

Minimal Manipulability: Anonymity and Unanimity

Stefan Maus, Hans Peters, Ton Storcken

Date of this version: June 29, 2004

University of Maastricht, Department of Quantitative Economics,

P.O. Box 616, 6200 MD Maastricht, THE NETHERLANDS

Tel.: +31433883761; fax: +31433884874

Email: s.maus@ke.unimaas.nl, h.peters@ke.unimaas.nl,

t.storcken@ke.unimaas.nl

Abstract This paper is concerned with the minimal number of profiles at which a unanimous and anonymous social choice function is manipulable. The lower bound is derived when there are three alternatives to choose from. Examples of social choice functions attaining the lower bound are given. We conjecture that these examples are in fact all minimally manipulable social choice functions. Since some of these examples are even Pareto optimal, we have also derived the lower bound for Pareto optimal and anonymous social choice functions. Some of the minimally manipulable Pareto optimal and anonymous social choice functions can be interpreted as status quo voting.

1 Introduction

A well-known result of Gibbard (1973) and Satterthwaite (1975) states that any strategy-proof surjective social choice function on more than two alternatives is dictatorial. This implies that anonymity and unanimity are not compatible with strategy-proofness. Here we investigate how incompatible these are, i.e. how much manipulability we have to allow at anonymous and unanimous social choice functions.

There are a few studies on the degree of vulnerability for strategic behaviour of classical social choice functions, such as Borda, Plurality, etc.. For example, Aleskerov (1999) contains simulation and enumeration results on 26 different social choice functions for different indices of manipulability. Slinko (2002) counts the number of instabile profiles of classical social choice functions, which is an upper bound for the number of manipulable profiles of these social choice functions. On the other hand, little is known about to what extent manipulability has to be admitted by social choice functions satisfying certain properties. Results on this may provide lower bounds to which then all social choice functions satisfying these properties can be compared. A first investigation was pioneered by Kelly (1988), who found the minimal

number of manipulable profiles for nondictatorial surjective social choice functions with three alternatives and two agents and also formulated several conjectures. This line of research was continued by Fristrup and Keiding (1998), who determined the minimal number of manipulable profiles for two agents and any number of alternatives. Maus et al. (2004b) consider the three alternative case and show that for any number of agents larger than two there are six minimally manipulable nondictatorial surjective social choice functions, and these are even anonymous. However, they also consider minimally manipulable unanimous and nondictatorial social choice functions. These turn out to be nondictatorial only in a minimal sense, namely at only one profile.

Here we will exclude such social choice functions by demanding that agents are treated anonymously and choice is unanimous. Maus et al. (2004a) contain results for this case if unanimity is replaced with surjectivity. In particular, anonymous social choice functions are nondictatorial. So, we are interested in the minimal number of manipulable profiles that an anonymous unanimous social choice function has to admit. It is well known that in case of two alternatives any monotonic social choice function is strategy-proof. As monotonicity and anonymity are compatible (see e.g. May (1952)) in case of two alternatives this minimal number is zero. We consider the three alternatives case and show that the minimal number is $2^n - 2$. Furthermore, we give examples of social choice functions reaching that lower bound and point out that some of these examples are even Pareto optimal. So, $2^n - 2$ is also the lower bound for Pareto optimal and anonymous social choice functions. After that we argue that status quo voting is an example of a Pareto optimal and anonymous social choice function that is minimally manipulable for three alternatives.

It should be noted that the way in which we measure manipulability, namely by counting manipulable profiles, is not the only possible one. In fact there are a lot of variations of this, where for example profiles are counted by the number of agents that can manipulate them, or the severity of the manipulation is taken into account. The severity is measured by the distance in the agent's ranking between what he achieves with and without manipulation. Aleskerov (1999) contains an overview of such variations and numerical results on the manipulability of 26 social choice functions according to these measurements.

The paper is organized as follows. Section 2 contains the basic definitions. Then, Section 3 shows some special results about the manipulability of two and three agent social choice functions that will be used in our proof of the main result. The main result, saying that any unanimous and anonymous social choice function over three alternatives has at least $2^n - 2$ manipulable profiles, where n is the number of agents, is proven in Section 4. After that we provide examples of social choice functions attaining that lower bound in Section 5. In Section 6 we see that some of the examples from Section 5 are even Pareto optimal. Pareto optimality is a stronger requirement than unanimity. This shows, given anonymity, that though Pareto optimality is stronger it does not raise the number of manipulable profiles one has to admit

compared to imposing unanimity. Finally, we show that the minimally manipulable Pareto optimal and anonymous social choice functions have a nice interpretation as status quo voting rules.

2 Preliminaries

We denote the cardinality of a set S by $|S|$. Let \mathbb{N} be the set of natural numbers, and let $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. Let $\lceil \cdot \rceil$ denote the upper entier. So for $n \in \mathbb{N}_0$ $\lceil \frac{n}{2} \rceil := \min\{l \in \mathbb{N}_0 \mid l \geq \frac{n}{2}\}$.

Let A be a finite set of alternatives, $m := |A| \geq 3$, and $N = \{1, \dots, n\}$ be a finite set of agents, $n \geq 2$.

Let $t \subset A \times A$. We call t *complete* if for all $x, y \in A$ $(x, y) \in t$ or $(y, x) \in t$. Note that completeness of t implies $(x, x) \in t$ for all $x \in A$. We call t *transitive* if for all $x, y, z \in A$ $(x, y) \in t$ and $(y, z) \in t$ implies $(x, z) \in t$. We call t *antisymmetric* if for all $x, y \in A$ $(x, y) \in t$ and $(y, x) \in t$ implies that $x = y$.

A *preference* $t \subset A \times A$ is a *linear order* (complete, transitive, antisymmetric) on A . Let P denote the set of all preferences. Suppose that $A = \{x_1, x_2, \dots, x_m\}$. By completeness, transitivity and antisymmetry we can write conveniently

$$t = x_1 x_2 \dots x_m$$

for the preference t such that $(x_i, x_j) \in t$ if and only if $i \geq j$, $i, j \in \{1, 2, \dots, m\}$,

$$t = \dots x \dots y \dots$$

if we want to express only that x is strictly preferred to y , and

$$t = x \dots$$

if we want to express only that x is preferred to all other alternatives.

A *profile* is a map $p : N \rightarrow P$. Thus, a profile assigns to every agent i a preference $p(i)$ over the alternatives. For a nonempty subset S of N we denote by $p|_S$ the restriction of the map p to the domain S . The set of all profiles is denoted by P^N .

For a permutation σ of N and a profile $p \in P^N$ let $p \circ \sigma$ be the profile given by $(p \circ \sigma)(i) := p(\sigma(i))$ for all $i \in N$. Two profiles $p, q \in P^N$ are called *anonymously equivalent* if there is a permutation σ of N such that $q = p \circ \sigma$. Then $|p^{-1}(t)| = |q^{-1}(t)|$ for all $t \in P$. We denote anonymous equivalence between p and q in P^N by $p \sim q$ and write $[p] := \{q \in P^N \mid q \sim p\}$. For all $p \in P^N$,

$$|[p]| := \frac{n!}{\prod_{t \in P} |p^{-1}(t)|!}.$$

A *social choice function* is a function $f : P^N \rightarrow A$. Hence, a social choice function selects a unique alternative $f(p)$ at every profile p .

A social choice function is called *surjective* if any alternative in A is chosen at least once, i.e. if $f(P^N) = A$. In literature this is also known as citizen-sovereignty. Moreover, a social choice function is called *unanimous* if $f(p) = x$ for all profiles $p \in P^N$ such that $p(i) = x \dots$ for all $i \in N$. Clearly, unanimity is stronger than surjectivity. Throughout this paper we assume that any social choice function is at least surjective.

A social choice function is called *anonymous* if $f([p]) = \{f(p)\}$ for all $p \in P^N$. Thus, anonymous social choice functions are symmetric in the arguments. In a sense they treat agents equally.

In contrast to anonymity, the following social choice functions respect only the preference of a single agent. Let $d \in N$ and for any profile p let

$$dict_d(p) := x$$

where x is such that $(x, y) \in p(d)$ for all $y \in A$. So x is the most preferred alternative of agent d in p . The social choice functions $f = dict_d$, $d \in N$, are called *dictatorial* social choice functions, and agent d is called a *dictator*. A social choice function f is called *nondictatorial* if there is no agent d such that $f = dict_d$.

We are interested in strategic behaviour of individuals when facing cooperative decision-making as captured by social choice functions. This is formalized by the following definitions. Let $f : P^N \rightarrow A$ be a social choice function. Let $p \in P^N$ be a profile. Then each profile q such that, for some $i \in N$, $q|_{N-\{i\}} = p|_{N-\{i\}}$ and $q(i) \neq p(i)$, is called an *i-deviation* from p . Letting $t := q(i)$ we use the notation $q = (p_{-i}, t)$. If it is not important which agent deviates from p to q we call q a *deviation* from p . A profile p is called *manipulable* (under f) if there is an agent that is better off by being dishonest about his preference, i.e. if there is an $i \in N$, and an *i-deviation* q such that

$$(f(p), f(q)) \notin p(i).$$

In this case we say that p is *manipulable towards* q (under f). Let

$$M_f := \{p \in P^N \mid p \text{ is manipulable under } f\}.$$

Note that $p \in M_f$ implies that $[p] \subseteq M_f$. A social choice function is called *strategy-proof* if $M_f = \emptyset$, otherwise it is said to be *manipulable*. Let F be a nonempty set of social choice functions. We call f^* *minimally manipulable* if $f^* \in \arg \min_{f \in F} |M_f|$.

