CAES Working Paper Series

Adjacent manipulation by weak orders

Shin Sato

Faculty of Economics

Fukuoka University, Japan

WP-2010-003

Center for Advanced Economic Study

Fukuoka University

(CAES)

8-19-1 Nanakuma, Jonan-ku, Fukuoka,
JAPAN 814-0180
+81-92-871-6631

Adjacent manipulation by weak orders

Shin Sato

Faculty of Economics, Fukuoka University

8-19-1 Nanakuma, Jonan-ku, Fukuoka,

JAPAN 814-0180

E-mail: `shinsato@adm.fukuoka-u.ac.jp`

February 26, 2010

Abstract

This paper considers adjacent manipulation by weak orders. The concept of adjacent manipulability of social choice functions is introduced by Sato (2010), but his analysis is limited to the cases where agents' preferences are formulated as linear orders. I show that the main result by Sato (2010) cannot be carried over to the circumstances where ties among alternatives are allowed in agents' preferences. However, it is also shown that when agents' preferences are linear or "almost linear", then the result by Sato (2010) revives.

Keywords: AM-proofness, social choice function, strategy-proofness, weak order.

JEL classification: D71.

1 Introduction

The robustness to strategic manipulation is one of the most important topics in the theory of social choice.[1] In this area of research, Sato (2010) incorporates agents' reluctance to make a "big lie" into the analysis. However, his analysis is limited to the cases where agents' preferences are formulated as linear orders, i.e., ties among alternatives are excluded from consideration. In this paper, I consider weak orders in which ties are allowed as agents' preferences and examine the same question as that of Sato (2010);

> Consider the two societies N and N' such that
>
> - Any agent in N considers only preferences "around" the true one as the candidates for misrepresentation. In other words, "big lies" are excluded from the options for misrepresentation.
> - Any agent in N' feels free to misreport any preferences, as assumed under strategy-proofness.
>
> Then, is there any difference between N and N' in terms of difficulty of constructing an SCF which is robust to strategic manipulation?

Sato (2010) argues that in many cases, there is a good reason to believe that we are often confronted with the societies like N in which the agents are reluctant to make a "big lie". As his main result, Sato (2010) finds a condition under which the reluctance to make a "big lie" is not helpful for the social choice rule designer at all to construct a social choice function (SCF) immune from strategic manipulation.

What I would like to show in this paper is twofold. First, the result by Sato (2010) collapses with weak orders. Second, when agents' preferences are linear or "almost linear", then the result by Sato (2010) revives.

[1] The seminal result is the Gibbard-Satterthwaite theorem (Gibbard, 1973; Satterthwaite, 1975).

In the theory of social choice, linear orders and weak orders are two major formulations of agents' preferences. Generally speaking, linear orders are easier to deal with because ties are excluded. Some results are stated with linear orders just for simplicity.[2] However, in some other cases, results critically rely on the assumption that ties are excluded. For example, on the set of linear orders, dictatorship implies strategy-proofness, but on the set of weak orders, this implication does not hold. My analysis will show that the main result by Sato (2010) is "near" the latter type.

This paper is organized as follows. Section 2 gives basic notation and definitions, and the core concept of this paper, AM-proofness, is introduced. Section 3 gives several examples of the nonequivalence between strategy-proofness and AM-proofness, and find causes of the nonequivalence. Section 4 contains the equivalence theorem, and Section 5 concludes.

2 Basic notation and definitions

Let $N = \{1, \dots, n\}$ be a finite set of agents, let X be a finite set of alternatives with $|X| = m$.[3] Let $\mathcal{W}$ be the set of all weak orders on X, and let $\mathcal{L}$ denote the set of all linear orders on X.[4] Elements of $\mathcal{W}$ are called *preference relations*. Typical notation for an element of $\mathcal{W}$ is R, and subscripts represent indices of agents; R_i is a preference relation of agent i. For each $R \in \mathcal{W}$ and for each $A \subset X$, let $R|A$ denote the restriction of R to A. The strict parts of R and R_i are represented as P

[2]A representative example is the Gibbard-Satterthwaite theorem stating that strategy-proofness implies dictatorship. The theorem is often stated over linear orders, but it holds over weak orders.

[3]For any set A, $|A|$ denotes its cardinality.

[4]Any subset of $X \times X$ is called a *binary relation* on X. A binary relation R on X is said to be *complete* if for any $x, y \in X$, either xRy or yRx holds, *transitive* if for any $x, y, z \in X$, $[xRy \; \& \; yRz]$ implies xRz, *antisymmetric* if for any $x, y \in X$, $[xRy \; \& \; yRx]$ implies $x = y$. A binary relation is called a *weak order* if it is complete and transitive, a *linear order* if it is complete, transitive, and antisymmetric.

and P_i, and the indifference parts are represented as I and I_i, respectively.

The indifference part I of $R \in \mathcal{W}$ is an equivalence relation,[5] and it induces a partition of X. Let $r_k(R)$ denote the kth indifference class. For example, $r_1(R)$ denotes the set $\{x \in X \mid xRy \; \forall y \in X\}$: the set of most preferred alternatives with respect to R. The set $r_2(R)$ denotes the set $\{x \in X \setminus r_1(R) \mid xRy \; \forall y \in X \setminus r_1(R)\}$, and so on. For notational simplicity, when $r_k(R)$ consists of one alternative $x \in X$, then I omit the braces and write $r_k(R) = x$ whenever the omission does not cause confusion.

An n-tuple $R_N = (R_1, \ldots, R_n) \in \mathcal{W}^N$ is called a *preference profile*. Given a preference profile R_N, (R'_i, R_{-i}) denotes the preference profile such that R_i is replaced by $R'_i \in \mathcal{W}$ in R_N.

Given a set $\mathcal{D} \subset \mathcal{W}$, a function f of $\mathcal{D}^N$ into X is called a *social choice function (SCF) on* $\mathcal{D}$, and $\mathcal{D}$ is called the domain of f.

In this paper, I measure the difference between two preference relations by the *Kemeny distance* (Kemeny, 1959; Kemeny and Snell, 1962). The Kemeny distance d_K is defined by for any $R, R' \in \mathcal{W}$,

$$d_K(R, R') = |R \triangle R'|,$$

where $R \triangle R'$ is the symmetric difference between R and R', defined by $(R \setminus R') \cup (R' \setminus R)$.

