# On Asymptotic Coalitional Strategy-Proofness of Social Choice Rules under the IAC Assumption

**Arkadii Slinko***

Department of Mathematics
Auckland University
New Zealand

**Abstract:** We consider scoring rules and multistage elimination rules under the Impartial Anonymous Culture assumption. We show that, when the number of participating agents $n$ tends to infinity, the proportion of voting situations manipulable by a coalition of $k$ voters to all voting situations is smaller than $D_m \frac{k}{n}$, where $D_m$ depends only on the number of alternatives $m$ but not on $k$ and $n$.

**Key words:** Social Choice Rule, Impartial Anonymous Culture, Manipulability, Asymptotic Strategy-Proofness.

**JEL classification:** D7

*Department of Mathematics, The University of Auckland, Private Bag 92019, Auckland New Zealand. Fax: 64-9-3737457, email: a.slinko@auckland.ac.nz

# 1 Definitions and Basic Concepts

Let $A$ and $\mathcal{N}$ be two finite sets of cardinality $m$ and $n$ respectively. The elements of $A$ will be called alternatives, the elements of $\mathcal{N}$ agents. We assume that the agents have preferences over the set of alternatives. By $\mathcal{L} = \mathcal{L}(A)$ we denote the set of all linear orders on $A$; they represent the preferences of agents over $A$. The cardinality of this set is $M = m!$ and we list these linear orders in some way $R_1, \ldots, R_M$ and fix this order.

The well-known Impartial Culture (IC) assumption stipulates that all voters are independent and that they can choose any linear order on $A$ with equal probability $1/M$. Unlike the IC, the Impartial Anonymous Culture (IAC) assumption does not operate in terms of individual voters but rather in terms of voting situations that may occur.

Suppose that $n_i$ voters chose the linear order $R_i$ as their preference over the set of alternatives $A$. Then we say that a voting situation $S = (n_1, n_2, \ldots, n_M)$ occurred. In other words, a voting situation is a multiset on $\mathcal{L}$ of cardinality $n$. The IAC assumes that all voting situations are equiprobable. Since the total number of voting situations is $\binom{n+M-1}{n}$, the probability of each particular voting situation under the IAC is $\binom{n+M-1}{n}^{-1}$.

Hence, unlike the IC, the IAC is a model for the society as a whole and not a model for the behavior of individual voters. For a comprehensive survey of various probability models used in Social Choice Studies see Berg and Lepelley [2].

Let $\mathcal{S}_n(A)$ be the set of all voting situations for $n$ voters and the set $A$ of alternatives. Under the IAC, any family of mappings $F = \{F_n\}$, $n \in \mathbb{N}$,

$$F_n \colon \mathcal{S}_n(A) \to A,$$

is called a social choice function (SCF). For historical reasons, SCFs are often called rules.

**Definition 1.** *Let $S$ be a voting situation. We say that a voting situation $S'$ occurred as a result of change of mind of $k$ voters, if some $k$ voters who previously submitted linear orders $R_{i_1}, \ldots, R_{i_k}$ now submit linear orders $R_{j_1}, \ldots, R_{j_k}$ while the remaining voters submit their original linear orders.*

**Definition 2.** *Let $F$ be an SCF and let $S$ be a voting situation. We say that $S$ is $k$-manipulable for $F$ if there is a voting situation $S'$, for which occurred as a result of change of mind of $k$ voters, with the linear orders $R_{i_1}, \ldots, R_{i_k}$*

*being replaced by the linear orders $R_{j_1}, \ldots, R_{j_k}$, such that $F(S')R_{i_s}F(S)$ for all $s = 1, 2, \ldots, k$. We also say a voting situation $S$ is $k$-unstable if there exists a voting situation $S^{'}$, which occurred as a result of change of mind of $k$ voters, such that $F(S^{'}) \neq F(S)$.*

Every $k$-manipulable voting situation is $k$-unstable, but the reverse is not always true. For $k = 1$ we get individual manipulability and individual stability.

The concept of an unstable profile was introduced for the IC assumption by Pazner and Wesley [9] and Peleg [10] and the concept of an unstable voting situation (individual and coalitional) was discussed in Lepelley and Mbih [8]. However, the concept of group manipulability discussed in [4, 7, 8] does not restrict the size of the manipulating coalition.

