the mechanism $\varphi$ is feasible if and only if $\alpha \in \left[\frac{s}{s^2 - s + 1}, \frac{1}{s}\right] \neq \emptyset$. 

82
Second, we give equivalent conditions for \( r \)-local partial strategyproofness of \( \varphi \), i.e., 2.

**Claim 8.** For sufficiently small \( r \), the following are equivalent:

- \( \varphi \) is feasible and \( r \)-locally partially strategyproof,
- \( \alpha \in I_s = \left[ \frac{s^3-s^2}{s^3+2s^2-s-1}, \frac{s^2+1}{s^3+s^2-s+1} \right] \).

Furthermore, for sufficiently small \( r > 0 \), \( I_s \neq \emptyset \).

**Proof of Claim 8.** We use Lemma 1 to establish \( r \)-partial dominance for any manipulation by just a swap, which in turn yields \( r \)-local partial strategyproofness. We only need to consider those swaps that lead to a change of the assignment, otherwise there is nothing to show. In the following, \( \delta_k \) denotes the adjusted \( k \)th partial sum, i.e., for \( P: j_1 > \ldots > j_m \),

\[
\delta_k = \sum_{l=1}^{k} s^{k-l} (\varphi_{ji}(P) - \varphi_{ji}(P')) = r^{-k} \left( \sum_{l=1}^{k} r^l (\varphi_{ji}(P) - \varphi_{ji}(P')) \right). \tag{131}
\]

Observe that positivity of \( \delta_1, \delta_2, \delta_3 \) is equivalent to \( r \)-partial dominance of \( \varphi_{ji}(P) \) over \( \varphi_{ji}(P') \) at \( P \) by Lemma 1. Table 3 lists all the cases we need to consider.

1. \( a > b > \ldots \nrightarrow b > a > \ldots : \)

\[
\begin{align*}
\delta_1 &= \alpha \geq 0, \tag{132} \\
\delta_2 &= s\alpha - \beta = 0 \geq 0, \tag{133} \\
\delta_3 &= (1 - \alpha) - (1 - \beta) = \beta - \alpha \geq 0. \tag{134}
\end{align*}
\]
For $\delta_3$, we assumed that the third choice was $d$, otherwise there is nothing to show.

- $b > a > \ldots \implies a > b > \ldots$:

$$
\begin{align*}
\delta_1 &= \beta \geq 0, \\
\delta_2 &= s\beta - \alpha = \alpha(s^2 - 1) \geq 0, \\
\delta_3 &= s^2\beta - s\alpha + \alpha - \beta = \alpha(s^3 - 2s + 1) \geq 0.
\end{align*}
$$

For $\delta_3$, we assumed that the third choice was $d$, otherwise there is nothing to show.

II. $a > \ldots \iff d > \ldots$: $\varphi(a > \ldots)$ first-order stochastically dominates $\varphi(d > \ldots)$ for all preference orders where $a$ is preferred to $d$, and vice versa.

III. $a > c > d > b \iff c > a > d > b$:

$$
\begin{align*}
\delta_1 &= \alpha - (1 - \gamma_c - \gamma_d) \\
&= \alpha - 1 + (1 - \alpha) \left(\frac{(s - 1)^{-1} + s(s + 1)}{s(s + 1) - 1}\right) \geq 0,
\end{align*}
$$

since

$$
(s - 1)^{-1} + s(s + 1) \geq s(s + 1) - 1 \iff (s - 1)^{-1} \geq -1.
$$

$$
\begin{align*}
\delta_2 &= s(\alpha - 1 + \gamma_c + \gamma_d) - \gamma_c \\
&= (1 - \alpha)s \left(\frac{(s - 1)^{-1} + s(s + 1) - (s - 1)^{-1}s^{-1}}{s(s + 1) - 1}\right) \\
&= (1 - \alpha)s \left(\frac{s(s + 1) + s^{-1}}{s(s + 1) - 1} - 1\right) \geq 0,
\end{align*}
$$

$$
\delta_3 = s\delta_2 \geq 0.
$$
• $c > a > d > b \iff a > c > d > b$:

\[
\begin{align*}
\delta_1 &= \gamma_c \geq 0, \\
\delta_2 &= s\gamma_c + 1 - \gamma_c - \gamma_d - \alpha \\
&= (a - \alpha) \left( 1 + \frac{1 - s(s + 1)}{s(s + 1) - 1} \right) = 0 \\
\delta_3 &= \gamma_d - 1 + \alpha, \\
&= (1 - \alpha) \left( \frac{s(s + 1)}{s(s + 1) - 1} \right) \geq 0.
\end{align*}
\]