Another (more intuitively appealing) definition of the Kemeny distance is the following. Let $\mathcal{X}$ denote the set $\{\{x, y\} \subset X \mid x \neq y\}$. Given $R, R' \in \mathcal{W}$ and $\{x, y\} \in \mathcal{X}$, there are three possible cases.

CASE 1: Alternatives x and y are oppositely ranked in R and R'. In other words, x is preferred to y in either of R and R', and y is preferred to x in the other.

[5] A binary relation R on X is said to be *reflexive* if for any $x \in X$, xRx holds, *symmetric* if for any $x, y \in X$, xRy implies yRx. A binary relation is called an *equivalence relation* if it is complete, symmetric, and transitive.

Table 1: Preference relations R^1, R^2, and R^3

R^1	R^2	R^3
x	xy	y
y	z	x
z		z

CASE 2: Alternatives x and y are strictly ranked in either of R and R' while x and y are indifferent in the other.

CASE 3: Alternatives x and y are ranked in the same way in R and R'.

For each $R, R' \in \mathcal{W}$ and for each $\{x, y\} \in \mathcal{X}$, define

$$s_{R,R'}(\{x,y\}) = \begin{cases} 2, & \text{if CASE 1 applies,} \\ 1, & \text{if CASE 2 applies,} \\ 0, & \text{if CASE 3 applies.} \end{cases}$$

Based on this preparation, define

$$d_K(R, R') = \sum_{\{x,y\} \in \mathcal{X}} s_{R,R'}(\{x,y\}).$$

This can be the definition of the Kemeny distance. Clearly, the minimum positive distance is 1, and the maximum distance is $m(m-1)$.

Example 2.1

As an example, consider R^1, R^2, and R^3 in Table 1. (Superscripts are used to distinguish different preference relations.) Here, the alternatives are x, y, and z. The alternative z is always bottom ranked, but x and y are differently ranked in three preference relations; x is preferred to y in R^1, they are indifferent in R^2, and y is preferred to x in R^3. Thus, $s_{R^1,R^2}(\{x,y\}) = s_{R^2,R^3}(\{x,y\}) = 1$ and $s_{R^1,R^3}(\{x,y\}) = 2$, and we have $d_K(R^1, R^2) = d_K(R^2, R^3) = 1$, and

$d_K(R^1, R^3) = 2$. Remember that 1 is the minimum positive distance in $\mathcal{W}$. Thus, I would like to say that R^1 and R^2 are adjacent while R^1 and R^3 are not in $\mathcal{W}$.

Definition 2.1

Two weak orders R and R' are said to be *adjacent* if $d_K(R, R') = 1$. Let $A(R)$ denote the set of all weak orders adjacent to R in $\mathcal{W}$.

Remark 2.1

Two distinct linear orders are never adjacent. If R and R' are two distinct linear orders, then there exist two alternatives x and y such that xPy and $yP'x$, i.e., $s_{R,R'}(\{x, y\}) = 2$. Thus, $d_K(R, R') \geq 2$, and R and R' are not adjacent.

Definition 2.2

An SCF f on $\mathcal{D}$ is said to be

- *strategy-proof* if for any $R_N \in \mathcal{D}^N$ and for any $i \in N$,

$$f(R_N) R_i f(R'_i, R_{-i}), \quad \forall R'_i \in \mathcal{D}.$$

- *AM-proof* if for any $R_N \in \mathcal{D}^N$ and for any $i \in N$,

$$f(R_N) R_i f(R'_i, R_{-i}), \quad \forall R'_i \in (A(R_i) \cap \mathcal{D}).$$

The only difference between strategy-proofness and AM-proofness is the set of reportable false preference relations. If an SCF f is AM-proof, agents cannot be better off by misreporting a preference relation adjacent to his sincere one. The concept of AM-proofness is an extreme one in the sense that it considers only adjacent preference relations as the candidates for misrepresentation. However, as Corollary 4.1 of Sato (2010) shows, once we find a condition under which strategy-proofness and AM-proofness are equivalent, the condition is also sufficient for the equivalence between strategy-proofness and any intermediate notion of the robustness to strategic manipulation. Thus, considering AM-proofness is effective in finding an answer to the motivating question of this paper in Introduction.

Definition 2.3

For any two preference relations R and R' in a domain $\mathcal{D} \subset \mathcal{W}$, a sequence $(R^1 \ldots R^\ell)$ is called a *path* from R to R' in $\mathcal{D}$ if

(i) $R^1 = R$ and $R^\ell = R'$,

(ii) $R^h \in \mathcal{D}$ for all $h \in \{1, \ldots, \ell\}$,

(iii) any two preference relations in the sequence are distinct from each other, and

(iv) $R^{h+1} \in A(R^h)$ for all $h \in \{1, \ldots, \ell - 1\}$.

A path from R to R' in $\mathcal{D}$ is also a path from R' to R in $\mathcal{D}$. Thus, when the direction is not important, we speak of a path *between* R and R' in $\mathcal{D}$.

Definition 2.4

A sequence $(R^1 \ldots R^\ell)$ of weak orders is said to be *with restoration* if there are distinct alternatives x and y such that

(i) $xP^k y$, $xI^{k'} y$, and $xP^{k''} y$ for some k, k', k'' with $k < k' < k''$, or

(ii) $xP^k y$, $xI^{k'} y$, $yP^{k''} x$, and $xI^{k'''} y$ for some k, k', k'', k''' with $k < k' < k'' < k'''$.

The sequence is said to be *without restoration* if there is no such pair of alternatives.

When a sequence is with restoration, at least one of the two cases in the above definition holds. In both cases, preferences over some pair of alternatives restore through the sequence. In the first case, x is preferred to y, then they become indifferent, and after that, x is preferred to y, again. In the second case, indifference between x and y restores through the sequence. In most cases, we deal with paths, and examples of a path with restoration will be given in the next section.

3 The nonequivalence between strategy-proofness and AM-proofness

In this section, I will give several examples of the nonequivalence between strategy-proofness and AM-proofness.