The prominence of the dictatorial rules arises from the following impossibility result due to Gibbard (1973) and Satterthwaite (1975).

Theorem 1 *Let A be a finite set of alternatives, $|A| \geq 3$. Let $f : P^N \rightarrow A$ be a nondictatorial surjective social choice function. Then*

$$|M_f| \geq 1.$$

However, this theorem makes no statement about the number of manipulable profiles such social choice functions admit. This question has been solved by Kelly (1988) for two agents and three alternatives, and by Fristrup and Keiding (1998) for two agents, and any number of alternatives larger than three. Maus et al. (2004b) consider the case of unanimous and nondictatorial social choice function for any numbers of agents and alternatives larger than or equal to three, and the case of surjective and nondictatorial social choice functions for three alternatives. Maus et al. (2004a) consider the case of anonymous and surjective social choice functions, but have to leave some cases open.

Here we address the case of unanimous and anonymous social choice functions over three alternatives. So, let F be the set of all such social choice functions with a set of agents N such that $|N| = n \geq 2$. We will need the following simple lemma.

Lemma 2 *Let $k, l \in \mathbb{N}$. Then*

$$k!! \leq (k + l - 1)!$$

Proof. $\frac{(k+l-1)!}{k!!} = \frac{1}{k+l} \binom{k+l}{k} \geq 1$ as $k + l > k > 0$. ■

3 Special results for two and three agents

Since the connection of the results of this section with the main result is far from obvious we sketch the proof of the main result beforehand.

Let $f \in F$. We will prove that there are natural numbers k_1, k_2 such that $\frac{n}{2} \leq k_1 \leq k_2 \leq n$, and profiles $q_i^k \in M_f$ for all $\frac{n}{2} \leq k \leq k_i - 1, k \in \mathbb{N}, i \in \{1, 2\}$ with the following properties. Each such q_i^k shows that there are at least $\binom{n}{k}$ manipulable profiles in M_f , which can be equal to the ones from another q_i^k only if $k = \frac{n}{2}$ for both profiles. Furthermore, we show that if $k_i < n, i \in \{1, 2\}$, then there is a $p_i \in M_f$ that yields at least another $\sum_{j=k_i}^{n-1} \binom{n}{j}$ different manipulable profiles in M_f . Altogether we will have that in any case

$$|M_f| \geq \sum_{j=1}^{n-1} \binom{n}{j} = 2^n - 2,$$

which is our main result. We find the q_i^k by applying Theorem 3 to two agent social choice functions $f_k, \frac{n}{2} \leq k < k_2$, defined from f . To show that the manipulable profiles arising from the same $f_k, k \neq \frac{n}{2}$, are different we need the statements of Lemma 4. To apply Theorem 3 and Lemma 4 we need that the $f_k, k < k_2$, are nondictatorial. This will be shown by applying Lemma 5 to three person social choice function g defined from f . This will also show that the p_i exist if $k_i < n$.

Now, we present the mentioned results for two and three agent social choice functions. Throughout the rest of this section let $N = \{1, 2\}$ if there are two agents, and $N = \{1, 2, 3\}$ if

there are three agents. Two agent social choice functions can be depicted in 6x6 tables, since we have only six preferences. We let agent 1's preference be constant in rows and agent 2's preference constant in columns. The following result is taken from Kelly (1988). It can also be proven by considering 6x6 tables.

Theorem 3 *Let $f : P^N \rightarrow A$ be a surjective two agent social choice function. Then f is either dictatorial or $|M_f| \geq 2$.*

Let $D_f : M_f \rightarrow 2^{N \times P}$ be the correspondence assigning to each manipulable profile the set of agents that can manipulate and the preferences they can use to manipulate, i.e. for all $p \in M_f$,

$$D_f(p) := \{(i, t) \in N \times P \mid (f(p), f(p_{-i}, t)) \notin p(i)\}.$$

The first result shows some properties of D_f for unanimous and nondictatorial two agent social choice function.

Lemma 4 *Let $f : P^N \rightarrow A$ be a unanimous and nondictatorial two agent social choice function. Then for all $p, q \in M_f$ there are $(i, t_p) \in D_f(p), (j, t_q) \in D_f(q)$ such that $t_p \notin p(N), t_q \notin q(N)$. Furthermore we can choose i, j, t_p and t_q in such a way that, either*

(a) $i \neq j$,

(b) $p(N) \cup \{t_p\} \neq q(N) \cup \{t_q\}$, or,

(c) there is a labeling of the alternatives in A such that

$$\begin{aligned} p &= (bca, cab), (i, cba) \in D_f(p) \\ q &= (cba, bca), (i, cab) \in D_f(q) \\ f(p) &= a. \end{aligned}$$

Proof. By Theorem 1 we can find $p, q \in M_f$. First of all we show that for any $p \in M_f$ and $(k, t_p) \in D_f(p)$ such that $t_p \in p(N)$, there is a preference $\tilde{t}_p \notin p(N)$ such that $(k, \tilde{t}_p) \in D_f(p)$. This shows that we can choose $(i, t_p) \in D_f(p), (j, t_q) \in D_f(q)$ such that $t_p \notin p(N), t_q \notin q(N)$. Suppose that without loss of generality $k = 1$. So $(1, t_p) \in D_f(p)$ and $t_p \in p(N)$, hence $p = (p(1), t_p)$. Let $\tilde{t}_p \neq t_p$ be the preference such that $t_p = x \dots$ and $\tilde{t}_p = x \dots$. Then $f(\tilde{t}_p, t_p) = f(t_p, t_p)$ by unanimity of f . So $(1, \tilde{t}_p) \in D_f(p)$ as $(1, t_p) \in D_f(p), \tilde{t}_p \notin p(N) = \{p(1), t_p\}$.

Note that in (a) $p = q$ is not excluded. Suppose that (a) and (b) are not true. Then all $p, q \in M_f$ are manipulable by the same agent i , without loss of generality let $i = 1$.

Let $O_2(t) := \{f(p) \mid p(1) = t\}$ be the set of all opportunities that agent 2 has when the preference of agent 1 is fixed to t . Let $A = \{x, y, z\}$. Consider $t = xyz$ in P . Then $x \in O_2(t)$ by unanimity. Note that fixing agent 1's preferences corresponds to looking at rows only in the

tables describing f . Agent 2's inability to manipulate implies then that we can only have a certain type of rows.

	2 :	xyz	xzy	yxz	yzx	zxy	zyx
1. $O_2(xyz) = \{x\}$	1 : xyz	x	x	x	x	x	x
2. $O_2(xyz) = \{x, y\}$	1 : xyz	x	x	y	y	x	y
3. $O_2(xyz) = \{x, z\}$	1 : xyz	x	x	x	z	z	z
4. $O_2(xyz) = A$	1 : xyz	x	x	y	y	z	z

Similar tables hold for any $O_2(t), t \in P$.

Now, we fix $\{x, y, z\}$ such that $p(1) = xyz$. Remember that $p, q \in M_f$ are such that agent 1 manipulates at p and q and that

$$\{xyz, p(2), t_p\} = \{q(1), q(2), t_q\}.$$

Step 0: $x \in O_2(xyz)$.

This is a consequence of unanimity.

Step 1: $O_2(xyz) \neq \{x\}$.

If $O_2(xyz) = \{x\}$ then agent 1 cannot manipulate at p .

Step 2: $O_2(xzy) \neq \{x\}$.

Suppose contrapositive that $O_2(xzy) = \{x\}$. By step 1 there is an $u \in O_2(xyz) - \{x\}$. Let $t_1 \neq t_2$ be the two preferences such that $t_i = u \dots, i = 1, 2$. Then $f(xyz, t_i) = u, i \in \{1, 2\}$, since agent 2 cannot manipulate, and so (xyz, t_i) is manipulable by agent 1 towards (xzy, t_i) for $i \in \{1, 2\}$. But then (b) holds, contradicting our assumptions. Hence, $O_2(xzy) \neq \{x\}$.

Step 3: $z \in O_2(xzy)$.

In view of step 2 suppose contrapositive that $O_2(xzy) = \{x, y\}$. Then $f(xzy, zyx) = y$, since agent 2 cannot manipulate, and $f(zxy, zyx) = z$ by unanimity. So $r = (xzy, zyx)$ is manipulable by agent 1 towards $(zxy, zyx) = (t_r, r(2))$. As $p(1) = xyz \notin r(N) \cup \{t_r\}$ this contradicts our assumption that (b) does not hold.

Step 4: $z \in O_2(xyz)$.

In view of step 1 suppose contrapositive that $O_2(xyz) = \{x, y\}$. Then $f(xzy, zxy) = z$ and $f(xyz, zxy) = x$, since agent 2 cannot manipulate. Hence, $r = (xzy, zxy)$ is manipulable by agent 1 towards $(xyz, zxy) = (t_r, r(2))$. As (b) does not hold we have $p(N) \cup \{t_p\} = \{xyz, p(2), t_p\} = r(N) \cup \{t_r\} = \{xyz, xzy, zxy\}$. If $p(2) \in \{xyz, xzy\}$ then $f(p) = x$ by unanimity. If $p(2) = zxy$, then $f(p) = x$ because $O_2(xyz) = \{x, y\}$ and agent 2 cannot manipulate. But $f(p) = x$ contradicts the manipulability of p by agent 1. Hence, $z \in O_2(xyz)$.