IV. $a > c > b > d \iff c > a > b > d$:

\[
\begin{align*}
\delta_1 &= \alpha - (1 - \gamma_c - \gamma_d) \geq 0, \\
\delta_2 &= s(\alpha - 1 + \gamma_c + \gamma_d) - \gamma_c \geq 0,
\end{align*}
\]

as in case III, and

\[
\begin{align*}
\delta_3 &= s^2(\alpha - 1 + \gamma_c + \gamma_d) - s\gamma_c + (1 - \alpha) - \gamma_d \\
&= (1 - \alpha) \left( 1 - s^2 + \frac{s + (s^2 - 1)s(s + 1)}{s(s + 1) - 1} \right) \\
&= (1 - \alpha) \left( \frac{s^2 + s - 1}{s(s + 1) - 1} \right) = 1 - \alpha \geq 0.
\end{align*}
\]

• $c > a > b > d \iff a > c > b > d$:

\[
\begin{align*}
\delta_1 &= \gamma_c \geq 0, \\
\delta_2 &= s\gamma_c + (1 - \gamma_c - \gamma_d) - \alpha \\
&= (1 - \alpha) \left( 1 + \frac{s - 1/2}{s(s + 1) - 1} \right) \\
&= (1 - \alpha) \left( 1 + \frac{1 - s(s + 1)}{s(s + 1) - 1} \right) = (1 - \alpha)(1 - 1) = 0 \geq 0, \\
\delta_3 &= 0 + \gamma_d \geq 0.
\end{align*}
\]

V. $b > \ldots \iff d > \ldots$ : $\varphi(b > \ldots)$ first-order stochastically dominates $\varphi(d > \ldots)$ for
all preference orders where \( b \) is preferred to \( d \), and vice versa.

VI.  

\( \bullet \ b > c \ldots \iff c > b \ldots \): We begin with \( \delta_3 \) as its positivity will also imply positivity of \( \delta_1 \) and \( \delta_2 \). Furthermore, the strictest condition arises from the preference order \( b > c > a > d \).

\[
\delta_3 = s^2(\beta - \gamma_d) + s(-\gamma_c) + (-1 + \gamma_c + \gamma_d) \\
= \alpha \left( \frac{s^5 + 2s^4 - s^2 - s - 1}{s(s + 1) - 1} \right) - \frac{s^4 - s^3}{s(s + 1) - 1} \geq 0
\]  

holds if and only if

\[
\alpha \geq \frac{s^4 - s^3}{s^5 + 2s^4 - s^2 - s - 1}.  \tag{162}
\]

\( \bullet \ c > b \ldots \iff b > c \ldots \):

\[
\delta_1 = \gamma_c \geq 0.  \tag{163}
\]

We can consider the case where \( d \) is the third choice as this condition is strictly stronger than if \( a \) is the third choice. It suffices to consider \( \delta_3 \) as its positivity implies positivity of \( \delta_2 \).

\[
\delta_3 = s^2 \gamma_c + s \gamma_d - s \beta - (1 - \beta) \\
= \alpha \left[ \frac{-s^4 + s^3 + s^2 - s - s^2}{s(s + 1) - 1} \right] + \left[ \frac{s^3 - s + \frac{s^2}{s-1} + 1}{s(s + 1) - 1} \right] \geq 0
\]  

holds if and only if

\[
\alpha \leq \frac{s^3 - s + \frac{s^2}{s-1} + 1}{s^4 + s^3 - s^2 + s + \frac{s^2}{s-1}}.  \tag{166}
\]

VII. \( d > c \ldots \iff c > d \ldots \): \( \varphi(d > \ldots) \) first-order stochastically dominates \( \varphi(c > d \ldots) \) for all preference orders where \( d \) is preferred to \( c \), and vice versa.