3.1 The universal domain

First, I show the nonequivalence on the universal domain $\mathcal{W}$. Assume $m \geq 3$, and let x, y, z be three alternatives distinct from each other. Let R^* be the weak order such that $r_1(R^*) = X$. At R^*, any two alternatives are indifferent. Thus, I will call R^* *the total indifference relation*. The important fact is $A(R^*) = \emptyset$ (see Appendix). Given $R_N \in \mathcal{W}^N$ and $a, b \in X$, let $N(R_N, a, b)$ denote the number of the agents who strictly prefer a to b at R_N. Then, for each preference profile R_N, define

$$f(R_N) = \begin{cases} x, & \text{if } R_i = R^* \text{ for all } i \in N, \\ y, & \text{if } N(R_N, y, z) \geq 1, \\ z, & \text{otherwise.} \end{cases}$$

This SCF chooses x iff every agent reports R^*, y iff some individual strictly prefers y to z, and z otherwise. I will show that the SCF f is AM-proof, but violates strategy-proofness. First, to see that f is not strategy-proof, consider the preference profile R_N such that xP_1yP_1z and $R_i = R^*$ for all $i \in N \setminus \{1\}$. By definition, $f(R_N) = y$. Let $R'_1 = R^*$. Then, $f(R'_1, R_{-1}) = x$, and hence $f(R'_1, R_{-1})P_1f(R_N)$. Thus, f is not strategy-proof.

Next, I show that f is AM-proof. Let R_N be any preference profile. I will show that agent 1 cannot be better off by misreporting a preference relation adjacent to R_i.

CASE 1: $R_i = R^*$ for all $i \in N \setminus \{1\}$. In this case, if $R_1 = R^*$, then agent 1 is indifferent among all alternatives, and hence he has no incentive to lie. Assume $R_1 \neq R^*$. Let R'_1 be any element of $A(R_1)$. Because $R'_1 \neq R^*$,[6] $f(R'_1, R_{-1})$ is either y or z. The social outcome under the sincere preference relation R_1 is y if yP_1z, and z otherwise, i.e., zR_1y. In any case, $f(R'_1, R_{-1})$ is not strictly preferred to $f(R_N)$ with respect to R_1.

CASE 2: $R_i \neq R^*$ for some $i \in N \setminus \{1\}$.

SUBCASE 2.1: yP_jz for some $j \in N \setminus \{1\}$. In this case, regardless of what agent 1 reports, the social outcome is y, and hence agent 1 has no incentive to lie.

SUBCASE 2.2: zR_jy for all $j \in N \setminus \{1\}$. If $R_1 = R^*$, then agent 1 is indifferent among all alternatives, and hence he has no incentive to lie. Assume $R_1 \neq R^*$. Let R'_1 be any element of $A(R_1)$. Because $R'_1 \neq R^*$, $f(R'_1, R_{-1})$ is either y or z. The social outcome under the sincere preference relation R_1 is y if yP_1z, and z otherwise, i.e., zR_1y. In any case, $f(R'_1, R_{-1})$ is not strictly preferred to $f(R_N)$ with respect to R_1.

For any agent other than agent 1, it can be established by a similar argument that he cannot be better off by misreporting an adjacent preference relation.

This example is sufficient to show the discrepancy between the two formulations of preferences; weak and linear orders. (Sato (2010) proves the equivalence between strategy-proofness and AM-proofness over the universal domain of linear orders.)

A more interesting implication follows from this example. Even if we allow misreporting preference relations R' within $m-2$ distance from the sincere one R (i.e., $d_K(R, R') \leq m-2$), it can be seen that the above SCF f is robust to strategic manipulation.[7] In this sense, the discrepancy between this result and that of Sato

[6]Because $R'_1 \in A(R_1)$, if $R_1 = R^*$, then $R^* \in A(R_1)$. This implies $R_1 \in A(R^*)$, which is a contradiction to $A(R^*) = \emptyset$.

[7]Let $B(R)$ denote the set $\{R' \in \mathcal{W} \mid d_K(R, R') \leq m-2\}$. Then, it can be seen that $B(R^*) = \emptyset$. (See Appendix.) Note that the nonequivalence on $\mathcal{W}$ is essentially due to the fact $A(R^*) = \emptyset$.

(2010) is *not* due to the difference of the concept of adjacency[8]; if we called two weak orders R and R' with $d_K(R, R') = 2$ adjacent to each other and two distinct linear orders could be adjacent as in Sato (2010), the nonequivalence would occur on $\mathcal{W}$ provided that $m \geq 4$.

The point of the nonequivalence on $\mathcal{W}$ is $A(R^*) = \emptyset$, i.e., the existence of an "isolated" preference relation. In the rest of this section, other causes of the nonequivalence will be discussed.

3.2 Causes of the nonequivalence

I will explore the causes of the nonequivalence between strategy-proofness and AM-proofness through several examples.

Example 3.1

Let $X = \{x, y, z, w\}$ and let $\mathcal{D}$ be the set of the preference relations from R^1 to R^7 in Table 2. It can be checked that the pairs of adjacent preferences relations are $(R^1, R^2), (R^2, R^3), (R^3, R^4), (R^4, R^5), (R^5, R^6), (R^6, R^7)$. Thus, no preference relation is "isolated". Consider an SCF f such that for any $R_{-1} \in \mathcal{D}^{N\setminus\{1\}}$, f chooses the alternative in the square brackets corresponding to the agent 1's preference relation in Table 2. For example, $f(R^1, R_{-1}) = x$, $f(R^2, R_{-1}) = y$, and so on. Then, f is AM-proof, but it is not strategy-proof.

To see that f is not strategy-proof, let R_N be a preference profile such that $R_1 = R^7$. Let $R'_1 = R^1$. Then, $f(R'_1, R_{-1}) P_1 f(R_N)$.

Next, I prove that f is AM-proof. The agents other than agent 1 cannot affect a social outcome, and hence they have no incentive to lie. Thus, it suffices to consider agent 1. Let R_N be any preference profile and let R'_1 be any element of $A(R_1)$. By the definition of f, $f(R'_1, R_{-1}) \neq f(R_N)$ only if $\{R_1, R'_1\} = \{R^1, R^2\}$.

