Incentive-Compatible Voting Rules with Positively Correlated Beliefs *

Mohit Bhargava †, Dipjyoti Majumdar ‡ and Arunava Sen §

April 10, 2011

Abstract

We study the consequences of positive correlation of beliefs in the design of voting rules in a model with an arbitrary number of voters. We propose two kinds of positive correlation, one based on the Kemeny distance between linear orders called K-correlation and another on the likelihood of agreement of the k best alternatives (for any k) of two orders called TS correlation. We show that K-correlation implies TS-correlation. We characterize the set of Ordinally Bayesian Incentive-Compatible (OBIC) (d'Aspremont and Peleg (1988)) voting rules with TS-correlated beliefs and additionally satisfying robustness with respect to local perturbations. We provide an example of a voting rule that satisfies OBIC with respect to *all* TS and K-correlated beliefs. However global robustness of OBIC with respect to either K or TS-correlation together with efficiency leads to dictatorship (provided that there are at least three alternatives). The generally positive results contrast sharply with the negative results obtained for the independent case by Majumdar and Sen (2004) and parallel similar results in the auction design model (Crémer and Mclean (1988)).

Keywords and Phrases: Voting rules, ordinal Bayesian incentive compatibility, positive correlation, local and global robustness with respect to beliefs.

JEL Classification numbers: C70, C72.

---

*The authors would like to thank numerous conference and seminar participants for their helpful suggestions and comments. The work of A. Sen and D. Majumdar is partially supported by research grant S01398 from the Social Science and Humanities Research Council, Canada. Majumdar also acknowledges support from Fonds Québécois de la Recherché sur la Société et la Culture (Grant. No. F00569). Part of the research was conducted when Sen was a visiting researcher at the Centre Interuniversitaire de Recherche en Économie Quantitative (CIREQ). He would like to thank CIREQ for their hospitality.

†J.P. Morgan, London, United Kingdom.

‡Concordia University, Montreal, Canada and CIREQ.

§Indian Statistical Institute, New Delhi, India.

1 Introduction

A widely-held belief is that difficulties associated with satisfactory group decision-making are significantly ameliorated if differences in the objectives of the members of the group are not "large". In the limit, if all agents have the same objectives, all conflicts of interest disappear and we may expect a trivial resolution of the problem. In mechanism design theory, agents have private information about their objectives or preferences (referred to as "types"); the theory seeks to analyze collective (or social) goals (referred to as social choice functions or SCFs) which are attainable subject to the constraint that all agents have the incentive to reveal their private information truthfully (referred to as incentive-compatibility). Here too, if the private information of all agents is perfectly correlated, the issue of incentives can be typically resolved [1]. More interestingly, an extensive literature initiated by Crémer and Mclean (1988) has pointed out that in environments where monetary compensation is feasible and preferences are quasi-linear (i.e. preferences over money are not dependent on type), even a little correlation in the beliefs over types leads to a dramatic enlargement of the class of incentive-compatible SCFs.

In this paper we explore the issue of correlated beliefs in the design of voting rules. In this environment, voters have opinions or preferences on the ranking of various candidates assumed to be finite in number. These preferences (types, in this model) expressed as linear orders over the set of candidates, are private information. A SCF or voting rule is a mapping which associates a candidate with a collection of types, one for each voter. The goal of the theory is to identify SCFs which induce voters to reveal their types truthfully for every conceivable realization of these types.