VIII.  

\( \bullet \ c > d > b > a \iff c > b > d > a \):

\[
\delta_1 = \gamma_c - \gamma_c \geq 0,  \tag{167}
\]
\[
\delta_2 = 1 - \gamma_c - 0 \geq 0,  \tag{168}
\]
\[
\delta_3 = s(1 - \gamma_c) + \gamma_d \geq 0.  \tag{169}
\]
\( c > b > d > a \iff c > d > b > a : \)

\[
\begin{align*}
\delta_1 &= \gamma_c - \gamma_c \geq 0, \\
\delta_2 &= \gamma_d - 0 \geq 0, \\
\delta_3 &= s \gamma_d - (1 - \gamma_c) \\
&= \alpha \left( \frac{-s^2(s + 1) - \frac{1}{s-1}}{s(s + 1) - 1} \right) + \left( \frac{s^2(s + 1) - \frac{1}{s-1} - s(s + 1) + 1}{s(s + 1) - 1} \right)
\end{align*}
\] (170)

which is positive if and only if

\[
\alpha \leq \frac{s^3 - s + 1 - \frac{1}{s-1}}{s^3 + s^2 - \frac{1}{s-1}}.
\] (174)

IX. \( c > a > d > b \iff c > a > b > d : \)

\[
\begin{align*}
\delta_1 &= \gamma_c - \gamma_c \geq 0, \\
\delta_2 &= 1 - \gamma_c - \gamma_d - 1 + \gamma_c + \gamma_d \geq 0, \\
\delta_3 &= \gamma_d \geq 0.
\end{align*}
\] (175)

\[
\begin{align*}
\delta_1 &= \gamma_c - \gamma_c \geq 0, \\
\delta_2 &= 1 - \gamma_c - \gamma_d - 1 + \gamma_c + \gamma_d \geq 0, \\
\delta_3 &= \gamma_d \geq 0.
\end{align*}
\] (177)

In summary, all local incentive constraints are satisfied if and only if

\[
\frac{s^4 - s^3}{s^5 + 2s^4 - s^2 - s - 1} \leq \alpha \leq \min \left\{ \frac{s^3 - s + 1 - \frac{1}{s-1}}{s^3 + s^2 - \frac{1}{s-1}}, \frac{s^3 - s + \frac{s^2}{s-1} + 1}{s^4 + s^3 - s^2 + s + \frac{s^2}{s-1}} \right\}.
\] (181)

The stronger upper bound is the second: asymptotically, as \( s \) grows, it behaves like \( \frac{1}{s+1} \), which converges to 0, while the first bound converges to 1. The stronger upper bound is also stronger than the upper bound for feasibility, since \( \frac{1}{s+1} \) is smaller than \( \frac{1}{s} \). The lower bound behaves like \( \frac{1}{s+2} \), which is greater than \( \frac{1}{s-1} \), the asymptotic of the lower
bound for feasibility. Finally, observe that the lower bound behaves like \( \frac{1}{s+2} \), which is strictly less than the asymptotic of the upper bound \( \frac{1}{s+1} \). Thus, for sufficiently large \( s \), \( \alpha \) can be chosen such that \( \varphi \) is \( r \)-locally partially strategyproof, which in turn implies feasibility.

It remains to show that for given \( \varepsilon > 0 \), there exist \( r \) and \( \alpha \) such that \( \varphi \) is \( r \)-locally partially strategyproof (and therefore feasible), but not \( r^{2-\varepsilon} \)-partially strategyproof, i.e., 3. To see this, we let \( \bar{s} = s^{2-\varepsilon} \) and consider the preference order \( a > b > c > d \) and the non-local misreport \( c > a > b > d \). If \( \varphi \) is \( \bar{r} \)-partially strategyproof, then in particular we must have \( \delta_3 \geq 0 \) for this manipulation. However, extensive algebraic transformations yield

\[
\delta_3 = \bar{s}^2 (\alpha - 1 + \gamma_c + \gamma_d) + \bar{s} (-\gamma_d) + (\gamma_c) \quad (182)
\]

\[
= (1 - \alpha) \left( -s^{5-\varepsilon} + s^{5-2\varepsilon} + s^{3-\varepsilon} - 1 \right) \quad (183)
\]

Since the leading term with exponent \( 5 - \varepsilon \) has negative sign, this value is negative for sufficiently large \( s \), and this negativity of \( \delta_3 \) is independent of \( \alpha \).