Step 5: $z \in O_2(yxz)$ and $z \in O_2(yzx)$.

Suppose contrapositive that there is a $t \in \{yxz, yzx\}$ such that $z \notin O_2(t)$, so $O_2(t) \subseteq \{x, y\}$. Let $r^1 = (xyz, zxy), r^2 = (xyz, zyx)$. By step 4 $z \in O_2(xyz)$, and so $f(r^1) = f(r^2) = z$. But then $f(r_{-1}^1, t) \in \{x, y\}$ and $f(r_{-1}^2, t) \in \{x, y\}$, so r^1 and r^2 are manipulable towards (r_{-1}^1, t)

and (r_{-1}^2, t) respectively. Clearly, this contradicts the assumption that (b) is not satisfied. This proves step 5.

Step 6: $O_2(xzy) = \{x, z\}$.

Suppose contrapositive that in view of step 0 and 3 $O_2(xzy) = A$. By step 0 and 2 either $O_2(xyz) = \{x, z\}$ or $O_2(xyz) = A$. Consider first the case in which $O_2(xyz) = \{x, z\}$. Let $r^1 = (xzy, yxz), r^2 = (xzy, yzx)$. Then $f(r^1) = f(r^2) = y$, whereas $f(r_{-1}^1, xyz) = x$ and $f(r_{-1}^2, xyz) = z$. Hence, r^1 and r^2 are manipulable by agent 1 towards $(r_{-1}^1, xyz), (r_{-1}^2, xyz)$ respectively. But then (b) is satisfied a contradiction. So, suppose that $O_2(xyz) = A$. We know that p is manipulable by agent 1 and that $p(1) = xyz$. By steps 3 to 5, unanimity and because agent 2 cannot manipulate $f(P \times \{zxy, zyx\}) = \{z\}$, hence no manipulation by agent 1 can occur if $p(2) \in \{zxy, zyx\}$. By unanimity agent 1 has no incentive to manipulate if $p(2) \in \{xyz, xzy\}$. Hence, $p(2) \in \{yxz, yzx\}$. Then $f(p) = y$ and $f(t_p, p(2)) = x$, since agent 1 manipulates p towards $(t_p, p(2))$. As $O_2(xzy) = A$ implies that $f(xzy, p(2)) = y$ (agent 2 cannot manipulate), and unanimity implies that $f(yxz, p(2)) = f(yzx, p(2)) = y$ we obtain $t_p \in \{zxy, zyx\}$. As $f(t_p, p(2)) = x$ the profile $r = (xzy, p(2)), f(r) = y$ ($O_2(xyz) = A$ and agent 2 cannot manipulate), is also manipulable by agent 1. Clearly, then with p and r (b) is satisfied, a contradiction. Hence, $O_2(xzy) = \{x, z\}$.

Step 7: $y \notin O_2(zyx)$.

Suppose contrapositive that $y \in O_2(zyx)$. Then $r = (zyx, yzx), f(r) = y$ (agent 2 cannot manipulate), is manipulable towards $(xzy, yzx) = (t_r, r(2))$, as $f(xzy, yzx) = z$ by step 6 and since agent 2 cannot manipulate. As $p(1) = xyz \notin r(N) \cup \{t_r\}$ this contradicts the assumption that (b) does not hold.

Step 8: $O_2(zyx) = \{z\}$

By step 7 $y \notin O_2(zyx)$. So suppose contrapositive that $O_2(zyx) = \{x, z\}$. Then $r = (zyx, yxz), f(r) = x$ (agent 2 cannot manipulate), is manipulable towards $(yzx, yxz) = (t_r, r(2))$, as $f(yzx, yxz) = y$ by unanimity. As $p(1) = xyz \notin r(N) \cup \{t_r\}$ this contradicts the assumption that (b) does not hold.

Step 9: $O_2(zxy) = \{z\}$.

Suppose contrapositive that there is an $t \in P$ such that $f(zxy, t) \neq z$. Because of unanimity $t \notin \{zxy, zyx\}$. Then $r = (zxy, t)$ is manipulable towards $(zyx, t) = (t_r, r(2))$ by step 8. As $p(1) = xyz \notin r(N) \cup \{t_r\}$ this contradicts the assumption that (b) does not hold unless $p(1) = t = xyz$ and $\{p(2), t_p\} = \{zxy, zyx\}$. But in that case $f(p_{-1}, t_p) = f(t_p, p(2)) = z$ by unanimity, but agent 1 in $p = (xyz, p(2))$ has no incentive to manipulate to a profile (p_{-1}, t_p) where outcome z obtains. So, this contradicts the assumption that agent 1 manipulates at p towards (p_{-1}, t_p) . This finishes the proof of step 9.

Step 10: $O_2(yzx) = \{y, z\}$.

In view of step 5 suppose contrapositive that $O_2(yzx) = A$. Then $r = (yzx, xzy), f(r) = x$ (agent 2 cannot manipulate), is manipulable towards $(zxy, xzy) = (t_r, r(2))$ by step 9. As

$p(1) = xyz \notin r(N) \cup \{t_r\}$ this contradicts the assumption that (b) does not hold.

We summarize now what we know about f by steps 1 to 10 in the following 6x6 table.

2 :	xyz	xzy	yxz	yzx	zxy	zyx
1 : xyz	x	x			z	z
1 : xzy	x	x	x	z	z	z
1 : yxz			y	y	z	z
1 : yzx	y	z	y	y	z	z
1 : zxy	z	z	z	z	z	z
1 : zyx	z	z	z	z	z	z

We finish the proof by considering the two cases for $O_2(xyz) \in \{\{x, z\}, A\}$.

If $O_2(xyz) = \{x, z\}$ then $r = (xyz, yzx)$, $f(r) = z$, is manipulable towards $(yxz, yzx) = (t_r, r(2))$, and r is the only manipulable profile such that $r(1) = xyz$. Hence, $p = r$. If (b) is not satisfied we must have $q(N) \cup \{t_q\} = \{xyz, yxz, yzx\}$. Since $q(1) \neq xyz$ we have $q(1) \in \{yxz, yzx\}$. Since q has to be manipulable this leads to $q = (yxz, xyz)$, $t_q = yzx$. But then (c) holds, letting $i = 1$, $x = b$, $y = c$ and $z = a$.

On the other hand if $O_2(xyz) = A$, then $r = (xyz, yxz)$, $f(r) = y$, is manipulable towards $(xzy, yxz) = (t_r, r(2))$, and r is the only manipulable profile such that $r(1) = xyz$. Hence, $p = r$. If (b) is not satisfied we must have $q(N) \cup \{t_q\} = \{xyz, xzy, yxz\}$. Since $q(1) \neq xyz$ and clearly also $q(1) \neq xzy$ (otherwise $q \in \{(xzy, xyz), (xzy, yxz)\} \implies f(q) = x$) we have $q(1) = yxz$. So, $q \in \{(yxz, xyz), (yxz, xzy)\}$. Now, if $q = (yxz, xyz)$ then $t_q = xzy$, so agent 1 manipulates q towards (xzy, xyz) , where he obtains the outcome $x = f(xzy, xyz)$. This implies that x is preferred to $f(q)$ under the preference $q(1) = yxz$, hence $f(q) = z$. This is a contradiction, since then agent 2 can manipulate q as well. So, only the case $q = (yxz, xzy)$, $t_q = xyz$, remains. For the same reasons as in the other case the manipulability of q towards (xyz, xzy) by agent 1 implies that $f(q) = z$. But then (c) holds, letting $i = 1$, $x = c$, $y = b$, $z = a$, and exchanging the roles of p and q .

This finishes the proof of the lemma. ■

Furthermore, we need the following result about the manipulability of unanimous three agent social choice functions where coalitions are dictators. Let $S \subseteq N$ be a nonempty coalition of agents. We call S a *dictator* if for all $t = xyz \in P$ and $q_{-S} \in P^{N-S}$ we have

$$f(t^S, q_{-S}) = x.$$

In this case we call f *S-dictatorial*.

Lemma 5 *Let $f : P^N \rightarrow A$ be a unanimous three agent social choice function. Suppose that f is $\{1, 2\}$ -dictatorial and $\{1, 3\}$ -dictatorial. Then, one of the following is true.*

(a) *Agent 1 is a dictator.*

(b) There are anonymously inequivalent $p, q \in M_f$, manipulable towards $(p_{-i}, t_p), (q_{-j}, t_q), i, j \in \{1, 2, 3\}$, respectively, such that $|p(N)| = |q(N)| = 3$ and $p(N) \cup \{t_p\} \neq q(N) \cup \{t_q\}$.

Proof. Let $f_t^2 : P^{\{1,2\}} \rightarrow A, t \in P$, be the two agent social choice function given by $f_t^2(p) := f(p(1), p(2), t)$ for all $p \in P^{\{1,2\}}$. Let $f_t^3 : P^{\{1,3\}} \rightarrow A, t \in P$, be the two agent social choice function given by $f_t^3(p) := f(p(1), t, p(3))$ for all $p \in P^{\{1,3\}}$. Suppose that (b) is not true. We show that then all f_t^2 are dictatorial with dictator 1, so (a) holds. By unanimity and $\{1, 3\}$ -dictatoriality ($\{1, 2\}$ -dictatoriality) of f we have some information about f_t^2 and f_t^3 . We summarize this information in the following table for $t = xyz$, where agent 1's preference is as usually described by the rows. The other agent can be either agent 2 or agent 3, and the remaining of the three agents is fixed to the preference xyz .