We consider the plausible case where beliefs over types are positively correlated. In doing so, we have to confront the issue of how to interpret positive correlation in distributions over linear orders. We propose two definitions of positive correlation. The first is based on the well-known notion of the Kemeny distance (Kemeny and Snell (1962), Kendall (1970)). A voter's beliefs are positively correlated in this sense if she assigns higher probability to *all* other voters' types being closer to her own in the Kemeny metric. The other notion of positive correlation is based on the likelihood of other voters top k alternatives (for any k) agreeing with one's own opinion of the top k alternatives. To illustrate this notion, consider the case of voting for the annual chess Oscar award by chess journalists and experts. Assume that the three players in serious contention are Anand (A), Carlsen (C) and Kramnik (K). Assume that a voter's opinion is A followed by C followed by K. Then she believes that the event where all other voters rank A best, is strictly more likely that the event where all other voters best alternative is either C or K. In addition she believes that the event

[1] The mechanism design problem is still non-trivial because the mechanism designer may be ignorant of the common type realized. However if there are at least three agents, the problem of inducing all agents to reveal their private information truthfully can be achieved under certain conditions. See Maskin (1999).

where all other voters' best two players is $\{A, C\}$ is more likely than the event where all other voters, best two alternatives are either $\{A, K\}$ or $\{C, K\}$. We call these two notions of positive correlation, K (or Kemeny) correlation and "Top-Set" or TS correlation respectively. We show that a K-correlated belief is TS-correlated. We note that the requirements for K as well as TS-correlation are weak in the sense that the conditions for positive correlation apply only in "exceptional" circumstances (all other voters have types which are at a greater distance from one's own or all other voters are unanimous about their best k-alternatives and so on). Our choice of definitions is deliberate because as we shall see, even these weak notions lead to a dramatic increase in the possibilities for the design of incentive-compatible social choice functions, at least in certain circumstances.

The equilibrium notion that we use is that of Ordinal Bayesian Incentive-Compatibility (OBIC) introduced in d'Aspremont and Peleg (1988). This requires the probability distribution over outcomes obtained by truth-telling to first-order stochastically dominate the distribution from mis-reporting for every voter type. These distributions are obtained from a voter's beliefs about the types of the other voters and the assumption that the other voters are telling the truth. The condition is equivalent to requiring that truth-telling be optimal in terms of expected utility for all possible utility functions which represent the voter's type.

In addition to OBIC we consider two kinds of robustness conditions of the mechanism with respect to beliefs. The first is local robustness which requires the mechanism to remain incentive-compatible if voter beliefs are perturbed slightly. This leads to two notions of local robustness depending on the definition of positive correlation used: we call these K-local robustness or K-LOBIC and TS-local robustness or TS-LOBIC. The second notion of robustness considered is global robustness where the mechanism remains incentive-compatible with respect to all beliefs that are positively correlated. Once again, we have two kinds of global robustness depending on the definition of positive correlation used and we call these K global robustness or K-ROBIC and TS global robustness or TS-ROBIC. The relationship between K correlation and TS-correlation leads to obvious relationships between K-LOBIC and TS-LOBIC mechanisms or SCFs and between K-ROBIC and TS-ROBIC SCFs. The motivation of imposing robustness requirements on beliefs is the well-known Wilson doctrine (Wilson, 1987). Robust mechanisms have the attractive feature that they continue to implement the objectives of the mechanism designer even if the designer or the voters make errors in their beliefs.

Our results are as follows. We characterize the class of TS-LOBIC SCFs subject to the weak requirement of unanimity. In particular, we provide a necessary and sufficient condition that a SCF needs to satisfy in order that there exist some neighborhood of TS-correlated beliefs such that the SCF is OBIC with respect to all beliefs in the neighborhood. It is clear that if truth-telling for a particular type is weakly dominated by a mis-report for a SCF, then the SCF cannot be locally robust incentive compatible with respect to any class of beliefs. We show that a minor modification of this condition to take into account the ordinal nature of OBIC, is also *sufficient* if TS-correlation is considered. We give an example to show that this condition is not sufficient for K-LOBIC. We also prove a general possibility result in this regard. We show that any SCF satisfying the property of neutrality and elementary monotonicity (a large class, including for instance, SCFs derived from scoring correspondences) is TS-LOBIC. Moreover they are incentive-compatible in a neighborhood of the *uniform prior*.