For our study we will use the following two indices of group manupulability. Given the rule $F$, the index of $k$-manipulability of $F$ under the IAC is

$$K_F(n, m, k) = \frac{d_F(n, m, k)}{\binom{n+M-1}{n}}, \tag{1}$$

where $d_F(n, m, k)$ is the total number of all $k$-manipulable voting situations, and the index of instability

$$L_F(n, m, k) = \frac{e_F(n, m, k)}{\binom{n+M-1}{n}}, \tag{2}$$

where $e_F(n, m, k)$ is the total number of all $k$-unstable voting situations. We note that under the IAC the set of all voting situations $\mathcal{S}_n(A)$ is assumed to be a discrete probability space with the uniform distribution, hence the indices $K_F(m, n, k)$ and $L_F(m, n, k)$ become the probabilities of drawing at random a $k$-manipulable voting situation, or a $k$-unstable voting situation, respectively. Since $K_F(m, n, k) \leq L_F(m, n, k)$, any upper bound that we can obtain for $L_F(n, m, k)$ will be an upper bound for $K_F(m, n, k)$.

In this paper we prove that if $F$ is any scoring rule or multistage elimination rule, then

$$L_F(n, m, k) \leq D_m \frac{k}{n}, \tag{3}$$

where $D_m$ is the constant that depends only on $m$ but not on $k$ and $n$. Therefore, if $k = o(n)$,[1] then $L_F(n, m, k) \to 0$ as $n \to \infty$. This means that any such $F$ is asymptotically cannot be manipulated by coalitions of size $k$.

[1] The notation $g(n) = o(f(n))$ means that $g(n)/f(n) \to 0$, when $n \to \infty$.

It is interesting to compare this result with the existing results for the IC. The corresponding indices are defined as follows. The index of $k$-manipulability of $F$ will be

$$K_F(n,m,k) = \frac{d_F(n,m,k)}{(m!)^n},$$

where $d_F(n,m,k)$ is the total number of all $k$-manipulable profiles, and the index of instability of $F$ will be

$$L_F(n,m,k) = \frac{e_F(n,m,k)}{(m!)^n},$$

where $e_F(n,m,k)$ is the total number of all $k$-unstable profiles. Peleg [10] proved that if $k = o(\sqrt{n})$, then $L_F(n,m,k) \to 0$. It also follows from Theorem 2 of the author's paper [12] that for any faithful scoring rule $F$ and any multistage elimination rule based on the scores of $F$

$$L_F(n,m,k) \le C_m \frac{k}{\sqrt{n}},$$

where $C_m$ is a constant which depends only on $m$ but not on $k$ and $n$.

## 2 The Main Combinatorial Result

**Definition 3.** *Let $n, \ell$ be positive integers. Any $\ell$-tuple $(n_1, n_2, \ldots, n_\ell)$ of nonnegative integers such that*

$$n_1 + n_2 + \ldots + n_\ell = n \tag{4}$$

*will be called an $\ell$-composition of $n$.*

Note that this definition slightly differs from the classical definition of a composition (see, for example, Andrews [1]) since the summands in our definition may be zero.

In this section $n$ and $\ell$ will be fixed.

**Definition 4.** *Let $k_1, \ldots, k_\ell$, and $p$ be integers. We say that an $\ell$-composition $(n_1, n_2, \ldots, n_\ell)$ of $n$ satisfies the equation*

$$\sum_{i=1}^{\ell} k_i x_i = p \tag{5}$$

*iff* $\sum_{i=1}^{\ell} k_i n_i = p$. *This equation will be said to be nontrivial if it does not follow from*

$$\sum_{i=1}^{\ell} x_i = n \tag{6}$$

*and* $k_i \neq 0$ *for some* $i$.

**Theorem 1.** *Let* $\ell, n$ *be fixed positive integers and let* $\mathcal{S}(\ell, n)$ *be the set of all possible* $\ell$*-compositions of* $n$*. Suppose that the set* $\mathcal{S}(\ell, n)$ *is given the structure of a discrete probability space with the uniform distribution. Then the probability* $P(\ell, n)$ *of choosing an* $\ell$*-composition satisfying a given nontrivial equation is less than or equal to* $C_\ell/n$*, where* $C_\ell$ *is the constant which depends only on* $\ell$ *but not on* $n$*.*

**Proof:** Let us fix a notrivial equation (5). The statement will be proved by induction on $\ell$. For $\ell = 2$ the equation (5) becomes

$$ax_1 + bx_2 = c, \tag{7}$$

where either $a$ or $b$ (or both) are nonzero integers. Since (7) does not follow from $x_1 + x_2 = n$, the determinant $\begin{vmatrix} a & b \\ 1 & 1 \end{vmatrix}$ must be nonzero. Hence we get at most one pair $(n_1, n_2)$, satisfying (7). The total number of 2-compositions is $n + 1$, hence the probability that a random pair satisfies (7) is less than $1/n$, as required. This gives a basis for the induction.