2 (3) :	xyz	xzy	yxz	yzx	zxy	zyx
1 : xyz	x	x	x	x	x	x
1 : xzy	x	x				
1 : yxz			y			
1 : yzx				y		
1 : zxy					z	
1 : zyx						z

Let $r \in P^{\{1,2\}}$. Suppose that (b) is false. For steps 1 to 4 we omit the superscript $k = 2$.

Step 1: The second row can be filled with x , i.e.

$$f_{xyz}(xzy, t) = x \text{ for all } t \in P.$$

Suppose contrapositive that $u := f_{xyz}(xzy, t) \in \{y, z\}$ for some $t \in P$. Let $A = \{u, v, w\}$, and $p = (xzy, uvw, xyz), q = (xzy, uvv, xyz)$. If $f(p) \neq u$ then p is manipulable towards $(p_{-2}, t) = (p_{-2}, t_p)$, and if $f(p) = u$ then p is manipulable towards $(p_{-1}, xyz) = (p_{-1}, t_p)$. The same holds for q . Clearly, p and q are anonymously inequivalent and satisfy $|p(N)| = |q(N)| = 3$. Furthermore, $p(N) \cup \{t_p\} = q(N) \cup \{t_q\}$ implies that $t_p = uvv$ and $t_q = uvw$. As $u \neq x$ this implies that agent 2 is manipulating at p and q and so $t_p = t = t_q$, contradicting $t_p = uvv, t_q = uvw$. Hence, (b) is satisfied, a contradiction.

Step 2: f_{xyz} is unanimous.

In view of step 1 it remains to be shown that $f_{xyz}(yxz, yzx) = f_{xyz}(yzx, yxz) = y$ and $f_{xyz}(zxy, zyx) = f_{xyz}(zyx, zxy) = z$. But for example by applying step 1 to f_{yxz}^3 we obtain $f_{xyz}(yzx, yxz) = f_{yxz}^3(yzx, xyz) = y$. Likewise the other parts of unanimity of f_{xyz} follow.

Step 3: In the third and fourth row z cannot be chosen, i.e.

$$f_{xyz}(r) \neq z \text{ if } r(1) \in \{yxz, yzx\}.$$

Suppose contrapositive that $r(1) \in \{yxz, yzx\}$ and $f_{xyz}(r) = z$. Let $p = (r(1), zxy, xyz)$ and $q = (r(1), zyx, xyz)$. If $f(p) \neq z$ then p is manipulable towards $(p_{-2}, r(2)) = (p_{-2}, t_p)$, and if

$f(p) = z$ then p is manipulable towards $(p_{-3}, r(1)) = (p_{-3}, t_p)$, as by $\{1, 3\}$ -dictatorship of f we have $f(r(1), zxy, r(1)) = y$. The same holds for q . Clearly, p and q are anonymously inequivalent and satisfy $|p(N)| = |q(N)| = 3$. Furthermore, $p(N) \cup \{t_p\} = q(N) \cup \{t_q\}$ implies that $t_p = zyx$ and $t_q = zxy$. As $r(1) \in \{yxz, yzx\}$ this implies that agent 2 is manipulating at p and q and so $t_p = r(2) = t_q$, contradicting $t_p = zyx, t_q = zxy$. Hence, (b) is satisfied, a contradiction.

Step 4: In the cell (yzx, zxy) , y has to be chosen, i.e.

$$f_{xyz}(yzx, zxy) = y.$$

Let $p = (yzx, zxy, xyz)$ and $q = (yzx, zxy, xzy)$. By step 3 $f(p) \in \{x, y\}$. Suppose contrapositive to the claim of step 4 that $f(p) = x$. Then p is manipulable towards $(p_{-1}, zxy) = (p_{-1}, t_p)$ by $\{1, 2\}$ -dictatorship of f . If $f(q) \neq x$ then q is manipulable towards $p = (q_{-3}, xyz) = (q_{-3}, t_q)$, and if $f(q) = x$ then q is manipulable towards (q_{-1}, zxy) by $\{1, 2\}$ -dictatorship of f . Clearly, p and q are anonymously inequivalent and satisfy $|p(N)| = |q(N)| = 3$. As $xzy \in q(N)$ but $xzy \notin p(N) \cup \{t_p\}$ (b) is satisfied, a contradiction.

Step 5: In the fourth row y has to be chosen, i.e.

$$f_{xyz}(yzx, t) = y \text{ for all } t \in P.$$

We show step 5 for $k = 2$. The case $k = 3$ works in the same way. By step 3 we have that $f_{xyz}^2(yzx, t) \in \{x, y\}$ for all $t \in P$. Furthermore, $f_{xyz}^2(yzx, zyx) = f_{zyx}^3(yzx, xyz) \in \{x, y\} \cap \{y, z\} = \{y\}$ by step 3 applied to f_{xyz}^2 and f_{zyx}^3 . Hence, in view of step 4, we know that the claim is true if $t \in \{yxz, yzx, zxy, zyx\}$. Now, suppose contrapositive to the claim that $f_{xyz}^2(yzx, t) = x$ for some $t \in \{xyz, xzy\}$. Let $p = (yzx, zxy, xyz)$. Then p is manipulable towards $(p_{-2}, t) = (p_{-2}, t_p)$, as $f(p) = f_{xyz}^2(yzx, yxz) = y$ and $f(p_{-2}, t) = f_{xyz}^2(yzx, t) = x$. Let $q^1 = (yxz, zxy, xyz)$, $q^2 = (yzx, xzy, xyz)$ and $q^3 = (yxz, xzy, xyz)$. By step 3 $f(q^i) \in \{x, y\}$ for all $i \in \{1, 2, 3\}$. We show that for all possible combinations of $f(q^i) \in \{x, y\}, i \in \{1, 2, 3\}$, (b) is satisfied, contradicting our assumptions.

If $f(q^1) = x$ then q^1 is manipulable towards $(q_{-1}^1, yzx) = (q_{-1}^1, t_{q^1}) = (yzx, zxy, xyz)$, as $f(yzx, zxy, xyz) = y$. Clearly, p and q^1 are anonymously inequivalent and satisfy $|p(N)| = |q^1(N)| = 3$. As $q^1(1) = yxz \notin p(N) \cup \{t_p\} \subset \{yzx, zxy, xyz, xzy\}$, (b) is satisfied. Hence, we are done unless $f(q^1) = y$.

If $f(q^2) = y$ then q^2 is manipulable towards $(q_{-2}^2, t) = (q_{-2}^2, t_{q^2}) = (yzx, t, xyz)$, as $f(yzx, t, xyz) = x$. Clearly, p and q^2 are anonymously inequivalent and satisfy $|p(N)| = |q^2(N)| = 3$. As $p(2) = zxy \notin q^2(N) \cup \{t_{q^2}\} \subset \{yzx, xzy, xyz\}$, (b) is satisfied, a contradiction. Hence, we are done unless $f(q^2) = x$.

If $f(q^3) = y$ then q^2 is manipulable towards $(q_{-1}^2, yxz) = (q_{-1}^2, t_{q^2}) = q^3$, as $f(q^2) = x$. Still, p and q^2 are anonymously inequivalent, satisfy $|p(N)| = |q^2(N)| = 3$, and $p(2) = zxy \notin q^2(N) \cup \{t_{q^2}\} \subset \{yzx, xzy, xyz, yxz\}$, so (b) is satisfied. Hence, $f(q^3) = x$. But then q^1 is manipulable

towards $(q_{-2}^1, xzy) = (q_{-2}^1, t_{q^1}) = q^3$, as $f(q^1) = y$. Still, p and q^1 are anonymously inequivalent, satisfy $|p(N)| = |q^2(N)| = 3$, and $q^1(1) = yxz \notin p(N) \cup \{t_p\} \subset \{yzx, zxy, xyz, xzy\}$, so (b) is satisfied. This finishes the proof of step 4.

Step 6: In the third row y has to be chosen, i.e.

$$f_{xyz}(yxz, t) = y \text{ for all } t \in P.$$

We show step 5 for $k = 2$. The case $k = 3$ works in the same way. By step 3 we have that $f_{xyz}^2(yxz, t) \in \{x, y\}$ for all $t \in P$. Furthermore, $f_{xyz}^2(yxz, zyx) = f_{zyx}^3(yxz, xyz) \in \{x, y\} \cap \{y, z\} = \{y\}$ by step 3 applied to f_{xyz}^2 and f_{zyx}^3 . Hence, we know that the claim is true if $t \in \{yxz, yzx, zyx\}$. We extend this to $\{xzy, yxz, yzx, zxy, zyx\}$. So suppose contrapositive that $t \in \{xzy, zxy\}$ and $f_{xyz}^2(yxz, t) = x$. Then $p = (yxz, t, xyz)$ is manipulable by agent 1 towards $(yzx, t, xyz) = (t_p, t, xyz)$. As $t \in \{xzy, zxy\}$, $|p(N)| = 3$. Consider $\tilde{t} \in \{xzy, zxy\} - \{t\}$ and let $q = (yxz, \tilde{t}, xyz)$. If $f(q) = x$ then q is manipulable by agent 1 towards $(yzx, \tilde{t}, xyz) = (t_q, \tilde{t}, xyz)$. If $f(q) = y$ then q is manipulable by agent 2 towards $(yxz, t, xyz) = (yxz, t_q, xyz)$. In both cases $|q(N)| = 3$ as $\tilde{t} \in \{xzy, zxy\}$. Furthermore, $p(N) \cup \{t_p\} \neq q(N) \cup \{t_q\}$, as $\tilde{t} \in q(N)$ but $\tilde{t} \notin p(N) \cup \{t_p\} = \{yxz, yzx, t, xyz\}$. Hence, (b) is satisfied, a contradiction. So, $f_{xyz}(yxz, t) = y$ for all $t \in P - \{xyz\}$. It remains to be shown that $f_{xyz}(yxz, xyz) = y$. Suppose contrapositive that $f_{xyz}(yxz, xyz) = x$. Then $p = (yxz, xzy, xyz)$ and $q = (yxz, zxy, xyz)$, $f(p) = f(q) = y$, are both manipulable by agent 2 towards (yxz, xyz, xyz) . So (b) is satisfied, a contradiction. This proves step 6.