In conclusion, given a value of \( \varepsilon > 0 \), we can find \( r > 0 \) and \( \alpha \in (0, 1) \) such that the resulting mechanism \( \varphi \) is feasible and \( r \)-locally partially strategyproof, but it is not \( r^{2-\varepsilon} \)-partially strategyproof.

### K. Proof of Theorem 8

**Proof of Theorem 8.** A mechanism \( \varphi \) is upper invariant, monotonic, and plain if and only if \( \varphi^p \) is upper invariant and swap monotonic for all distributions \( \mathbb{P} \) with full support.

First, we show sufficiency (“\( \Rightarrow \)”: for some fixed preference reports \( P_i \), the mechanism \( f_i(\cdot, P_{-i}) \) is upper invariant. Thus, \( f^p \) is the convex combination of a finite number of upper invariant mechanisms.

Consider preference orders \( P_i \in \mathcal{P} \) and \( P'_i \in N_{P_i} \) with \( P_i : a_k > a_{k+1} \) and \( P'_i : a_{k+1} > a_k \). If \( f^p(P_i) = f^p(P'_i) \), then there is nothing to show for swap monotonicity. Else, if \( f^p(P_i) \neq f^p(P'_i) \), then there must exist some \( P_{-i} \in \mathcal{P}^{N\setminus\{i\}} \) with \( f_i(P_i, P_{-i}) \neq f_i(P'_i, P_{-i}) \). We get from monotonicity of \( f \) that \( f_{i,a_k}(P_i, P_{-i}) \geq f_{i,a_k}(P'_i, P_{-i}) \) and \( f_{i,a_{k+1}}(P_i, P_{-i}) \leq f_{i,a_{k+1}}(P'_i, P_{-i}) \) for all \( P_{-i} \in \mathcal{P}^{N\setminus\{i\}} \). Moreover, by plainness, there exist
The Probabilistic Serial mechanism is implemented via the Simultaneous Eating algorithm; with $P$ is exhausted under report $\tau$, objects are continuously consumed as time progresses. Let $\tau$ be the time when object $j$ is exhausted under report $P_i$, and $\tau_j$ the time when $j$ is exhausted under report $P_i'$.

Proof of Proposition 6. PS is swap monotonic.

Omitted Proofs and Examples from Section 12

L.1. Proof of Proposition 6

Proof of Proposition 6. PS is swap monotonic.

Suppose agent $i$ is considering the following two reports that only differ by the ordering of two objects $x$ and $y$:

- $P_i : a_1 > \ldots > a_K > x > y > b_1 > \ldots > b_L,$
- $P_i' : a_1 > \ldots > a_K > y > x > b_1 > \ldots > b_L.$

The Probabilistic Serial mechanism is implemented via the Simultaneous Eating algorithm; objects are continuously consumed as time progresses. Let $\tau_j$ be the time when object $j$ is exhausted under report $P_i$, and $\tau_j'$ the time when $j$ is exhausted under report $P_i'$.

89
If $\tau_A = \max(\tau_{a_k}, k \leq K) \geq \min(\tau_x, \tau_y)$, the last of the objects $a_k$ is exhausted only after the first of $x$ and $y$ is exhausted. By upper invariance, $\tau_A = \tau'_A$. This means that by the time $i$ arrives at $x$ (under report $P_i$) or at $y$ under report $P'_i$, one of them is already exhausted. Thus, $i$ will proceed directly to the respective other object. The consumption pattern does not differ between the two reports, i.e., the assignment does not change.