Therefore, the above arguments hold when we replace $A(R)$ by $B(R)$ for any $R \in \mathcal{W}$.

[8] In Sato (2010), two linear orders R and R' are said to be adjacent if $d_K(R, R') = 2$. Note that in $\mathcal{L}$, 2 is the minimum positive distance.

Table 2: Preference relations in Example 3.1

R^1	R^2	R^3	R^4	R^5	R^6	R^7
$[x]$	$x[y]$	$[y]$	$[y]$	$[y]$	$x[y]$	x
y	z	x	x	x	w	$[y]$
z	w	z	zw	w	x	w
w		w		z		x

Table 3: Preference relations in Example 3.2

R^1	R^2	R^3	R^4	R^5	R^6
x	$x[y]$	$[y]$	$[y]$	$[y]$	$[x]y$
$[y]$	z	x	x	x	w
z	w	z	zw	w	z
w		w		z	

Therefore, profitable misrepresentation can occur only if $(R_1, R_1') = (R^1, R^2)$ or $(R_1, R_1') = (R^2, R^1)$. However, in both cases, $f(R_N)R_1f(R_1', R_{-1})$ holds, which shows that f is AM-proof.

The point of this example is that between R^1 and R^7, there is only a path *with* restoration in $\mathcal{D}$. (xP^1y, xI^2y, and xP^7y.) This restoration of preferences over x and y is one of the causes of the nonequivalence between strategy-proofness and AM-proofness.

The following example shows that another kind of restoration can also be a cause of the nonequivalence.

Example 3.2

Let $X = \{x, y, z, w\}$ and let $\mathcal{D}$ be the set of the preference relations from R^1 to R^6 in Table 3. The pairs of adjacent preference relations are (R^1, R^2), (R^2, R^3),

Table 4: Preference relations in Example 3.3

R^1	R^2	R^3
x	$x[y]$	$[x]y$
$[y]$	z	zw
z	w	
w		

(R^3, R^4), (R^4, R^5), and (R^5, R^6). From R^1 to R^6, there is only one path $(R^1 R^2 R^3 R^4 R^5 R^6)$ in $\mathcal{D}$. Note that this path is *with* restoration: xP^1y, xI^2y, yP^3x, and xI^6y. As in Example 3.1, consider an SCF f such that for any $R_{-1} \in \mathcal{D}^{N\setminus\{1\}}$, f chooses the alternative in the square brackets corresponding to the agent 1's preference relation in Table 3. Then, it can be seen that f is AM-proof, but it violates strategy-proofness.

Assume that between any two preference relations, there exists a path without restoration. Then, is this sufficient for the equivalence? The equivalence theorem of Sato (2010) suggests that the answer would be 'yes'. However, the answer is 'no', as the following example shows.

Example 3.3

Let $X = \{x, y, z, w\}$ and let $\mathcal{D}$ be the set of preference relations R^1, R^2, and R^3 in Table 4. Then, in $\mathcal{D}$, the pairs of adjacent preference relations are (R^1, R^2) and (R^2, R^3). Note that every preference relation has an adjacent one and between any preference relations, there exists a path without restoration in $\mathcal{D}$. As in previous examples, consider an SCF f such that for any $R_{-1} \in \mathcal{D}^{N\setminus\{1\}}$, f chooses the alternative in the square brackets corresponding to the agent 1's preference relation in Table 2:

$$f(R_N) = \begin{cases} y, & \text{if } R_1 \in \{R^1, R^2\} \\ x, & \text{if } R_1 = R^3. \end{cases}$$

It is clear that f is AM-proof, but violates strategy-proofness. Therefore, the nonequivalence between strategy-proofness and AM-proofness arises.

This last example suggests that the existence of multiple "thick" indifference classes can be another cause of the nonequivalence between strategy-proofness and AM-proofness.

Based on these examples, I will give a sufficient condition for the equivalence in the next section.

4 The equivalence between strategy-proofness and AM-proofness: almost linear preferences

In this section, I introduce the concept of almost linear preferences and find a sufficient condition for the equivalence between strategy-proofness and AM-proofness.

Definition 4.1

A weak order R is said to be *almost linear* if one indifference class contains two alternatives, and all the other indifference classes consist of one alternative.

For example, R^2 in Table 4 is almost linear while R^1 and R^3 are not. An almost linear order is very close to linear orders in the sense that every almost linear order has an adjacent linear order. Let $\mathcal{AL}$ denote the set of all linear and almost linear orders. Considering domains contained in $\mathcal{AL}$ means that only a slight deviation from linear orders is allowed.

The following theorem is a main result of this paper.

Theorem 4.1

Let $\mathcal{D}$ be any domain contained in $\mathcal{AL}$. Assume that for any two preference relations in $\mathcal{D}$, there exists a path without restoration between them. Then, any SCF on $\mathcal{D}$ satisfying AM-proofness satisfies strategy-proofness.

Because strategy-proofness trivially implies AM-proofness, this theorem finds a sufficient condition for the logical equivalence between strategy-proofness and AM-proofness.

First, I prove the following lemma.

Lemma 4.1

Let f be an SCF on $\mathcal{D} \subset \mathcal{AL}$. For any preference profile $R_N \in \mathcal{D}^N$, for any $i \in N$, and for any $R'_i \in (A(R_i) \cap \mathcal{D})$ such that for any $y \in X$,

- *$f(R_N)R_i y$ implies $f(R_N)R'_i y$, and*
- *$f(R_N)P_i y$ implies $f(R_N)P'_i y$,*

we have $f(R_N) = f(R'_i, R_{-i})$.

Proof of Lemma 4.1. Suppose $f(R_N) \neq f(R'_i, R_{-i})$. Because R_i is a weak order, there are three possible cases.