Step 7: f_{xyz}^2 can be described by the following 6x6 table.

2 :	xyz	xzy	yxz	yzx	zxy	zyx
1 : xyz	x	x	x	x	x	x
1 : xzy	x	x	x	x	x	x
1 : yxz	y	y	y	y	y	y
1 : yzx	y	y	y	y	y	y
1 : zxy		z		z	z	z
1 : zyx		z		z	z	z

The first four rows and the $\{zxy, zyx\}^{\{1,2\}}$ -cells follow by steps 1,5 and 6 applied to f_{xyz}^2 . By steps 5 and 6 applied to f_{yzx}^3 we see that $f_{xyz}^2(t, yzx) = f_{yzx}^3(t, xyz) = z$ for all $t \in \{zxy, zyx\}$. Likewise, $f_{xyz}^2(t, xzy) = z$ for all $t \in \{zxy, zyx\}$ by applying steps 5 and 6 to f_{xzy}^3 .

We prove now that f is dictatorial with dictator 1. This is equivalent to dictatoriality of all $f_t^2, t \in P$. To the contrary suppose that for some $t \in P$ f_t^2 is nondictatorial. Without loss of generality $t \in \{xyz, yxz\}$. Note that for both of these t the table of f_t^2 is similar. In view of step 7 this implies that there is a $r \in V := \{zxy, zyx\} \times \{xyz, yxz\}$, such that $u = f_t^2(r) \in \{x, y\}$. Without loss of generality $r(1) = zxy$, otherwise exchange the roles of x

and y which is allowed as we only fixed $t \in \{xyz, yxz\}$. Let $A = \{u, v, z\}$, $p = (r(1), uzv, t)$. Then p is manipulable towards $(p_{-2}, r(2)) = (p_{-2}, t_p) = (r, t)$, as $f_t^2(r) = u$ and $f(p) = z$. If there is an $\tilde{r} \in \{zyx\} \times \{xyz, yxz\}$ such that $\tilde{u} = f_t^2(\tilde{r}) \in \{x, y\}$, $A = \{\tilde{u}, z, \tilde{v}\}$ then, likewise, one obtains that $q = (\tilde{r}(1), \tilde{u}z\tilde{v}, t)$ is manipulable towards $(q_{-2}, \tilde{r}(2)) = (q_{-2}, t_q) = (\tilde{r}, t)$. Clearly, such p, q, t_p and t_q satisfy (b), a contradiction. Hence, we are done unless

$$f_t^2(zyx, \tilde{t}) = z \text{ for all } \tilde{t} \in P.$$

Let $q = (zxy, yxz, t)$. If $f(q) = z$ then q is manipulable towards $(q_{-2}, r(2)) = (q_{-2}, t_q) = (r, t)$, as $f(r, t) \in \{x, y\}$. Clearly, p and q are anonymously inequivalent and satisfy $|p(N)| = |q(N)| = 3$. As $p(2) = uzv \notin q(N) \cup \{t_q\} \subset \{zxy, yxz, xyz\}$ (b) is satisfied, a contradiction. Hence, we are done unless $f(q) \in \{x, y\}$. But then q is manipulable towards $(q_{-1}, zyx) = (q_{-1}, t_q)$, as $f(zyx, q) = f_t^2(zyx, q(2)) = z$. Still, p and q are anonymously inequivalent, satisfy $|p(N)| = |q(N)| = 3$, and $p(2) = uzv \notin q(N) \cup \{t_q\} \subset \{zxy, yxz, xyz, zyx\}$. Hence, (b) is satisfied, and we have obtained contradictions in all cases. So, there is no nondictatorial f_t and this finishes the proof. ■

4 The lower bound $2^n - 2$

In this section we will prove $\min_{f \in F} |M_f| \geq 2^n - 2$, using the results from Section 3.

Let $k : P^N \rightarrow \{\lceil \frac{n}{2} \rceil, \dots, n\}$ be given by

$$k(p) := \min\{|p_{-1}(T)| \mid T \subseteq P, |p^{-1}(T)| \geq \frac{n}{2}\}$$

and $l : P^N \rightarrow \{0, \dots, \lceil \frac{n}{2} \rceil\}$ by $l(p) := n - k(p)$. Then the following lemma holds for profiles that contain at least three different preferences.

Lemma 6 *Let p be a profile such that $|p(N)| \geq 3$. Then*

$$|[p]| \geq \sum_{j=0}^{l(p)-1} \binom{n}{k(p) + j}.$$

Proof. Let T be such that $|p^{-1}(T)| = k(p)$. Then

$$\begin{aligned} |[p]| &= \frac{n!}{\prod_{t \in P} |p^{-1}(t)|} \\ &= \frac{n!}{\prod_{t \in T} |p^{-1}(t)|! \prod_{t \in P-T} |p^{-1}(t)|!} \\ &\geq \frac{n!}{(k(p) - |T| + 1)! (l(p) - |P| + |T| + 1)!} \\ &= \frac{k(p)!}{(k(p) - |T| + 1)!} \frac{l(p)!}{(l(p) - |P| + |T| + 1)!} \binom{n}{k(p)} \\ &\geq l(p) \binom{n}{k(p)} \geq \sum_{j=0}^{l(p)-1} \binom{n}{k(p) + j}, \end{aligned}$$

The first inequality holds as

$$\prod_{t \in \tilde{T}} |p^{-1}(t)|! \leq (k(p) - |\tilde{T}| + 1)!$$

for all $\tilde{T} \subseteq P$ by applying Lemma 2 repeatedly. Furthermore, by $|p(N)| \geq 3$ either $|P - T| \geq 2$ and

$$\frac{l(p)!}{(l(p) - |P| + |T| + 1)!} \geq \frac{l(p)!}{(l(p) - 1)!} = l(p)$$

or $|T| \geq 2$ and

$$\frac{k(p)!}{(k(p) - |T| + 1)!} \geq \frac{k(p)!}{(k(p) - 1)!} = k(p) \geq l(p),$$

which justifies the second inequality. Finally the third inequality is valid as

$$\binom{n}{k(p)} \geq \binom{n}{k(p) + j}$$

for all $j \in \{0, \dots, l(p) - 1\}$, since $k(p) \geq \frac{n}{2}$. ■

Let $M_f^3 := \{p \in M_f \mid |p(N)| \geq 3\}$, and if $M_f^3 \neq \emptyset$ let

$$p_1 \in \arg \min_{p \in M_f^3} k(p), k_1 := k(p_1), l_1 := l(p_1),$$

and if $M_f^3 = \emptyset$ let $k_1 = k_2 = n$. Furthermore, if $k_1 < n$ and $M_f^3 - [p_1] \neq \emptyset$ let

$$p_2 \in \arg \min_{p \in M_f^3 - [p_1]} k(p), k_2 := k(p_2), l_2 := l(p_2),$$

and if $M_f^3 - [p_1] = \emptyset$ let $k_2 = n$.

The next lemma shows, using the results from Section 3, that for all $k \in \{\lceil \frac{n}{2} \rceil, \dots, k_1 - 1\}$ we have two profiles $q_1^k, q_2^k \in M_f$ such that $|q_1^k(N)| = |q_2^k(N)| = 2$ and $k(q_1^k) = k(q_2^k) = k$. If $k \neq \frac{n}{2}$ we show that these two profiles are anonymously inequivalent. For all $k \in \{k_1, \dots, k_2 - 1\}$ we show that there is one profile $q_1^k \in M_f$ such that $|q_1^k(N)| = 2$ and $k(q_1^k) = k$. Now note that profiles $p, q \in P^N$, such that $|p(N)| = |q(N)| = 2$ and $k(p) \neq k(q)$ are anonymously inequivalent. Altogether, this then shows that, if n is odd,

$$\begin{aligned} |M_f| &\geq \left| \bigcup_{k=\lceil \frac{n}{2} \rceil}^{k_2-1} [q_1^k] \cup \bigcup_{k=\lceil \frac{n}{2} \rceil}^{k_1-1} [q_2^k] \cup [p_1] \cup [p_2] \right| \\ &= \sum_{k=\lceil \frac{n}{2} \rceil}^{k_2-1} \binom{n}{k} + \sum_{k=\lceil \frac{n}{2} \rceil}^{k_1-1} \binom{n}{k} + |[p_1]| + |[p_2]| \\ &\geq \sum_{k=\lceil \frac{n}{2} \rceil}^{k_2-1} \binom{n}{k} + \sum_{k=\lceil \frac{n}{2} \rceil}^{k_1-1} \binom{n}{k} + \sum_{j=0}^{l_1-1} \binom{n}{k_1 + j} + \sum_{j=0}^{l_2-1} \binom{n}{k_2 + j} \\ &= \sum_{k=1}^{n-1} \binom{n}{k} = 2^n - 2, \end{aligned}$$

where we use Lemma 6 for the second inequality. Likewise one shows that $|M_f| \geq 2^n - 2$ if n is even, the only difference is that the two profiles $q_1^{\frac{n}{2}}, q_2^{\frac{n}{2}}$ do not have to be anonymously inequivalent.