Now suppose that $\tau_A < \tau_y \leq \tau_x$. Then $i$ received no shares of $y$ under $P_i$. But under $P'_i$, it consumes shares of $y$ from $\tau_A$ until $\tau'_y > \tau_A$. Thus, $i$’s share in $y$ strictly increases. Furthermore, $i$ consumed shares of $x$ from $\tau_A$ until $\tau_x$ under report $P_i$. Under report $P'_i$, $i$ arrives at $x$ only later at $\tau'_y > \tau_A$. The same agents that consumed $x$ under report $P_i$ will also be consuming $x$ under report $P'_i$ and at the same times. In addition, there may be some agents who arrive together with $i$ from $y$. Thus, under report $P'_i$ agent $i$ faces strictly more competition for weakly less capacity of $x$, implying that its share of $x$ will strictly decrease. Note that if $i$ faced no competition at $y$, it was the only agent at $y$, and thus consumes it until time 1. In this case the assignment will also decrease, because $i$ arrived later under report $P'_i$.

Finally, suppose that $\tau_A < \tau_x < \tau_y$. Under report $P'_i$, agent $i$ will arrive strictly earlier at $y$, i.e., the competing agents will be the same and arrive at the same times or later (if they arrived from $x$). Thus, the assignment for $y$ will strictly increase under report $P'_i$. Furthermore, $i$ might not receive any shares of $x$ under report $P'_i$, a strict decrease. Otherwise, the argument why $i$ receives strictly less shares of $x$ under $P'_i$ is the same as for the case “$\tau_A < \tau_y \leq \tau_x$”.

L.2. Examples for Rank Efficient Mechanism

Example 4. Consider the setting $N = \{1, 2, 3, 4\}$, $M = \{a, b, c, d\}$, $q_j = 1$, and the preference profile

\[
\begin{align*}
P_1 & : \quad a > d > c > b, \\
P_2 & : \quad a > b > d > c, \\
P_3 & : \quad b > c > d > a, \\
P_4 & : \quad c > a > b > b.
\end{align*}
\]
The unique rank efficient assignment is \( d \to 1, a \to 2, b \to 3, c \to 4 \). Suppose agent 1 changes its report to

\[
P_{1}^{''} : a > c > b > d.
\]

Now the only rank efficient assignment is \( a \to 1, d \to 2, b \to 3, c \to 4 \). The reports \( P_1 \) and \( P_{1}^{''} \) differ by two swaps: \( d \leftrightarrow c \) and \( d \leftrightarrow b \). Thus, at least one of these swaps must have increased the likelihood of getting object \( a \) for agent 1. This contradicts upper invariance. Also, under no report out of \( P_1, P_{1}^{'} : a > c > d > b \), \( P_{1}^{''} \) did agent 1 have any probability of getting objects \( b \) or \( c \). Hence, the swap that changes the assignment involved a change of position of either object \( b \) or \( c \), but the probability for each remained zero, a contradiction to swap monotonicity.

**Example 5.** Consider the setting \( N = \{1, 2, 3, 4, 5\} \), \( M = \{a, b, c, d, e\} \), \( q_j = 1 \), and the preference profile

\[
\begin{align*}
P_1 & : a > c > b > d > e, \\
P_2 & : c > b > a > d > e, \\
P_3 & : c > a > b > e > d, \\
P_4 & : a > c > b > e > d, \\
P_5 & : e > a > b > c > d.
\end{align*}
\]

The unique rank efficient assignment is \( d \to 1, b \to 2, c \to 3, a \to 4, e \to 5 \).

Agent 1 could change its report to

\[
P_{1}^{''} : b > a > c > d > e,
\]

in which case \( b \to 1, d \to 2, c \to 3, a \to 4, e \to 5 \) is the unique rank efficient assignment. Hence, either the swap \( c \leftrightarrow b \) or the swap \( a \leftrightarrow b \) changed the assignment for \( d \), a contradiction to lower invariance.
L.3. Multi-unit Assignment Mechanisms

To introduce $K$-unit assignment mechanisms formally, we must extend our model slightly. Instead of one object, each agent should receive a bundle of $K$ objects. We assume that the agents have additive valuations over bundles. Then it is meaningful to consider ordinal mechanism, where each agent submits a preference order over objects. An assignment is represented by an $n \times m$-matrix $x$, where $\sum_{i \in N} x_{i,j} = q_j$ for all $j \in M$, $\sum_{j \in M} x_{i,j} = K$ for all $i \in N$, and with the additional constraint that each agent should receive at most one copy of each object (i.e., $x_{i,j} \in [0,1]$). By virtue of the Birkhoff-von Neumann Theorem and its extensions (Budish et al., 2013), these assignments are implementable via lotteries. We restrict attention to the assignment of “scarce” objects, i.e., objects for which the capacity $q_j$ is lower than the total number of agents $n$. This is justified, because for any objects with $q_j \geq n$ we can simply distribute one copy to each agent, independent of their preferences.