CASE 1: $f(R_N)I_i f(R'_i, R_{-i})$. This implies that R_i is almost linear. Then, its adjacent preference relation $R'_i \in \mathcal{DAL}$ should be a linear order, and hence it is antisymmetric. Because $f(R_N)$ and $f(R'_i, R_{-i})$ are assumed to be distinct alternatives, either $f(R_N)P'_i f(R'_i, R_{-i})$ or $f(R'_i, R_{-i})P'_i f(R_N)$ holds. The former case is a contradiction to AM-proofness. Consider the latter case. Note that, by assumption, for any $y \in X$, $yP'_i f(R_N)$ implies $yP_i f(R_N)$.[9] Thus, $f(R'_i, R_{-i})P'_i f(R_N)$ implies $f(R'_i, R_{-i})P_i f(R_N)$, which is a contradiction to AM-proofness.

CASE 2: $f(R_N)P_i f(R'_i, R_{-i})$. By assumption, $f(R_N)P'_i f(R'_i, R_{-i})$ holds, which is a contradiction to AM-proofness.

[9] This is the contrapositive of $[f(R_N)R_i y \Rightarrow f(R_N)R'_i y]$.

Table 5: Two cases where a social outcome can change

Case 1		Case 2	
R^h	R^{h+1}	R^h	R^{h+1}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$f(R^h, R_{-i})$	$f(R^h, R_{-i})\ x$	$f(R^h, R_{-i})\ x$	x
x	$\vdots$	$\vdots$	$f(R^h, R_{-i})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

CASE 3: $f(R'_i, R_{-i}) P_i f(R_N)$. This is a contradiction to AM-proofness. ■

Proof of Theorem 4.1. Let f be an SCF on $\mathcal{D} \subset \mathcal{AL}$ satisfying AM-proofness. I will show that f is also strategy-proof. Let R_N be any preference profile in $\mathcal{D}^N$, let i be any agent, and let R'_i be any preference relation in $\mathcal{D}$. Let L denote the set $\{y \in X \mid f(R_N) R_i y\}$, i.e., the lower contour set of $f(R_N)$ with respect to R_i. The goal is to prove $f(R'_i, R_{-i}) \in L$.

Let $(R^1 \ldots R^\ell)$ be a path in $\mathcal{D}$ without restoration from R_i to R'_i. By Lemma 4.1, for each h $(1 \leq h \leq \ell - 1)$, when agent i changes his preference relation from R^h to R^{h+1}, a social outcome can change only if either

Case 1: an alternative just below $f(R^h, R_{-i})$, say x, at R^h becomes indifferent to $f(R^h, R_{-i})$ at R^{h+1}, or

Case 2: an alternative indifferent to $f(R^h, R_{-i})$, say x, at R^h is just above $f(R^h, R_{-i})$ at R^{h+1}.

(Note that in Cases 1 and 2, from R^h to R^{h+1}, there should be no other changes because they are adjacent.) These two cases are described in Table 5. In both cases, by AM-proofness, the only candidates for $f(R^{h+1}, R_{-i})$ are $f(R^h, R_{-i})$ and x.

In the following, I prove that for each h $(1 \leq h \leq \ell - 1)$, if $f(R^h, R_{-i}) \in L$, then $f(R^{h+1}, R_{-i}) \in L$. Thus, assume $f(R^h, R_{-i}) \in L$. If $f(R^h, R_{-i}) = f(R^{h+1}, R_{-i})$, then $f(R^{h+1}, R_{-i})$ is trivially in L. Thus, assume $f(R^h, R_{-i}) \neq f(R^{h+1}, R_{-i})$. By the above arguments, it suffices to consider Case 1 and Case 2 of Table 5.

First, consider Case 1. Remember that the only candidates for $f(R^{h+1}, R_{-i})$ are $f(R^h, R_{-i})$ and x, the alternative just below $f(R^h, R_{-i})$. Then, if $x \in L$, whichever alternative $f(R^{h+1}, R_{-i})$ is, $f(R^{h+1}, R_{-i}) \in L$. Thus, assume $x \notin L$. In other words, $xP^1 f(R_N)$. (Remember that $R^1 = R_i$, and L is the lower contour set of $f(R_N)$ with respect to R_i.) Because $f(R_N)R^1 f(R^h, R_{-i})$, by transitivity, $xP^1 f(R^h, R_{-i})$ holds. Also, because x is ranked lower than $f(R^h, R_{-i})$ at R^h, x and $f(R^h, R_{-i})$ should be indifferent at some preference relation between R^1 and R^h in the path; there exists $k < h$ such that $xI^k f(R^h, R_{-i})$. In sum, $xP^1 f(R^h, R_{-i})$, $xI^k f(R^h, R_{-i})$, $f(R^h, R_{-i})P^h x$, and $xI^{h+1} f(R^h, R_{-i})$, which is a contradiction to the assumption that the path under consideration is *without* restoration. Therefore, the case $x \notin L$ cannot occur.

Next, consider Case 2. By the same reasoning as in Case 1, if $x \in L$, then $f(R^{h+1}, R_{-i}) \in L$. Thus, assume $x \notin L$, i.e., $xP^1 f(R_N)$. Because $f(R_N)R^1 f(R^h, R_{-i})$, by transitivity, $xP^1 f(R^h, R_{-i})$ holds. Then, $xP^1 f(R^h, R_{-i})$, $xI^h f(R^h, R_{-i})$, and $xP^{h+1} f(R^h, R_{-i})$, which is a contradiction to the assumption that the path under consideration is *without* restoration. Therefore, the case $x \notin L$ cannot occur.

Thus, in any case, $f(R^h, R_{-i}) \in L$ implies $f(R^{h+1}, R_{-i}) \in L$. Trivially, $f(R^1, R_{-i}) = f(R_N) \in L$ holds. Then, by the induction on h, $f(R^\ell, R_{-i}) = f(R'_i, R_{-i}) \in L$ holds, which completes the proof. ■

Based on Theorem 4.1, I prove that on $\mathcal{AL}$, the set of all linear and almost linear preferences, any SCF satisfies AM-proofness iff it satisfies strategy-proofness. Thus, the equivalence between strategy-proofness and AM-proofness which does

Table 6: Raising x one position from R to R'.

Case 1		Case 2	
R	R'	R	R'
$\vdots$	$\vdots$	$\vdots$	$\vdots$
y	xy	xy	x
x	$\vdots$	$\vdots$	y
$\vdots$			$\vdots$

not hold on $\mathcal{W}$ revives if deviation from linear orders is limited to the slightest one.