We also analyze the structure of SCFs satisfying various ROBIC requirements. In view of the weak notion of positive correlation, one may be conjecture that imposing robustness with respect to all positively correlated beliefs on the complete domain of preferences, will lead to truth-telling being a weakly dominant strategy, i.e. dictatorship when there are at least three alternatives. This is not true - we provide an example of a non-dictatorial SCF satisfying unanimity which is TS-ROBIC (and hence K-ROBIC). However if we additional impose the requirement of efficiency, the K-ROBIC (and hence TS-ROBIC) requirement precipitates dictatorship provided that there are at least three alternatives.

Our results contrast sharply with the negative results obtained in Majumdar and Sen (2004) for the case of independent beliefs. In this case, there is a generic set of beliefs for each voter such that OBIC with respect to *any* belief in this set is equivalent to dictatorship where truth-telling is of course, a weakly dominant strategy. There are beliefs, such as the uniform prior with respect to which a wide class of SCFs are OBIC. However, even local robustness cannot be satisfied for any non-dictatorial SCF because of the generic impossibility result. In the positively correlated case on the other hand, we demonstrate significant possibility results with local robustness. There even exist non-dictatorial SCFs satisfying unanimity which are OBIC with respect to *all* positively correlated beliefs although they must be inefficient.

Our results are in the same spirit as the possibility results in auction design theory with correlated valuations (Crémer and Mclean, 1988). However, our results and arguments bear no resemblance to their auction theory counterparts because of at least two significant differences between the models. The first is that monetary transfers which are at the heart of the possibility results in the auction model, are not permitted in the voting model. The second is that the nature of types in the voting model (linear orders) is very different from its counterpart in the auction model (a non-negative real number or vector). The notion of correlation in the voting model is therefore more delicate. Several alternative approaches and definitions are possible and the results depend on the choices made. Finally, our results are different because we address a different question from that in (Crémer and Mclean, 1988). In particular, we investigate the structure of social choice functions that satisfy certain robustness properties with respect to beliefs in addition to standard incentive-compatibility requirements.

The paper is organized as follows. In Section 2 we try to explain why correlation of types may help in mechanism design in our model. Section 3 introduces the notations and definitions Section 4 discusses alternative notions of positive correlation while Sections 5 and

6 deal with incentive-compatibility with local and global robustness respectively. Section 7 concludes. The proof of the main theorem in Section 6 is contained in the Appendix.

2 Why does Correlation of types help in Mechanism Design in Voting Models?

Consider the case where there are two voters 1 and 2 and three alternatives a, b and c to choose from. A voter's type is one of the six orderings of the alternatives. These types will represented by abc etc which signifies "a is preferred to b is preferred to c". A social choice function or voting rule is a 6×6 matrix where each entry in the matrix is an alternative.

Consider a "partial" social choice function described in the array below. Thus, if the row voter's type is abc, the outcome is a if the column voter's type is abc, bca or cab. If the row voter's type is acb, the outcome is a if the column voter's type is bac or bca.

$$\begin{array}{ccccccc} & abc & acb & bac & bca & cab & cba \\ abc & a & . & . & a & a & . \\ acb & . & . & a & a & . & . \end{array} \tag{1}$$

Suppose the row voter's type is abc. Suppose further that she has a cardinal representation of her type where the utility of alternative a is 1, that of c is 0 and that of b is arbitrarily close to 0. What is the expected utility of this voter from truth-telling assuming that the column voter tells the truth? It clearly depends on the prior beliefs of the row voter of type abc about the type of the column voter. It is, in fact

$$\mu_1(abc|abc) + \mu_1(bca|abc) + \mu_1(cab|abc) \tag{2}$$

where $\mu_1(abc|abc)$ is the row voter's belief that the column voter's type is abc conditional on the row voter's type being abc etc. By deviating to acb, the row voter of type abc will obtain the expected utility