L.3.1. Probabilistic Serial Mechanism for $K$-unit Assignment

For the $K$-unit assignment problem, the Probabilistic Serial mechanism takes as input a preference profile $P$ and determines an assignment as follows:

- The objects are treated as if they were divisible. All agent begin consuming probability shares of their respective most preferred object at equal speeds.
- When an agent has consumed a total of 1 in probability shares from an object, this agent leaves the object and continues consuming probability shares of its next most preferred object that still has remaining capacity.
- When an object is completely consumed, the agents from this object move on to their respective next most preferred objects that still have remaining capacity.
- This continues until all agents have collected probability shares that sum to $K$.
- The entry $x_{i,j}$ of the resulting assignment is given by the amount of shares of $j$ that $i$ managed to consume in this process.

L.3.2. HBS Draft Mechanism for $K$-unit Assignment

For the $K$-unit assignment problem, the HBS Draft mechanism is defined with respect to fixed priority order $\pi$ over agents. Given $\pi$, the mechanism takes a preference profile
\(P\) as input and determines an assignment as follows:

- At the beginning of the first pass, the agent with the highest priority (according to \(\pi\)) draws one copy of the object that it prefers most.
- The agent with the second highest priority draws one copy of the object that it prefers most out of all available objects.
- In the order given by \(\pi\), all agents draw copies of their respective most preferred available objects.
- The first pass ends, when the last agent has drawn its first object.
- The second pass works like the first pass, but the order of \(\pi\) is reversed: the agent with the lowest priority draws a second object (but no second copy of object that it has drawn in the first pass).
- Then the agent with the second-lowest priority draws a second object, and so on.
- The third, fourth, etc. pass are analogous. In odd passes, the drawing order is \(\pi\), while in even passes, the drawing order is the reverse of \(\pi\).
- The process continues until all agents have received a total of \(K\) objects, which is the case after exactly \(K\) passes.

The HBS Draft mechanism can also be understood as a random mechanism if the priority order \(\pi\) is determined randomly.

Proof of Proposition 7. Given a setting \((N, M, q)\) with \(\sum_{j \in M} q_j = n \cdot K\), the Probabilistic Serial mechanism and the HBS Draft mechanism for the \(K\)-unit assignment problem are upper invariant, monotonic, and plain. However, neither of them is swap monotonic.

Upper invariance and monotonicity of these mechanisms are straightforward. The proofs of plainness are more challenging. We first show plainness of Probabilistic Serial for multi-unit assignment.

Consider any preference orders \(P_i \in P, P'_i \in N_{P_i}\) with \(P_i : a_k > a_{k+1}\) and \(P'_i : a_{k+1} > a_k\). Construct the preference profile \((P_i, P_{-i})\) by setting \(P'_{i'} = P_i\) for all \(i' \neq i\). The assignment \(PS(P_i, P_{-i})\) simply gives all agents the same assignment vector. Precisely, \(PS_{i,j}(P_i, P_{-i}) = q_j/n\). Suppose that all agents start consuming shares of \(a_k\) at \(\tau_{k-1}\) and start consuming shares of \(a_{k+1}\) at \(\tau_{k+1}\). If \(i\) reports \(P'_i\) instead, it will start consuming
shares of $a_{k+1}$ already at $\tau_{k-1}$, but all other agents will arrive strictly later. Thus, $i$’s assignment of $a_{k+1}$ increases strictly. Furthermore, $a_k$ will be exhausted by the other agents. Therefore, by the time $i$ finishes consuming $a_{k+1}$, it will not receive any more shares of $a_k$. Thus, $i$’s assignment of $a_k$ decreases strictly. Since this construction holds for any preference order $P_i$, plainness of PS follows.