For this purpose, I would like to make clear the following terminology which will be repeatedly used in the construction of a path in $\mathcal{AL}$. Given a preference relation $R \in \mathcal{AL}$, the resulting preference relation R' from "raising an alternative x by one position" is one of the following:

Case 1: when $r_k(R) = x$ and $r_{k-1}(R) = y$ for some k and for some $y \in X$, then x and y becomes indifferent, i.e., $xI'y$, and the other parts of R are unchanged.

Case 2: when $r_k(R) = \{x, y\}$ for some k and for some $y \in X$, then $r_k(R') = x$ and $r_{k+1}(R') = y$, and the other parts of R are unchanged.

I represent these two cases in Table 6. It can be seen that if R is linear, then R' is almost linear, and if R is almost linear, then R' is linear. I would like to note that in Case 2, x is a kth ranked alternative in both R and R' though I say "x is raised by one position". In either case, I will say that x *improves the position relative to* y from R to R'.

Remark 4.1

Let R be any linear order and let x be any alternative. If x is raised by one position

Table 7: Preference relations in Example 4.1

R	Raise x by one position			
y	y	y	$\mathbf{x}y$	$\mathbf{x}$
z	$\mathbf{x}z$	$\mathbf{x}$	z	y
$\mathbf{x}$		z		z

until it becomes a unique top, then all preference relations appear in this process are in $\mathcal{AL}$, and the starting and the end points are both linear orders.

Example 4.1

Let $X = \{x, y, z\}$, and consider R in Table 7. Then, from the second to the last column of Table 7, x is raised by one position at a time until it becomes a unique top.

Corollary 4.1

On $\mathcal{AL}$, AM-proofness and strategy-proofness are logically equivalent.

Proof. To prove the statement of this corollary, by Theorem 4.1, it suffices to show that for any two preference relations in $\mathcal{AL}$, there is a path without restoration between them in $\mathcal{AL}$.

Let R and R' be any two distinct elements of $\mathcal{AL}$. First, assume that R and R' are both linear orders, and I construct a path without restoration from R to R'. Let $R^1 = R$.

STEP 1: Compare $r_1(R^1)$ and $r_1(R')$. If they are identical, then proceed to the next step. Assume $r_1(R^1) \neq r_1(R')$. Then, in R^1, raise $r_1(R')$ one position at a time until $r_1(R')$ becomes a unique top. Then, number every preference relation in this process in order of appearance.

I explain this process by an example. Consider R^1 and R' in Table 8. Then, $r_1(R')$ in R^1 is raised by one position from R^1 to R^2, and one more position from

Table 8: Step 1 of the proof of Corollary 4.1

		Raise $r_1(R')$ one position at a time in R			
$R(=R^1)$	R'	R^2	R^3	R^4	R^5
x	$r_1(R')$	x	x	$r_1(R')\ x$	$r_1(R')$
y	z	$r_1(R')\ y$	$r_1(R')$	y	x
$r_1(R')$	w	$\vdots$	y	$\vdots$	y
$\vdots$	$\vdots$		$\vdots$		$\vdots$

R^2 to R^3, and so on. Note that the process does not end until $r_1(R')$ becomes a unique top. Thus, R^4 is not the endpoint. The last preference relation, R^5, is the outcome of Step 1. These preference relations (R^2 to R^5 in Table 8) will constitute a part of a path from R to R'.

STEP k: Let R^a denote the resulting preference relation from Step $(k-1)$. Compare $r_k(R^a)$ and $r_k(R')$. If they are identical, then proceed to the next step. Assume $r_k(R^a) \neq r_k(R')$. Then, raise $r_k(R')$ one position at a time in R^a until $r_k(R')$ becomes a unique kth ranked alternative. Then, number the preference relations from $a+1$ to $a+b$, where b is the number of preference relations constructed in this step.

This step-by-step transformation terminates at Step $m-1$. Let R^ℓ denote the outcome of Step $m-1$. I claim that the sequence $(R^1 \dots R^\ell)$ forms a path without restoration from R to R' in $\mathcal{AL}$.

Note that in this transformation process, an alternative x improves the position relative to an alternative y only if x is ranked higher than y in R'. (Assume that x improves the position relative to y in Step k. Then, $x = r_k(R')$ and y is *not* one of the top k alternatives with respect to R'.) Thus, if $xP'y$ and xP^ky for some k, then $xP^{k'}y$ for all $k' > k$. This fact will be used repeatedly in the following.

First, I show that $R^1 = R$ and $R^\ell = R'$. The relation $R^1 = R$ follows from the definition of R^1. Thus, I prove $R^\ell = R'$. Let k be any element of $\{1, \dots, m-2\}$. Assume that the restriction of the outcome of Step k, say R^a, to the set $\{r_1(R^a), \dots, r_k(R^a)\}$ is equal to $R'|\{r_1(R'), \dots, r_k(R')\}$. Let $R^{a'}$ denote the outcome of Step $k+1$. The top k alternatives of R^a are unchanged in Step $k+1$, i.e., $R^{a'}|\{r_1(R^{a'}), \dots, r_k(R^{a'})\} = R'|\{r_1(R'), \dots, r_k(R')\}$. Moreover, the $(k+1)$th ranked alternative of $R^{a'}$ is equal to that of R'. Thus, $R^{a'}|\{r_1(R^{a'}), \dots, r_k(R^{a'}), r_{k+1}(R^{a'})\} = R'|\{r_1(R'), \dots, r_k(R'), r_{k+1}(R')\}$. By the induction on k, we can conclude that $R^\ell = R'$.

I show that all preference relations are in $\mathcal{AL}$. By the construction, a linear order and an almost linear order appear alternately in the sequence $(R^1 \dots R^\ell)$. Thus, $R^h \in \mathcal{AL}$ for all $h \in \{1, \dots, \ell\}$.

Because every R^h $(h \in \{2, \dots, \ell\})$ is the resulting preference relation from raising some alternative by one position in R^{h-1}, it follows that $R^h \in A(R^{h-1})$.