Lemma 7 *Let $f : P^N \rightarrow A$ be a unanimous and anonymous social choice function. Let k_1, k_2, p_1, p_2 be as defined above. Then for all $k \in \{\lceil \frac{n}{2} \rceil, \dots, k_1 - 1\}$ there are two profiles $q_1^k, q_2^k \in M_f$ such that $|q_1^k(N)| = |q_2^k(N)| = 2$ and $k(q_1^k) = k(q_2^k) = k$. If $k \neq \frac{n}{2}$ these two profiles are anonymously inequivalent. For all $k \in \{k_1, \dots, k_2 - 1\}$ there is a profile $q_1^k \in M_f$ such that $|q_1^k(N)| = 2$ and $k(q_1^k) = k$.*

Proof. We will apply Lemma 4 to the following two agent social choice functions to show the existence of the q_i^k .

Let $k \in K := \{\lceil \frac{n}{2} \rceil, \dots, n\}$, and $l := n - k$. Define $f_k : P^{\{1,2\}} \rightarrow A$ by

$$f_k(p) := f(p(1)^k, p(2)^l)$$

for all $p \in P^{\{1,2\}}$, where $(p(1)^k, p(2)^l)$ denotes a profile q in P^N , such that $|q^{-1}(p(1))| = k$ and $|q^{-1}(p(2))| = l$. As f is anonymous f_k is well defined and as f is unanimous, f_k is unanimous.

To apply Lemma 4 to a $f_k, k \leq k_2$, we have to show that f_k is nondictatorial. Let \underline{d} be the smallest $k \in K$ such that f_k is dictatorial,

$$\underline{d} := \min\{k \in K \mid f_k \text{ is dictatorial}\}.$$

Note that f_n is dictatorial with dictator 1, so $\underline{d} \leq n$. We show that $k_2 \leq \underline{d}$, which implies that all $f_k, k \leq k_2$, are nondictatorial.

Claim 1:

$$k_2 \leq \underline{d}.$$

If $\underline{d} = n$ this is trivial. Suppose that $\underline{d} < n$. Let $\underline{l} := n - \underline{d}$. As f is anonymous $\underline{d} > \frac{n}{2} \geq \underline{l}$. Consider the following three agent social choice function. Let $g : P^{\{1,2,3\}} \rightarrow A$ be given by $g(p) := f(p(1)^{\underline{d}-\underline{l}}, p(2)^{\underline{l}}, p(3)^{\underline{l}})$, where $(p(1)^{\underline{d}-\underline{l}}, p(2)^{\underline{l}}, p(3)^{\underline{l}})$ can be understood similarly to $(p(1)^k, p(2)^l)$ above. As f is anonymous g is well defined and as f is unanimous, g is unanimous. Furthermore, as $f_{\underline{d}}$ is dictatorial, g is $\{1, 2\}$ and $\{1, 3\}$ -dictatorial. So we can apply Lemma 5 to g .

If part (a) of Lemma 5 holds, i.e. if agent 1 is a dictator, then $f_{\tilde{d}}$, where $\tilde{d} := \max\{2\underline{l}, \underline{d} - \underline{l}\} \in K$, is dictatorial. If $\tilde{d} < \underline{d}$ then, as $\tilde{d} \geq \frac{n}{2}$, this contradicts the minimality of \underline{d} . So, $\tilde{d} \geq \underline{d}$, which implies that $2\underline{l} \geq \underline{d}$. If $2\underline{l} = \underline{d}$ then g is anonymous, contradicting dictatoriality. So, $\underline{l} > \underline{d} - \underline{l}$. Consider the profiles

$$r^i := (abc^{\underline{d}-\underline{l}+i}, bca^{\underline{l}-i}, cab^{\underline{l}}), i \in \{0, \dots, 2\underline{l} - \underline{d}\}.$$

Then, $f(r^0) = a$, $f(r^{2\underline{d}}) = b$, as f is anonymous and g is dictatorial with dictator 1. But any r^{i+1} is a j -deviation of r^i for an agent j such that $(b, a) \in r^i(j)$. Hence, one of the r^i is manipulable. Clearly, $r^i \in M_f^3$ and $k(r^i) \leq \underline{d}$ for this r^i . For the same reasons a \bar{r}^j ,

$$\bar{r}^j := (acb^{\underline{d}-\underline{l}+j}, bca^{\underline{l}-j}, cab^{\underline{l}}), j \in \{0, \dots, 2\underline{d} - \underline{d}\},$$

is manipulable, in M_f^3 , and $k(\bar{r}^j) \leq \underline{d}$ for this \bar{r}^j . As this \bar{r}^j is anonymously inequivalent to r^i we have $k_2 \leq \max\{k(r^i), k(\bar{r}^j)\} \leq \underline{d}$.

So, suppose that part (b) of Lemma 5 is true. Then we have anonymously inequivalent $p, q \in M_g$, manipulable towards $(p_{-i}, t_p), (q_{-j}, t_q), i, j \in \{1, 2, 3\}$, respectively, such that $|p(N)| = |q(N)| = 3$ and $p(N) \cup \{t_p\} \neq q(N) \cup \{t_q\}$. If $i \in \{2, 3\}$ then by anonymity of f $i = 2$ without loss of generality. The same holds for j . Let

$$r^s := \begin{cases} (p(1)^{\underline{d}-\underline{l}-s}, t_p^s, p(2)^{\underline{l}}, p(3)^{\underline{l}}) & \text{if } i = 1, s \in \{0, \dots, \underline{d} - \underline{l} - 1\}, \\ (p(1)^{\underline{d}-\underline{l}}, p(2)^{\underline{l}-s}, t_p^s, p(3)^{\underline{l}}) & \text{if } i = 2, s \in \{0, \dots, \underline{l} - 1\}, \end{cases}$$

and

$$\bar{r}^u := \begin{cases} (q(1)^{\underline{d}-\underline{l}-s}, t_q^s, q(2)^{\underline{l}}, q(3)^{\underline{l}}) & \text{if } j = 1, u \in \{0, \dots, \underline{d} - \underline{l} - 1\}, \\ (q(1)^{\underline{d}-\underline{l}}, q(2)^{\underline{l}-s}, t_q^s, q(3)^{\underline{l}}) & \text{if } j = 2, u \in \{0, \dots, \underline{l} - 1\}. \end{cases}$$

By manipulability of p and q under g there are such r^s and \bar{r}^u that are manipulable under f . As $|p(N)| = |q(N)| = 3$, $r^s, \bar{r}^u \in M_f^3$. As p and q are anonymously inequivalent and $p(N) \cup \{t_p\} \neq q(N) \cup \{t_q\}$ we have $[r^s] \approx [\bar{r}^u]$. So, $k_2 \leq \max\{k(r^s), k(\bar{r}^u)\} \leq \underline{d}$. This finishes the proof of claim 1.

Let $k \in \{\lceil \frac{n}{2} \rceil, \dots, k_2 - 1\}$. By claim 1 f_k is nondictatorial. So, by Theorem 3 there are $p, q \in M_{f_k}$. Note that we do not demand $p \neq q$ here, because we want to allow for the moment for the possibility that p and q fall under case (a) of Lemma 4. However, if necessary we can always choose p and q to satisfy $p \neq q$ by the same theorem. By the first part of Lemma 4 there are $(i, t_p) \in D_{f_k}(p), (j, t_q) \in D_{f_k}(q)$ such that $t_p \notin p(N), t_q \notin q(N)$. Let

$$r^s := \begin{cases} (p(1)^{k-s}, p(2)^{\underline{l}}, t_p^s) & \text{if } i = 1, s \in \{0, \dots, k - 1\}, \\ (p(1)^k, p(2)^{\underline{l}-s}, t_p^s) & \text{if } i = 2, s \in \{0, \dots, \underline{l} - 1\}, \end{cases}$$

and

$$\bar{r}^u := \begin{cases} (q(1)^{k-u}, q(2)^{\underline{l}}, t_q^u) & \text{if } i = 1, u \in \{0, \dots, k - 1\}, \\ (q(1)^k, q(2)^{\underline{l}-u}, t_q^u) & \text{if } i = 2, u \in \{0, \dots, \underline{l} - 1\}. \end{cases}$$

By manipulability of p and q under f_k there are s and u such that $r^s, \bar{r}^u \in M_f$ that are manipulable towards $r^{\bar{s}+1}$ or $\bar{r}^{\bar{u}+1}$ respectively. Let \bar{s} be the smallest index s such that an r^s is manipulable, i.e. the manipulation does not have to occur towards r^{s+1} . Let \bar{u} be the smallest index u such that \bar{r}^u is manipulable towards \bar{r}^{u+1} . The reason for these choices will become apparent in case (c).