Suppose now that $P(k, n) \leq C_k/n$ for all $k < \ell$. Let us note also that we may assume that in the equation (5) at least two coefficients among the $k_1, \ldots, k_\ell$ are nonzero and at least two coefficients are different from 1. If we had just one nonzero coefficient, then the statement of the theorem would be true. Indeed, in this case one of the $n_i$'s will be fixed and the total number of $\ell$-compositions with the fixed $n_i$ can be estimated as follows:

$$\binom{n - n_i + \ell - 2}{n - n_i} \leq \binom{n + \ell - 2}{n} \leq \frac{\ell}{n}\binom{n + \ell - 1}{n}.$$

Since the total number of $\ell$-compositions is $\binom{n+\ell-1}{n}$, the probability of this event is not greater than $\ell/n$. If only one coefficient is different from 1, this has the same effect, since subtracting the equation $x_1 + x_2 + \ldots + x_\ell = n$ we would get an equation, where only one coefficient in the left-hand-side is nonzero and the previous case would apply. Hence we may assume that for any $\ell$-composition $S = (n_1, n_2, \ldots, n_\ell)$ of $n$ satisfying (5) and for any $i = 1, 2, \ldots, \ell$, the $(\ell-1)$-composition

$$S' = (n_1, \ldots, n_{i-1}, n_{i+1}, \ldots, n_\ell)$$

of $n - n_i$ satisfies the nontrivial equation

$$k_1 x_1 + \ldots + k_{i-1} x_{i-1} + k_{i+1} x_{i+1} + \ldots + k_\ell x_\ell = p - k_i n_i. \tag{8}$$

Hence we may assume that any $\ell$-composition $S = (n_1, n_2, \ldots, n_\ell)$ of $n$ satisfying a nontrivial equation (5), will still satisfy a nontrivial equation if we remove any of the $n_i$'s from it.

Let us now estimate $P(\ell, n)$. This estimation will be based on the observation that $n_i \leq n/\ell$ for at least one $i \in \{1, \ldots, \ell\}$. Because of that

$$P(\ell, n) \leq \sum_{i=1}^{\ell} \sum_{k=0}^{n/\ell} Prob\{n_i = k\} P(\ell - 1, n - k).$$

Since, for $k \leq n/\ell$, by the induction hypothesis we obtain

$$P(\ell - 1, n - k) \leq \frac{C_{\ell-1}}{n-k} \leq \frac{C_{\ell-1}}{n - n/\ell} = \frac{\ell}{\ell - 1} \frac{C_{\ell-1}}{n},$$

we get

$$P(\ell, n) = \sum_{i=1}^{\ell} \sum_{k=0}^{n/\ell} Prob\{n_i = k\} P(\ell - 1, n - k) \leq \frac{\ell^2 C_{\ell-1}}{\ell - 1} \cdot \frac{1}{n} = \frac{C_\ell}{n}.$$

This proves the theorem.

# 3 Asymptotic Strategy-Proofness

One of the most important classes of SCFs was introduced by Gärdenfors [5], they are called *representable voting functions.*

**Definition 5.** *A representation function is a function* $f\colon \mathcal{L}(A) \times A \to \mathbb{R}$ *such that*

$$aR_i b \Longrightarrow f(R_i, a) \geq f(R_i, b).$$

*It is called faithful if*

$$(aR_i b \text{ and } a \neq b) \Longrightarrow f(R_i, a) > f(R_i, b).$$

**Definition 6.** *A representation function* $f\colon \mathcal{L}(A)\times A \to \mathbb{R}$ *is called positionalist, if* $f(R,a)$ *depends only on the cardinality of the lower contour set* $L(R,a) = \{b \in A \mid aRb \text{ and } a \neq b\}$.

The simplest positionalist representation function (which is used to define the Borda rule) can be defined by setting $f(R,a) = card\,(L(R,a))$. It is easy to see that it is faithful.