$$\mu_1(bac|abc) + \mu_1(bca|abc) \tag{3}$$

Incentive Compatibility will then require

$$\mu_1(abc|abc) + \mu_1(cab|abc) \geq \mu_1(bac|abc) \tag{4}$$

Now consider a row voter of type acb with a utility representation where the utility of a is 1, that of b is 0 and that of c arbitrarily close to 0. In order for this type not to deviate to abc, we require

$$\mu_1(bac|acb) \geq \mu_1(abc|acb) + \mu_1(cab|acb) \tag{5}$$

If the row voter's beliefs are *independent*, then the probabilities are not conditional on her type realization. Removing the dependence of beliefs on the row voter's types, inequalities 4 and 5 yield the following equality:

$$\mu_1(bac) = \mu_1(abc) + \mu_1(cab) \tag{6}$$

Observe that the equality above cannot hold for a "generic" belief over the column voter's type. If it does for some belief, a small "perturbation" will destroy it. The only way for incentive-compatibility to be maintained for a generic belief is for all the a's to line up along the same column, i.e. if the outcome is a when the row voter's type is abc and the column voter's type is t_2, then the outcome is also a when the row and column voter's types acb and t_2 respectively. In fact, there are several restrictions of this sort implied by the independence and genericity assumptions. Majumdar and Sen (2004) demonstrate that if there are at least three alternatives, incentive-compatibility implies dictatorship.

A critical observation is that if beliefs are correlated then the distribution of the column voter's type conditional on different realizations of the row voter's types, are *distinct*. Hence inequalities such as 4 and 5 can hold without precipitating a restriction such as 6. Consequently a much wider class of social choice functions are incentive-compatible. The rest of the paper explores the class of incentive-compatible social choice functions under different notions of positive correlation.

We now proceed to details.

3 Notation and Definitions

The set of voters is $N = \{1, \ldots, n\}$. Individual voters are denoted by i, j etc. The set of outcomes is the set A with $|A| = m$. Elements of A will be denoted by a, b, c, d etc. Let $\mathbb{P}$ denote the set of strict orderings[2] of the elements of A. A typical preference ordering or type for a voter will be denoted by P_i and for all $a, b \in A$ and $a \neq b$, aP_ib will be interpreted as "a is strictly better than b according to P_i". A preference profile is an element of the set $\mathbb{P}^n$. Preference profiles will be denoted by $P, \bar{P}, P'$ etc and their i-th components as $P_i, \bar{P}_i, P'_i$ respectively with $i \in N$.

For all $P_i \in \mathbb{P}$ and $k = 1, \ldots, M$, let $r_k(P_i)$ denote the k^{th} ranked alternative in P_i, i.e., $r_k(P_i) = a$ implies that $|\{b \neq a | bP_ia\}| = k - 1$. For all $i \in N$, for any $P_i \in \mathbb{P}$ and for any $a \in A$, let $B(a, P_i) = \{b \in A | bP_ia\} \cup \{a\}$. Thus $B(a, P_i)$ is the set of alternatives that are weakly preferred to a under P_i. For any $k = 1, \ldots, m$, $B(r_k(P_i), P_i)$ is the set of alternatives which are ranked k or higher in the ordering P_i. In order to economize on notation, we shall denote $B(r_k(P_i), P_i)$ simply as $B_k(P_i)$.

Definition 1 *A Social Choice Function or (SCF) f is a mapping $f : \mathbb{P}^n \rightarrow A$.*

[2]A strict ordering is a complete, transitive and antisymmetric binary relation.

We now state some familiar axioms on SCFs which we will use at various places in the paper.

Definition 2 *A SCF f is unanimous or satisfies unanimity if $f(P) = a_j$ whenever $a_j = r_1(P_i)$ for all voters $i \in N$.*

The axiom states that in any situation where all individuals agree on some alternative as the best, the SCF must respect this consensus. A stronger requirement than unanimity is the notion of Pareto-efficiency or simply, efficiency. This requires that it should not be possible to make all voters better-off relative to the outcome of the SCF at any preference profile.