Next, I prove that the sequence is without restoration. Assume that there exist $x, y \in X$ and k, k' with $k < k'$ such that xP^ky and $xI^{k'}y$. This implies $yP'x$. To see this, suppose $xP'y$. Then, as I noted above, $xP'y$ and xP^ky together imply $xP^{k'}y$, which is a contradiction. Thus, $yP'x$ holds. The relation $xI^{k'}y$ implies either $xP^{k'+1}y$ or $yP^{k'+1}x$. In the former case, x improves the position relative to y. However, x improves the position relative to y only if x is ranked higher than y in R', which is a contradiction to $yP'x$. Thus, $yP^{k'+1}x$ holds. Then, $yP'x$ and $yP^{k'+1}x$ imply $yP^{k''}x$ for all $k'' > k'$. Therefore, the sequence $(R^1 \dots R^\ell)$ is without restoration.

Finally, I prove that any two preference relations in the sequence are distinct from each other. Let k be any element of $\{1, \dots, \ell - 1\}$. From R^k to R^{k+1}, some alternative, say x, improves the position relative to some other alternative, say y. This implies that x is ranked higher than y in R', which implies that y never improves the position relative to x in the sequence. Thus, $R^k|\{x, y\} \neq R^{k'}|\{x, y\}$ for all $k' > k$, and hence $R^k \neq R^{k'}$ for all $k' > k$. Because k was arbitrary, any two preference relations in the sequence are distinct from each other.

So far, I have assumed that R and R' are linear orders. Now, consider the case where R or R' is not a linear order, i.e., almost linear order. Let $\succ$ be an arbitrary fixed linear order on X. If R is almost linear, then let $\bar{R}$ denote the linear order defined by for each distinct $x, y \in X$,

$$x\bar{P}y \iff \begin{bmatrix} xPy \\ \text{or} \\ xIy \text{ and } xP'y \\ \text{or} \\ xIy \text{ and } xI'y \text{ and } x \succ y \end{bmatrix}$$

The linear oder $\bar{R}$ is obtained by breaking the tie according to R' if possible, and if the tie cannot be broken by R', then it is broken by the fixed linear order $\succ$.

Similarly, if R' is almost linear, then let $\bar{R}'$ denote the linear order defined by for each distinct $x, y \in X$,

$$x\bar{P}'y \iff \begin{bmatrix} xP'y \\ \text{or} \\ xI'y \text{ and } xPy \\ \text{or} \\ xI'y \text{ and } xIy \text{ and } x \succ y \end{bmatrix}$$

The linear oder $\bar{R}'$ is obtained by breaking the tie according to R if possible, and if the tie cannot be broken by R, then it is broken by $\succ$.

Then, when R is almost linear, replace R by $\bar{R}$ in the above transformation process. Similarly, when R' is almost linear, replace R' by $\bar{R}'$. There are three possible cases.

Case 1: R and R' are almost linear.

Case 2: R is linear and R' is almost linear.

Case 3: R is almost linear and R' is linear.

First, I consider Case 1. Assume that R and R' are almost linear. Then, by replacing R and R' by $\bar{R}$ and $\bar{R}'$, respectively, in the transforming process, we have a path without restoration $(R^1 \dots R^\ell)$ from $\bar{R}$ to $\bar{R}'$ in $\mathcal{AL}$. (Note that $R^1 = \bar{R}$ and $R^\ell = \bar{R}'$.) I claim that the sequence $(RR^1 \dots R^\ell R')$ is a path without restoration from R to R' in $\mathcal{AL}$.

Because $R, R' \in \mathcal{AL}$, the sequence is in $\mathcal{AL}$. Because $R^1 \in A(R)$ and $R' \in A(R^\ell)$, every preference relation is adjacent to the preceding one in the sequence.

I prove that the sequence $(RR^1 \dots R^\ell R')$ is without restoration. Let a and b denote the distinct alternatives such that aIb and let c and d denote the distinct alternatives such that $cI'd$. Then, either aP^1b or bP^1a holds. Similarly, either $cP^\ell d$ or $dP^\ell c$ holds. Without loss of generality, assume that aP^1b and $cP^\ell d$ hold. Note that for any distinct two alternatives x and y such that $\{x, y\} \neq \{a, b\}$ and $\{x, y\} \neq \{c, d\}$, $R|\{x, y\} = R^1|\{x, y\}$ and $R'|\{x, y\} = R^\ell|\{x, y\}$. This implies that any restoration of preferences does not occur over such x and y. (Remember that $(R^1 \dots R^\ell)$ is a path without restoration.) Thus, it suffices to check whether the restoration of preferences occurs over $\{a, b\}$ or $\{c, d\}$. First, consider whether the restoration of preferences occurs over a and b in the sequence $(RR^1 \dots R^\ell R')$. The relation aP^1b implies either $aP'b$ or [$aI'b$ and $a \succ b$]. In either case, $aP^\ell b$ holds. aP^1b and $aP^\ell b$ together imply that aP^kb for all $k \in \{1, \dots, \ell\}$. Then, it is clear that the restoration of preferences cannot occur over a and b in the sequence $(RR^1 \dots R^\ell R')$. By the symmetric argument, it can be proved that the restoration of preferences cannot occur over c and d.

Finally, I prove that any two preference relations in the sequence are distinct from each other. Because it is known that any two preference relations in the path $(R^1 \dots R^\ell)$ are distinct from each other, it suffices to prove that R and R' are distinct from any one in $\{R^1, \dots, R^\ell\}$. I prove that $R \notin \{R^1, \dots, R^\ell\}$. Let a and b denote the distinct alternatives such that aIb. At R^1, the tie is broken and either aP^1b or bP^1a holds. Without loss of generality, assume aP^1b. There are two possible cases. The first case is $aP'b$, and the second case is $aI'b$ and $a \succ b$. In either case, $aP^\ell b$ holds. aP^1b and $aP^\ell b$ together imply that aP^kb for all $k \in \{1, \dots, \ell\}$. Thus, $R \neq R^k$ for all $k \in \{1, \dots, \ell\}$. By the symmetric argument, it can be proved that $R' \neq R^k$ for all $k \in \{1, \dots, \ell\}$.