Next, we show plainness of the HBS Draft mechanism for multi-unit assignment. Consider any preference orders $P_i \in \mathcal{P}, P'_i \in N_{P_i}$ with $P_i : a_k > a_{k+1}$ and $P'_i : a_{k+1} > a_k$. Suppose that some preference reports $P_{-i} \in \mathcal{P}^N \setminus \{i\}$, the assignment of $i$ changes between $P_i$ and $P'_i$.

First, observe that this change must involve either $a_k$ or $a_{k+1}$: towards contradiction, assume that $\text{HBSD}_i(P_{i}, P_{-i}) \neq \text{HBSD}_i(P'_{i}, P'_{-i})$, but $\text{HBSD}_{i,a_k}(P_i, P_{-i}) = \text{HBSD}_{i,a_k}(P'_{i}, P'_{-i})$ and $\text{HBSD}_{i,a_{k+1}}(P_i, P_{-i}) = \text{HBSD}_{i,a_{k+1}}(P'_{i}, P'_{-i})$. Since all agents get exactly one object in each pass, the passes in which $i$ received the different objects, including $a_k$ and $a_{k+1}$, must be the same under $P_i$ and $P'_i$. But then the order in which all agents received their objects must also remain the same, which implies that the assignment of $i$ has not changes; a contradiction.

There are different ways in which the change can involve $a_k$ and $a_{k+1}$. The following table gives an overview:

<table>
<thead>
<tr>
<th>$P_i$</th>
<th>$P'_{i}$</th>
<th>$\text{HBSD}<em>i(P</em>{i}, P_{-i})$</th>
<th>$\text{HBSD}<em>i(P'</em>{i}, P'_{-i})$</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_k$</td>
<td>$a_{k+1}$</td>
<td>$a_k$</td>
<td>$a_{k+1}$</td>
<td>Case 1 (impossible)</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>Case 1 (impossible)</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Case 1 (impossible)</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>Case 2</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>Case 3</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>Symmetric to Case 3</td>
</tr>
</tbody>
</table>

The omitted cases can be ruled out, because they violate monotonicity or because there is no change. The remaining cases are discussed next.

**Case 1:** When $i$ gets neither $a_k$ nor $a_{k+1}$, $i$ tried to draw both objects in the same pass. Since $i$ did not get $a_k$, it immediately attempted to draw $a_{k+1}$, which it did not get either. Thus, if $i$ tries to draw $a_{k+1}$ first (in the same pass), it will still not receive it there, move on to $a_k$ immediately, and also not get it. This shows that Case 1 is impossible.
Case 2: If $\text{HBSD}_{i,a_k}(P_i, P_{-i}) = \text{HBSD}_{i,a_{k+1}}(P_i', P_{-i}) = 1$ and $\text{HBSD}_{i,a_{k+1}}(P_i, P_{-i}) = \text{HBSD}_{i,a_k}(P_i', P_{-i}) = 0$, then we have already found preference reports $P_{-i}$ where the change in the assignment of $a_k$ and $a_{k+1}$ is strict for both. This is what we need to show for plainness.

Case 3: This is the most complex case. By reporting $P_i$, $i$ gets $a_k$, but by reporting $P_i'$, it gets both $a_k$ and $a_{k+1}$. First, we “reduce” the preference reports $P_{-i}$ to $P'_{-i}$ in such a way that agents first rank all the objects that they actually received. All other objects, which they do not receive are ranked below. The application process and resulting assignment of $\text{HBSD}(P_i, P_{-i})$ and $\text{HBSD}(P_i', P'_{-i})$ are exactly the same, except that no agent (except $i$) tries to draw an object that is already exhausted.

Now, consider the pass in which $i$ draws $a_k$ (when reporting $P_i$). Note that $i$ is not the last agent to draw an object in this pass (otherwise, $i$ would draw $a_k$ and then immediately draw again; thus, $i$ would receive $a_k$ and $a_{k+1}$, independent of the order in which it ranks them). At the beginning of the pass, $a_k$ is either the object that $i$ prefers most of all the objects with remaining capacity, or there are other objects with remaining capacity that $i$ would prefer to $a_k$.