Definition 3 *A SCF f is efficient or satisfies efficiency if for all profiles $P \in \mathbb{P}^n$, there does not exist an alternative $x \in A$ such that $xP_if(P)$ for all $i \in N$.*

A dictatorial SCF picks a particular voter's best alternative at every preference profile.

Definition 4 *A SCF f is dictatorial if there exists a voter $i \in N$ such that for all profiles $P \in \mathbb{P}^n$, $f(P) = r_1(P_i)$.*

The fundamental assumption in strategic voting theory is that a voter's preference ordering is her private information. The objective of a mechanism designer is to design SCFs which provide appropriate incentives for voters to reveal their private information. A standard requirement (for example Gibbard (1973) and Satterthwaite (1975)) is for SCFs to be dominant strategy incentive-compatible or strategy-proof. In such a SCF no voter can profitably misrepresent her preferences irrespective of what (the) other voter(s) reveal as their preferences.

Definition 5 *A SCF f is dominant strategy incentive-compatible or strategy-proof if, for all $P_i, P_i' \in \mathbb{P}$, and for all $P_{-i} \in \mathbb{P}^{n-1}$ either $f(P_i, P_{-i}) = f(P_i', P_{-i})$ or $f(P_i, P_{-i})P_if(P_i', P_{-i})$ holds.*

Gibbard (1973) and Satterthwaite (1975) show that if $|A| \geq 3$, every strategy-proof SCF satisfying unanimity is dictatorial. We employ a weaker notion of incentive-compatibility.

Definition 6 *A belief for voter i is a probability distribution on the set $\mathbb{P}^n$, i.e. it is a map $\mu_i : \mathbb{P}^n \to [0, 1]$ such that $\sum\limits_{P \in \mathbb{P}^n} \mu_i(P) = 1$.*

Clearly μ_i belongs to the unit simplex of dimension $m!^n - 1$. For all μ_i, for all $(P_i, P_{-i}) \in \mathbb{P}^n$, we shall let $\mu_i(P_{-i}|P_i)$ denote the conditional probability of P_{-i} given P_i. A *belief system* is a n-tuple of beliefs $(\mu_1, \cdots, \mu_n)$, one for each voter.

DEFINITION 7 *The utility function* $u : A \to \Re$ *represents* $P_i \in \mathbb{P}$, *if and only if for all* $a, b \in A$, *we have* $aP_ib \Leftrightarrow u(a) > u(b)$.

The notion of Ordinal Bayesian Incentive Compatibility or OBIC was introduced by d'Aspremont and Peleg (1988).

DEFINITION 8 *A SCF* f *is Ordinally Bayesian Incentive Compatible (OBIC) with respect to the belief system* $(\mu_1, \cdots, \mu_n)$ *if for all* $i \in N$, *for all* $P_i, P'_i \in \mathbb{P}$, *for all* u *representing* P_i, *we have*

$$\sum_{P_{-i} \in \mathbb{P}^{n-1}} u\left(f(P_i, P_{-i})\right) \mu_i(P_{-i}|P_i) \geq \sum_{P_{-i} \in \mathbb{P}^{n-1}} u\left(f(P'_i, P_{-i})\right) \mu_i(P_{-i}|P_i) \tag{7}$$

Suppose f is a SCF which is OBIC with respect to the belief system $(\mu_1, \cdots, \mu_n)$. Consider voter i with preference P_i. Then reporting truthfully is optimal in the sense that it yields a higher *expected* utility than that obtained by any misrepresentation. In computing this expected utility, it is assumed that voters other than i will reveal truthfully so that a probability distribution over outcomes is induced by f and voter i's beliefs, conditional on P_i, i.e. $\mu_i(.|P_i)$. Furthermore, higher expected utility from truth-telling occurs for *all* representations of the true preference P_i. An equivalent way of stating the same requirement is that truth-telling is a Bayes-Nash equilibrium of the revelation game induced by f for all possible utility representation of true preferences.