Suppose that

$$k \in \left\{ \left\lceil \frac{n}{2} \right\rceil, \dots, k_1 - 1 \right\}.$$

As $k(r^s) \leq k < k_1$, $k(\bar{r}^u) \leq k < k_1$, $|r^s(N)| \geq 3$ if $s \geq 1$ and $|\bar{r}^u(N)| \geq 3$ if $u \geq 1$, we must have $\bar{s} = 0$ and $\bar{u} = 0$. So $q_1^k := (p(1)^k, p(2)^l) = r^0 \in M_f$ and $q_2^k := (q(1)^k, q(2)^l) = \bar{r}^0 \in M_f$ satisfy $q_1^k, q_2^k \in M_f$, $|q_1^k(N)| = |q_2^k(N)| = 2$ and $k(q_1^k) = k(q_2^k) = k$. If $k = \frac{n}{2}$ we are done. If $k \neq \frac{n}{2}$ then $k > l$ and choosing $p \neq q$ implies that q_1^k and q_2^k are anonymously inequivalent. So, for $k \in \left\{ \left\lceil \frac{n}{2} \right\rceil, \dots, k_1 - 1 \right\}$ we are done.

Suppose that

$$k \in \{k_1, \dots, k_2 - 1\}.$$

We are done if $\bar{s} = 0$ or $\bar{u} = 0$. So, suppose contrapositive that $\bar{s} > 0$ and $\bar{u} > 0$. Using cases (a) to (c) of Lemma 4 we show that p and q can be chosen in such a way that $r^{\bar{s}} \approx \bar{r}^{\bar{u}}$. As $k(r^{\bar{s}}) \leq k$, $k(\bar{r}^{\bar{u}}) \leq k$, $|r^{\bar{s}}(N)| \geq 3$ and $|\bar{r}^{\bar{u}}(N)| \geq 3$ if $u \geq 1$, this contradicts $k < k_2$.

Case (a): We can choose p and q in M_{f_k} that are manipulable by different agents, say without loss of generality p by agent 1 and q by agent 2. Note that here we have to allow for $p = q$. Then

$$\max_{t \in P} |(r^{\bar{s}})^{-1}(t)| = \max\{k - \bar{s}, l, \bar{s}\} \leq k - 1 < k = (\bar{r}^{\bar{u}})^{-1}(q(1)),$$

so $r^{\bar{s}} \approx \bar{r}^{\bar{u}}$.

Case (b): We can choose p and q in M_{f_k} such that $p(N) \cup \{t_p\} \neq q(N) \cup \{t_q\}$. Then $r^{\bar{s}} \approx \bar{r}^{\bar{u}}$ is obvious.

Case (c): Either $i = j = 1$ and

$$\begin{aligned} p &= (bca, cab), (1, cba) \in D_{f_k}(p) \\ q &= (cba, bca), (1, cab) \in D_{f_k}(q) \\ f_k(p) &= a \end{aligned}$$

or $i = j = 2$ and

$$\begin{aligned} p &= (cab, bca), (2, cba) \in D_{f_k}(p) \\ q &= (bca, cba), (2, cab) \in D_{f_k}(q) \\ f_k(p) &= a. \end{aligned}$$

Suppose that $i = j = 1$ and $r^{\bar{s}} \sim \bar{r}^{\bar{u}}$. Then

$$r^{\bar{s}} = (bca^{k-\bar{s}}, cba^{\bar{s}}, cab^l) \sim (cba^{k-\bar{u}}, cab^{\bar{u}}, bca^l) = \bar{r}^{\bar{u}}.$$

This implies that $k - \bar{s} = l$, $\bar{s} = k - \bar{u}$ and $l = \bar{u}$. Note that $f(r^0) = f_k(p) = a$ is the least preferred alternative of the manipulating agent, who has the preference bca . By the minimality of \bar{s} and the anonymity of f this implies that $f(r^0) = f(r^1) = \dots f(r^{\bar{s}-1}) = f(r^{\bar{s}}) = f(\bar{r}^{\bar{u}}) = a$.

Furthermore, as $(cba^{k-\bar{u}}, cab^{\bar{u}}, bca^l)$ is manipulable by an agent with preference cba , we have $f(cba^{k-\bar{u}-1}, cab^{\bar{u}+1}, bca^l) \in \{b, c\}$. Hence,

$$f(cba^{k-\bar{u}-1}, cab^{\bar{u}+1}, bca^l) = f(cba^{\bar{s}-1}, cab^{l+1}, bca^{k-\bar{s}}) \in \{b, c\}.$$

But then $r^{\bar{s}-1}$ is manipulable, as $f(r^{\bar{s}-1}) = f(bca^{k-\bar{s}+1}, cba^{\bar{s}-1}, cab^l) = a$, and by deviating to cab any agent with preference bca can manipulate. This contradicts the minimality of \bar{s} .

Suppose that $i = j = 2$ and $r^{\bar{s}} \sim \bar{r}^{\bar{u}}$. Then

$$r^{\bar{s}} = (bca^k, cab^{l-\bar{s}}, cba^{\bar{s}}) \sim (cba^k, bca^{l-\bar{u}}, cab^{\bar{u}}) = \bar{r}^{\bar{u}}.$$

This implies by $k > l$ that $bca = cba$, a contradiction.

Hence, in any of the cases (a) to (c) p and q can be chosen in such a way that $r^{\bar{s}} \approx \bar{r}^{\bar{u}}$. This contradicts then $k < k_2$ if $\bar{s} > 0$ and $\bar{u} > 0$. So, without loss of generality $\bar{u} = 0$, and letting $q_1^k = r^0 \in M_f$ we have $|q_1^k(N)| = 2$ and $k(q_1^k) = k$.

This finishes the proof of the lemma. ■

By the remarks before Lemma 7 we can now state the following theorem.

Theorem 8 *Let $f : P^N \rightarrow A$ be a unanimous and anonymous social choice function. Then $|M_f| \geq 2^n - 2$.*

5 Social choice functions f^* that attain the lower bound

$$2^n - 2$$

Let $F^* := \{f \in F \mid |M_f| = 2^n - 2\}$ be the set of all social choice functions attaining the lower bound. In this section we formulate a conjecture about F^* , saying that $f^* \in F^*, n \geq 4$, if and only if f^* belongs to one of three classes of social choice functions which we describe completely. There will be an exception for $n \in \{2, 3\}$. First of all one class is only defined for $n \geq 3$, where its definition in the case $n = 3$ compared to $n \geq 4$ alters a little. Furthermore, for $n = 3$ a fourth extra class of social choice functions is needed. Altogether we define four classes $C_i, i \in \{1, 2, 3, 4\}$, of social choice functions such that $C_i \subseteq F^*$. For any $x \in A$ and $p \in P^N$ let

$$Pa(x, p) := \{y \in A - \{x\} \mid (y, x) \in p(i) \text{ for all } i \in N\}$$

be the set of all alternatives that *Pareto dominate* x in p . In 1. up to 5. it is to be understood that indefiniteness is resolved in such a way that anonymity and unanimity hold. Let $A = \{x, y, z\}$.

1. Let $C_1 \subseteq F$ be the set of all social choice functions f such that

$$f(p) = \begin{cases} z & \text{if } Pa(z, p) = \emptyset \\ x & \text{if } Pa(z, p) = \{x\} \\ y & \text{if } Pa(z, p) = \{y\} \end{cases}$$

and $f(p) = f(q) \in \{x, y, z\}$ for all $q \in P^N$ such that

$$Pa(z, p) = Pa(z, q) = \{x, y\}, [q] = [p].$$

2. Let $C_2 \subseteq F$ be the set of all social choice functions f such that

$$f(p) = \begin{cases} z & \text{if } Pa(z, p) = \emptyset \\ x & \text{if } Pa(z, p) = \{x\} \\ y & \text{if } Pa(z, p) = \{y\} \text{ and } p \notin [(xyz^{n-1}, yzx)] \\ z & \text{if } p \in [(xyz^{n-1}, yzx)] \\ x & \text{if } p \in [(xyz^{n-1}, yxz)] \end{cases}$$

and $f(p) = f(q) \in \{x, y, z\}$ for all $p, q \in P^N$ such that $Pa(z, p) = Pa(z, q) = \{x, y\}, [p] = [q] \neq [(xyz^{n-1}, yzx)]$.

3. If $n \geq 4$ let $C_3 \subseteq F$ be the set of all social choice functions f such that

$$f(p) = \begin{cases} z & \text{if } Pa(z, p) = \emptyset \\ x & \text{if } Pa(z, p) = \{x\} \text{ and } p \notin [(yxz^{n-1}, xzy)] \\ y & \text{if } Pa(z, p) = \{y\} \text{ and } p \notin [(xyz^{n-1}, yzx)] \\ z & \text{if } p \in [(yxz^{n-1}, xzy)] \\ z & \text{if } p \in [(xyz^{n-1}, yzx)] \\ x & \text{if } p \in [(xyz^{n-1}, yxz)] \cup [(xyz^{n-2}, yxz^2)] \\ y & \text{if } p \in [(yxz^{n-1}, xyz)] \cup [(yxz^{n-2}, xyz^2)] \end{cases}$$

and $f([p]) \in \{x, y, z\}$ if $Pa(z, p) = \{x, y\}$ and $p \notin [(xyz^{n-1}, yxz)] \cup [(xyz^{n-2}, yxz^2)] \cup [(yxz^{n-1}, xyz)] \cup [(yxz^{n-2}, xyz^2)]$.