I complete the proof for Case 1 in which both R and R' are almost linear. The proofs for the other cases, Cases 2 and 3, are "subsets" of that of Case 1, and I omit them. ■

5 Concluding remarks

In this paper, weak orders are considered as agents' preferences. It is shown that the main result by Sato (2010) critically relies on the fact that only linear orders are taken into account. At the same time, it is also shown that, if the extent of the deviation from linear orders is slightest, the equivalence between strategy-proofness and AM-proofness holds under a simple condition.

This paper would be a starting point for finding an answer to the question in the Introduction with weak orders. A natural line of research following this paper is to consider the case where agents are willing to misreport preferences within some fixed distance (greater than 1) from the sincere preference relation.

Appendix The minimum distance from the total indifference relation

As in Section 3.1, let R^* denote the weak order such that any two alternatives are indifferent at R^*, i.e., $r_1(R^*) = X$. In this appendix, I will show that for any weak order $\tilde{R}$ nearest to R^* in $\mathcal{W}$, $d_K(\tilde{R}, R^*) = m - 1$ holds, where m is the number of alternatives.

Let R be any weak order. Let $\mathcal{X} = \{\{x, y\} \mid x, y \in X,\ x \neq y\}$. Let ℓ denote the number of the indifference classes induced by R.

Define

- $\mathcal{X}_1 = \{\{x, y\} \in \mathcal{X} \mid \{x, y\} \cap r_1(R) \neq \emptyset\}$.
- $\mathcal{X}_2 = \{\{x, y\} \in \mathcal{X} \setminus \mathcal{X}_1 \mid \{x, y\} \cap r_2(R) \neq \emptyset\}$.

  $\cdots$
- $\mathcal{X}_k = \{\{x, y\} \in \mathcal{X} \setminus (\mathcal{X}_1 \cup \cdots \cup \mathcal{X}_{k-1}) \mid \{x, y\} \cap r_k(R) \neq \emptyset\}$.

  $\cdots$
- $\mathcal{X}_\ell = \{\{x, y\} \in \mathcal{X} \setminus (\mathcal{X}_1 \cup \cdots \cup \mathcal{X}_{\ell-1}) \mid \{x, y\} \cap r_\ell(R) \neq \emptyset\}$.

Then, $\mathcal{X}_1, \ldots, \mathcal{X}_\ell$ form a partition of $\mathcal{X}$, and $d_K(R, R^*)$ can be written as

$$d_K(R, R^*) = \sum_{\{x,y\} \in \mathcal{X}_1} s_{R,R^*}(\{x, y\}) + \sum_{\{x,y\} \in \mathcal{X}_2} s_{R,R^*}(\{x, y\}) + \cdots + \sum_{\{x,y\} \in \mathcal{X}_\ell} s_{R,R^*}(\{x, y\}). \tag{A.1}$$

Consider the following transformation process from R^* to R.

STEP 1: From X, lift the set $r_1(R)$ above.

STEP 2: From $X \setminus r_1(R)$, lift the set $r_2(R)$ between $r_1(R)$ and $X \setminus (r_1(R) \cup r_2(R))$.

$\cdots$

STEP k: From $X \setminus (r_1(R) \cup \cdots \cup r_{k-1}(R))$, lift the set $r_k(R)$ between $r_{k-1}(R)$ and $X \setminus (r_1(R) \cup \cdots \cup r_{k-1}(R) \cup r_k(R))$.

. . .

STEP $\ell-1$: From $X \setminus (r_1(R) \cup \cdots \cup r_{\ell-2}(R))$, lift the set $r_{\ell-1}(R)$ between $r_{\ell-2}(R)$ and $r_\ell(R)$.

Each step transforms a weak order into another weak order. Because R was arbitrary in $\mathcal{W}$, R^* can be transformed to any weak order according to the above steps.

I show that the kth term in the right hand side of (A.1) corresponds to the distance arises in Step k. The distance arises in Step k is the cardinality of the set $\{\{x,y\} \in \mathcal{X} \mid x \in r_k(R), xPy\}$. Let A denote this set. For any $\{x,y\} \in A$, $s_{R,R^*}(\{x,y\}) = 1$. Thus, $\sum_{\{x,y\}\in A} s_{R,R^*}(\{x,y\}) = |A|$. It can be seen that $A \subset \mathcal{X}_k$ and that for any $\{x,y\} \in \mathcal{X}_k \setminus A$, xIy holds, i.e., $s_{R,R^*}(\{x,y\}) = 0$. Therefore, $\sum_{\{x,y\}\in \mathcal{X}_k} s_{R,R^*}(\{x,y\}) = \sum_{\{x,y\}\in A} s_{R,R^*}(\{x,y\}) = |A| =$ [the distance arises in Step k].[10] Moreover, the distance arises in each step is positive. (For any $k \in \{1, \ldots, \ell-1\}$, there is at least one element $\{x,y\}$ of $\mathcal{X}$ such that $x \in r_k(R)$ and xPy.)

Now, let $\tilde{R}$ be one of the nearest weak orders to R^*. Then, it is clear that from R^* to $\tilde{R}$, there is only one step in the above transformation process. (If there are more than one step, then by stopping at Step 1, we can have a weak order nearer to R^* than $\tilde{R}$, which is a contradiction.) Let a denote the cardinality of $r_1(\tilde{R})$. Then, the distance between $\tilde{R}$ and R^* is $a(m-a)$. Because $\tilde{R}$ is nearest to R^*, this a should be either 1 or $m-1$. In any case, $d_K(\tilde{R}, R^*) = m-1$.

References

Gibbard, Allan (1973) Manipulation of voting scheme: A general result. Econometrica 41, 587–601.

[10]Note that the last term of (A.1) is 0. (For any $\{x,y\} \in \mathcal{X}_\ell$, $\{x,y\} \subset r_\ell(R)$ holds. Thus, xIy, and $s_{R,R^*}(\{x,y\}) = 0$.)

Kemeny, John G. (1959) Mathematics without numbers. Daedalus 88, 577–591.

Kemeny, John G.; Snell, J. Laurie (1962) Mathematical models in the social sciences. Blaisdell, New York.

Sato, Shin (2010) Adjacent manipulation. mimeo.

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