Let $f$ be a representation function and $S = (n_1, n_2, \ldots, n_M) \in \mathcal{S}_n(A)$ be a voting situation. We define the *score function* $Sc_f\colon \mathcal{S}_n(A)\times A \to \mathbb{R}$ by

$$Sc_f(S,a) = \sum_{i=1}^{M} n_i f(R_i, a), \quad a \in A. \tag{9}$$

**Definition 7.** *An SCF $F$ is called (faithfully) representable if there exists a (faithful) representation function $f$ such that for every $S \in \mathcal{S}_n(A)$ we have $F(S) = a_i$ if and only if*

$$j < i \implies Sc_f(S,a_i) > Sc_f(S,a_j), \tag{10}$$

$$j > i \implies Sc_f(S,a_i) \geq Sc_f(S,a_j). \tag{11}$$

*When $f$ is positionalist, any SCF is also called a (faithful) scoring rule (or point-voting scheme).*

Every scoring rule $F$ with a representation function $f$ is characterised by the respective vector of weights $W_f = (w_1, \ldots, w_m)$. The weights are chosen so that for any $Q \in \mathcal{L}(A)$ the equation $w_i = f(Q,a)$ holds if and only if $card(L(Q,a)) = m - i$. The weights must satisfy the condition

$$w_1 \geq w_2 \geq \ldots \geq w_m = 0,$$

and we can consider them to be integers. It is clear that the scoring rule $F$ is faithful iff $w_i \neq w_{i+1}$ for all $i = 1, 2, \ldots, m-1$.

For each voting situation $S \in \mathcal{S}_n(A)$ the value of the scoring function $f$ on $a$, which we will simply call the score of $a$, can be now computed as follows. Let $I_a = (i_1, \ldots, i_m)$ be the vector such that the number $i_k$ shows how many times the alternative $a$ was ranked $k$th. Then

$$Sc_f(S,a) = W_f \cdot I_a = \sum_{\ell=1}^{m} w_\ell i_\ell.$$

The most commonly used scores are the Plurality score $Sc_P(S, a)$, when $P$ is the Plurality rule, defined by the vector of weights $W_P = (1, 0, \ldots, 0)$ and the Borda score $Sc_B(S, a)$, where $B$ is the Borda rule, defined by the vector of weights $W_B = (m-1, m-2, \ldots, 1, 0)$. The Approval Voting score with a fixed number $k$ of approvals is given by the vector of weights $W_A = (1, 1, \ldots, 1, 0, \ldots, 0)$ ($k$ ones). When $k = m - 1$ it bears the name of the Antiplurality score.

Another related class of social choice rules are multistage elimination rules for which the winner is determined in several stages. At every stage one (or more) of the alternatives is eliminated on the basis of a certain "global" information, i.e. the eliminated alternative must be the worst one relative to a certain global characteristic. This characteristic is normally related to the scores of alternatives defined in the previous section.

Among these the best known is the Run-off Procedure for which at the first stage $m - 2$ alternatives with the lowest Plurality scores are eliminated. Hare's Rule (or Single Transferable Vote) stipulates that at every stage only one alternative with the minimal Plurality score is eliminates. The Inverse Borda rule acts exactly as Hare's rule but the Borda scores are used instead of the Plurality scores. Coombs' Procedure eliminates the alternative with the lowest Antiplurality score. Nanson's Procedure eliminates all alternatives which Borda scores are lower than the average Borda score. Clearly, for each of these rules, given an unstable voting situation certain scores for some two alternatives must be close, and it is easy to specify how close.

The following obvious lemma explains how the scores can be changed during a manipulation attempt.

**Lemma 1.** *Let $F$ be any scoring rule with a representation function $f$. Let $S$ be a voting situation and let $S'$ be another voting situation which occurred as a result of change of mind of $k$ voters. Then*

$$|Sc_f(S, a) - Sc_f(S', a)| \leq kw_1.$$

*Proof.* Obvious. □

**Theorem 2.** *For any scoring rule or multistage elimination rule $F$, under the IAC, there exists a constant $D_m$, which depends only on $m$ but not on $n$ and $k$, such that*

$$L_F(m, n, k) \leq D_m \frac{k}{n}.$$

*Proof.* Due to Lemma 2 a voting situation $S = (n_1, n_2, \ldots, n_M)$, where $M = m!$, may be unstable if for some two alternatives $a$ and $b$