4. If $n = 3$ let $C_3 \subseteq F$ be the set of all social choice functions f such that

$$f(p) = \begin{cases} z & \text{if } Pa(z, p) = \emptyset \\ x & \text{if } Pa(z, p) = \{x\} \text{ and } p \notin [(yxz^2, xzy)] \\ y & \text{if } Pa(z, p) = \{y\} \text{ and } p \notin [(xyz^2, yzx)] \\ z & \text{if } p \in [(yxz^2, xzy)] \\ z & \text{if } p \in [(xyz^2, yzx)] \\ x & \text{if } p \in [(xyz^2, yxz)] \\ y & \text{if } p \in [(yxz^2, xyz)] \end{cases}$$

and $f(p) \in \{x, y, z\}$ if $Pa(z, p) = \{x, y\}$ and $p \notin [(xyz^2, yxz)] \cup [(yxz^2, xyz)]$.

5. If $n = 3$ let $C_4 \subseteq F$ be the set of all social choice functions f for which there are $x, y \in A, x \neq y$, such that $f(p)$ is determined by majority voting between x and y , i.e.

$$f(p) = \begin{cases} x & \text{if } |\{i \in N \mid (x, y) \in p(i)\}| > |\{i \in N \mid (y, x) \in p(i)\}|, \\ y & \text{if } |\{i \in N \mid (y, x) \in p(i)\}| > |\{i \in N \mid (x, y) \in p(i)\}|, \end{cases}$$

for all profiles p where $f(p)$ is not determined already by unanimity.

The following lemma can be proven by counting manipulable profiles and evaluating the freedom left in the definitions of the C_i . The counting argument can be simplified by noting that no agent will want to manipulate to a Pareto dominated alternative. Since the rest is elementary counting we do not present the proof here.

Lemma 9 *Let F^* be the set of all minimally manipulable unanimous and anonymous social choice function with n agents and three alternatives. Then, $C_1 \cup C_2 \subseteq F^*$ if $n = 2$, $C_1 \cup C_2 \cup C_3 \cup C_4 \subseteq F^*$ if $n = 3$ and $C_1 \cup C_2 \cup C_3 \subseteq F^*$ if $n \geq 4$. Furthermore, the C_i that are defined in a particular case are disjoint. We have $|C_1| = 3^n$, $|C_2| = 2 * 3^{n-1}$, and $|C_3| = 3^{n-2}$, $|C_4| = 3$, if they are defined, and*

$$|F^*| \geq \begin{cases} 15 & \text{if } n = 2, \\ 51 & \text{if } n = 3, \\ 3^n + 2 * 3^{n-1} + 3^{n-2} & \text{if } n \geq 4. \end{cases}$$

We conjecture that equalities hold in the lemma. We believe that the proof of the main result in Section 4 can be used to show this conjecture and explain how this might be done.

First of all we note that in the statement of Lemma 6

$$[p_i] \geq \sum_{j=0}^{l(p_i)-1} \binom{n}{k(p_i) + j}, p_i \in M_f^3,$$

equality cannot hold. Going through the proof of that statement one can see that equality would imply that $l(p_i) = 1$, $k(p_i) = n - 1$, and $|p_i(N)| = 2$, a contradiction. So, $M_{f^*}^3 = \emptyset$ for any $f^* \in F^*$. This also implies that $\underline{d} = n$. Hence, all manipulable profiles arise from the f_k . In particular for each f_k we must have that it yields only two anonymously inequivalent manipulable profiles. Now, classifying all such f_k , we know the possible M_f 's and what f can look like on $\{p \in P^N \mid |p(N)| \leq 2\}$. Connecting this information by strategy-proofness, in particular on $B^3 := \{p \in P^N \mid |p(N)| \geq 3\}$, it might be possible to characterize F^* . The reason why we do not pursue this here is the classification of all such f_k . There are 125 unanimous and nondictatorial two agent social choice functions with two manipulable profiles which have to be checked first, and there might be even more with more manipulable profiles, but only two that give anonymously inequivalent profiles for f .

We mention also that we have confirmed the following conjecture up to 8 agents by computer, which is also how we found the 125 unanimous and nondictatorial two agent social choice functions over three alternatives.

Conjecture 10 Let F^* be the set of all minimally manipulable unanimous and anonymous social choice functions for a fixed size of the agent set n . Then

$$F^* = \begin{cases} C_1 \cup C_2 & \text{if } n = 2, \\ C_1 \cup C_2 \cup C_3 \cup C_4 & \text{if } n = 3, \\ C_1 \cup C_2 \cup C_3 & \text{if } n \geq 4, \end{cases}$$

where the C_i are defined as in Lemma 9. Furthermore, as the C_i are disjoint,

$$|F^*| = \begin{cases} 9 + 6 = 15 & \text{if } n = 2, \\ 27 + 18 + 3 + 3 = 51 & \text{if } n = 3, \\ 3^n + 2 * 3^{n-1} + 3^{n-2} & \text{if } n \geq 4. \end{cases}$$

6 Pareto optimality and anonymity

In this section we show how our results and the conjecture apply when unanimity is replaced by the stronger requirement of Pareto optimality.

A social choice function is called *Pareto optimal* if it does not choose Pareto dominated alternatives, i.e. if $Pa(x, p) \neq \emptyset$, $x \in A$, implies that $f(p) \neq x$. Pareto optimality is a stronger requirement than unanimity. Let p be such that $f(p) = x$ is determined by unanimity. Then $Pa(y, p) \supseteq \{x\}$ for all $y \in A - \{x\}$, implying that $f(p) \neq y$ for all $y \in A - \{x\}$. So $f(p) = x$ also by Pareto optimality. On the other hand $f(abc, bca) \in \{a, b\}$ by Pareto optimality, but unanimity does not reduce the choice at this profile. Let G be the set of Pareto optimal and anonymous social choice functions. Then $G \subset F$.

Let $\overline{C_1} \subseteq G$ be the set of all social choice functions $g \in G$ for which there is a labeling of the alternatives $A = \{x, y, z\}$ such that

$$g(p) = \begin{cases} z & \text{if } Pa(z, p) = \emptyset \\ x & \text{if } Pa(z, p) = \{x\} \\ y & \text{if } Pa(z, p) = \{y\} \end{cases}$$

and $g(p) = g(q) \in \{x, y\}$ for all $q \in P^N$ such that

$$Pa(z, p) = Pa(z, q) = \{x, y\}, [q] = [p].$$

Note that $\overline{C_1} \subset G \subset F$ and $\overline{C_1} \subset C_1 \subset F^*$. So the following lemma is implied by Lemma 9 and counting $\overline{C_1}$.

Lemma 11 Let G^* be the set of minimally manipulable social choice functions in G . Then

$$\overline{C_1} \subseteq G^*, |\overline{C_1}| = 3 * 2^n \text{ and } |G^*| \geq 3 * 2^n.$$

Furthermore, $G \subseteq F$ implies then the following corollary to Theorem 8.

Corollary 12 *Let $g : P^N \rightarrow A$ be a Pareto optimal and anonymous social choice function. Then $|M_g| \geq 2^n - 2$.*

It is easily checked that there are no Pareto optimal social choice functions in $C_i, i \in \{2, 3, 4\}$, whenever these sets are defined. So we can state the following conjecture based on conjecture 10.

Conjecture 13 *Let G^* be the set of all minimally manipulable Pareto optimal and anonymous social choice functions for a fixed size of the agent set n . Then*

$$G^* = \overline{C}_1,$$

and

$$|G^*| = 3 * 2^n.$$

Note that for any $g \in \overline{C}_1$ there is an alternative z that is chosen whenever it is not Pareto dominated by one of the other alternatives. We call this alternative the *status quo*. If z is Pareto dominated, then the choice is taken among one of the dominating alternatives in an arbitrary way. In practice this arbitrariness might be replaced by a voting rule between x and y that seems reasonable in the given situation. For example it might be that y is declared as a second status quo, so that x is only chosen if $p = x \dots^n$.

Anyway these explanations show that our result has a nice interpretation if a status quo z among the alternatives exists, e.g. a current jurisdiction z about which there is consensus that it should only be altered to one of two new jurisdictions x and y if every voter prefers the new jurisdiction. Then, the least manipulable social choice functions that guarantee anonymity and Pareto optimality of the outcome are as follows. Choose z in the situation where there is consensus that no change should be made. In the other situations choose an alternative Pareto dominating z ; if necessary, i.e. if there are two such alternatives, according to any unanimous social choice function over two alternatives.

References

- [1] Aleskerov, F, Kurbanov, E (1999) Degree of manipulability of social choice procedures. In: Proceedings of the Third International Meeting of the Society for the Advancement of Economic Theory. Berlin/Heidelberg/New York: Springer.
- [2] Fristup, P, Keiding, H (1998) Minimal manipulability and interjacency for two-person social choice functions. Social Choice and Welfare 15: 455-467.
- [3] Gibbard A (1973) Manipulation of voting schemes: a general result. Econometrica 41: 587-602.

- [4] Kelly JS (1988) Minimal Manipulability and Local Strategy-Proofness. *Social Choice and Welfare* 5: 81-85.
- [5] Maus, S, Peters, H, Storcken, A (2004a) Minimal manipulability: Anonymity and Surjectivity. Research Memoranda Universiteit Maastricht RM /04/007, University of Maastricht.
- [6] Maus, S, Peters, H, Storcken, A (2004b) Minimal manipulability: Unanimity and Nondictatorship. Research Memoranda Universiteit Maastricht RM /04/006, University of Maastricht.
- [7] May KO (1952) A Set of Independent, Necessary and Sufficient Conditions for Simple Majority voting. *Econometrica* 20: 680-684.
- [8] Satterthwaite M (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.
- [9] Slinko A (2002) On asymptotic strategy-proofness of classical social choice rules. *Theory and Decision* 52: 389-398.