Case (I): Suppose that $a_k$ is the object that $i$ prefers most of all objects with remaining capacity at the beginning of the pass. Let $q_{a_k}'$ be the remaining capacity of $a_k$ at the beginning of the pass. There are exactly $q_{a_k} - q_{a_k}'$ agents who have already received $a_k$ in prior rounds. All other $(n-1) - (q_{a_k} - q_{a_k}')$ agents (except $i$) would attempt draw $a_k$ if they ranked it in this round; and $(n-1) - (q_{a_k} - q_{a_k}') \geq 1$ since the initial capacity $q_{a_k}$ of $a_k$ was at most $n-1$. We obtain the preference reports $P'_{-i}$ by changing the draw of $q_{a_k}'$ agents in this round to $a_k$ (if they do not already draw $a_k$). Finding such agents is possible, because $(n-1) - (q_{a_k} - q_{a_k}') \geq q_{a_k}'$. In particular, we ensure that the last agent to draw in this round (as argued above, this agent cannot be $i$) draws the last copy of $a_k$.

With $P'_{-i}$ constructed in this way, $i$ can draw either $a_k$ or $a_{k+1}$ in this round. But in the subsequent round, $a_k$ will be exhausted. Thus, $\text{HBSD}_{i,a_k}(P_i, P'_{-i}) = 1$, but $\text{HBSD}_{i,a_k}(P_i, P^1_{-i}) = 0$. This shows that the assignment of $a_k$ may be affected, which is what we needed for directness.
Case (II): Now suppose that at the beginning of the pass where \( i \) draws \( a_k \) there are other objects with remaining capacity that \( i \) would prefer to \( a_k \). The reason for \( i \) to draw \( a_k \) instead must be that some other agents draw these objects between the beginning of the pass and \( i \)'s draw. We can select an agent who draws one of the objects that \( i \) prefers to \( a_k \) (\( x \), say) and change this agent’s draw: denote by \( c \) the object that \( i \) draws after drawing \( a_k \) under HBSD(\( P_i, P_{-i} \)). Instead of \( x \), the agent draws the copy of \( c \) that must still be available. Furthermore, there exists an agent who draws the last copy of \( a_k \) between the draw of \( i \) in this and the next pass. We change this agent’s draw to \( a_k \). This yields the preference reports \( P^3_{-i} \).

Observe that under the preference profile \( (P_i, P^3_{-i}) \), \( i \) will draw \( x \) instead of \( a_k \), one agent will draw \( c \) instead of \( x \), and one agent will draw \( a_k \) instead of \( a_{k+1} \). Thus, at the beginning of the next pass, there are no copies of objects that \( i \) prefers to \( a_k \). Moreover, at the time when \( i \) draws next, there is one copy of \( a_{k+1} \) and one copy of \( a_k \). This means that the next pass is a Case (I) pass, so that we can complete the proof by the construction as in Case (I).

This concludes the proofs of plainness for both mechanisms.

Finally, we show that neither of the mechanisms is swap monotonic. For strict, fixed priorities, NBM and ABM are deterministic mechanisms. On these mechanisms, swap monotonicity is equivalent to strategyproofness. Conversely, since the mechanisms are not strategyproof, they cannot be swap monotonic. To see that Probabilistic Serial and the HBS Draft mechanism are not swap monotonic, consider the following: let there be four objects in unit capacity, \( a, b, c, d \), two agents with preferences

\[
P_1 : a > b > c > d,
\]

\[
P_2 : c > b > d > a,
\]

and each agent should get two objects. Under both mechanisms, the assignments are

\[
PS(P_1, P_2) = HBSD(P_1, P_2) = \begin{pmatrix}
1 & 1/2 & 0 & 1/2 \\
0 & 1/2 & 1 & 1/2
\end{pmatrix}.
\]  

(184)
If agent 1 reports $P'_1 : b > a > c > d$ instead, the assignments change to

$$\text{PS}(P'_1, P_2) = \text{HBSD}(P'_1, P_2) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}. \quad (185)$$

Thus, agent 1’s assignments of $b$ increase strictly under the swap, but its assignments for $a$ do not decrease strictly.