$$|Sc_f(S,a) - Sc_f(S,b)| \leq 2kw_1.$$

This means that $S$ satisfies one of the following nontrivial equations:

$$\sum_{i=1}^{M} (f(R_i,a) - f(R_i,b))x_i = p,$$

$a, b \in A$, $-2kw_1 \leq p \leq 2kw_1$. We have $4kw_1 + 1$ such equations. Hence by Theorem 1

$$L_F(m,n,k) \leq (4kw_1+1)\frac{C_M}{n} \leq D_m \frac{k}{n}$$

for $D_m = (4w_1+1)C_M$. $\square$

**Corollary 1.** *For $k = o(n)$ any scoring rule or multistage elimination rule $F$, under the IAC, is asymptotically $k$-coalitionally stable, i.e.,*

$$L_F(m,n,k) \to 0,$$

*when $n \to \infty$.*

**Corollary 2.** *Let $F$ be any scoring rule or multistage elimination rule under the IAC. Suppose that $k \sim n^\alpha$, where $0 \leq \alpha < 1$. Then the probability that a random voting situation is manipulable for $F$ by a coalition of size $k$ is in the order of $O(n^{1-\alpha})$. In particular, a single inividual can manipulate with probability which is in the order of $O(1/n)$.*

# 4 Concluding Remarks

It is remarkable that for the Copeland and the Borda rules the probability of a 1-manipulable voting situation occurring in three-alternative elections was possible to calulate exactly. This was done by Favardin, Lepelley and Serais [4] using the technique developed in Huang and Chua [6]. We do not give here the precise formula (it is rather complicated) but note that asymptotically, for the Borda rule, this probability is equal to $\frac{25}{12n}$.

In papers [8] and [4] it was shown that if we do not restrict the size of the manipulating coalition, then, for $m = 3$, the proportion of manipulable voting situations does not go to zero but to a nonzero limit, which is $\frac{14}{48}$ in the case of Plurality and approximately $\frac{1}{2}$ in the case of Borda. We see that allowing manipulating coalitions of an arbitrary size changes the situation radically.

The situation, when $k \sim \alpha n$ with sufficiently small $\alpha$ was studied in [11], where an attempt was made, for various values of $m$, to find the vectors of weights for which the corresponding scoring rule is less susceptible to $k$-manipulation.

We note that in [3] several other interesting indices of manipulability were introduced and experimentally estimated. The most interesting one is the average minimum coalition size that can successfully manipulate. To the best of my knowledge no theoretical results in this direction are known.

## 5 Aknowledgements

This paper was started when the author was visiting Bilkent University of Ankara for the autumn semester of 2001. I am greatful to Prof. Semih Koray and to Prof. Mefharet Kocatepe for making it possible. I appreciate the stimulating discussions that we had with Prof. Semih Koray.

## References

[1] Andrews G.E., 1976. The Theory of Partitions, Addison-Wesley, Massachusets.

[2] Berg,S. and Lepelley,D., 1994. On Probability Models in Voting Theory. Statist. Neerlandica 48, 133–146.

[3] Chamberlin, J., 1985. Investigation into the relative manipulability of four voting systems. Behavior Science 30, 195–203.

[4] Favardin, P., Lepelley D., Serias, J., 2002. Borda Rule, Copeland Method and Strategic Manipulation. Review of Economic Design, to appear.

[5] Gärdenfors, P., 1973. Positionalist Voting Functions. Theory and Decision 4, 1–24.

[6] Huang, H.C., Chua, V.C.H., 2000. Analytical representation of probabilities under the IAC condition. Social Choice and Welfare 17, 143–156.

[7] Kim, K.H., and Roush, F.W., 1996. Statistical Manipulability of Social Choice Functions. Group Decision and Negotiation 5, 263–282.

[8] Lepelley D. and Mbih B., 1987. The proportion of Coalitionally Unstable Situations under the Plurality Rule. Economic Letters 24, 311–315.

[9] Pazner, E.A., 1978. Cheatproofness Properties of the Plurality Rule in Large Societies. The Review of Economic Studies 45, 85–91.

[10] Peleg B., 1979. A Note on Manipulability of Large Voting Schemes. Theory and Decision 11, 401–412.

[11] Saari D., 1990. Susceptibility to manipulation. Public Choice 64, 21–41.

[12] Slinko, A., 2002. On Asymptotic Strategy-Proofness of Classical Social Choice Rules. Theory and Decision, to appear.