The OBIC notion is a natural and minimal way to incorporate the weaker notion of truth-telling as optimal in expectation, relative to truth-telling as a dominant strategy, in an ordinal model (which is the standard model in voting theory). A fairly obvious relationship between OBIC and dominant strategies is the following:

OBSERVATION 1 *Suppose* f *is OBIC with respect to all belief systems* $(\mu_1, \cdots, \mu_n)$. *Then* f *is strategy-proof.*

In other words, if we require f to satisfy a robustness condition that it be OBIC with respect to all belief systems, then we are requiring nothing less than f to be strategy-proof.

An aspect of OBIC which may be regarded as somewhat unsatisfactory in some quarters, is that it requires truth-telling to be optimal for every type of a voter *for all* cardinalizations of the type. A partial response to this criticism is that OBIC can be defined in terms of stochastic dominance without explicit reference to utility functions.

DEFINITION 9 *The SCF* f *is OBIC with respect to the belief system* $(\mu_1, \cdots, \mu_n)$ *if for all* $i \in N$, *for all integers* $k = 1, \ldots, m$ *and for all* P_i *and* P'_i,

$$\begin{aligned} & \mu_i(\{P_{-i}|f(P_i, P_{-i}) \in B_k(P_i)\}|P_i) \\ \geq \quad & \mu_i(\{P_{-i}|f(P'_i, P_{-i}) \in B_k(P_i)\}|P_i) \end{aligned} \tag{8}$$

Suppose f satisfies OBIC with respect to $(\mu_1, \cdots, \mu_n)$. Consider voter i with preferences P_i. Then the aggregate probability induced by f on the first k alternatives of her true preference P_i for any $k = 1, \ldots, m$, is maximized by truth-telling.

We now turn our attention to the issue of positively correlated beliefs.

4 Positive Correlation

In this section we introduce two different notions of positive correlation. The first one (K-correlation) is in terms of a distance function on the set of preference orderings. Perhaps the best-known distance metric in finite, ordinal models is the *Kemeny metric* (Kemeny and Snell (1962),Kendall (1970)). It has been used widely in the literature on social welfare functions, for instance Bossert and Storcken (1992), Baigent (1987).

The Kemeny Metric: Let $P_i \in \mathbb{P}$. Two alternatives $a, b \in A$ are said to be adjacent in P_i if there does not exist any other alternative between them in P_i; formally, if there exists $k \in \{1, \ldots, m-1\}$ such that either $r_k(P_i) = a$ and $r_{k+1}(P_i) = b$ or $r_k(P_i) = b$ and $r_{k+1}(P_i) = a$. A *transposition* of a and b in P_i is the ordering obtained by switching the ranks of a and b in P_i leaving all other alternatives unchanged. The Kemeny distance between two orderings P_i and P'_i, denoted by $d(P_i, P'_i)$ is the number of transpositions required to change P_i to P'_i. For instance, if $A = \{a, b, c\}$, and P_i, P'_i are given by aP_ibP_ic and $cP'_iaP'_ib$, then $d(P_i, P'_i) = 2$. Generally, $d(P_i, P'_i) \in \{0, 1, \ldots, \binom{m}{2}\}$ for any $P_i, P'_i \in \mathbb{P}$.

Definition 10 *(K-correlation) A belief μ_i for voter i is said to be positively K-correlated if, for all preference profiles (P_i, P_{-i}) and (P_i, P'_{-i}),*

$$[d(P_i, P_j) < d(P_i, P'_j) \ \ \textit{for all } j \neq i] \Rightarrow \mu_i(P_{-i}|P_i) > \mu_i(P'_{-i}|P_i)$$

Thus μ_i is positively correlated in this sense if the following holds: voter i of type P_i considers it more likely that the types of the other voters is P_{-i} rather than P'_{-i} if for all $j \neq i$ the Kemeny distance between P_i and P_j is less than the Kemeny distance between P_i and the corresponding P'_j.

We denote by K^*, the set of all positively K-correlated beliefs.

The notion of K-correlation may be regarded as a weak requirement because it imposes restrictions the beliefs of voter i only when the preferences of *all* voters other than i are "farther away" for P_i.

We propose an alternative notion of positive correlation which we call "Top-Set" or TS-correlation. Consider a voter with beliefs μ_i and type P_i. Consider the set of k-best alternatives in P_i, $B_k(P_i)$ for some $k = 1, \ldots, m$. Let $D \subset A$ be such that $|D| = k$ and $D \neq B_k(P_i)$. Now consider the following two events:

- Event I: The k- best alternatives for all voters $j \neq i$ is $B_k(P_i)$.

- Event II: The k-best alternatives for all voters $j \neq i$ is D.

The belief μ_i is TS-correlated if Event I is considered strictly more likely than Event II according to the conditional distribution $\mu_i(.|P_i)$. In other words, if all voters other than i have the same k-best alternatives, then i considers it more likely this set coincides with her k-best alternatives than the case when it does not.

Definition **11** *(TS-Correlation) A belief for voter i, μ_i is positively TS-correlated if for all P_i and for all $k = 1, \ldots, m-1$*

$$\sum_{\{P_{-i}|B_k(P_j)=B_k(P_i)\ \forall j\neq i\}} \mu(P_{-i}|P_i) > \sum_{\{P_{-i}|B_k(P_j)=D\ \forall j\neq i\}} \mu(P_{-i}|P_i) \tag{9}$$

where $D \subset A$, $D \neq B_k(P_i)$ and $|D| = k$.

We denote by TS^* the set of all μ satisfying TS -correlation.

The following examples illustrate both notions of correlation.

Example **1** Let $N = \{1, 2\}$ and $A = \{a, b, c\}$. Consider the following belief μ_i which generates the conditional beliefs $\mu_i(.|abc)$ specified below: [3]

$$\begin{array}{ccccccc} & abc & acb & bac & bca & cab & cba \\ abc & \mu_i^1 & \mu_i^2 & \mu_i^3 & \mu_i^4 & \mu_i^5 & \mu_i^6 \end{array} \tag{10}$$

where $\mu_i^1 = \mu_i(abc|abc), \ldots, \mu_i^6 = \mu_i(cba|abc)$.

Observe that

$$\mu_i \in K^* \Rightarrow \begin{cases} \mu_i^1 > \mu_i^2, \mu_i^3, \mu_i^4, \mu_i^5, \mu_i^6 \\ \mu_i^2 > \mu_i^4, \mu_i^5, \mu_i^6 \\ \mu_i^3 > \mu_i^4, \mu_i^5, \mu_i^6 \\ \mu_i^4 > \mu_i^6 \\ \mu_i^5 > \mu_i^6 \end{cases} \tag{11}$$

On the other hand,

$$\mu_i \in TS^* \Rightarrow \begin{cases} \mu_i^1 + \mu_i^2 > \mu_i^3 + \mu_i^4 \\ \mu_i^1 + \mu_i^2 > \mu_i^5 + \mu_i^6 \\ \mu_i^1 + \mu_i^3 > \mu_i^2 + \mu_i^5 \\ \mu_i^1 + \mu_i^3 > \mu_i^4 + \mu_i^6 \end{cases} \tag{12}$$

It is easy to verify if μ_i satisfies the system 11, then μ_i satisfies the system 12. The converse is not true; for instance, pick $\mu_i^1 = 0.5$, $\mu_i^2 = 0.05$, $\mu_i^3 = 0.05$, $\mu_i^4 = 0.05$, $\mu_i^5 = 0.05$ and $\mu_i^6 = 0.3$.

[3] Here abc denotes the ordering "a is preferred to b preferred to c" etc.

Example 2 Let $N = \{1,2,3\}$ and $A = \{a,b,c\}$. Suppose voter 1's type is abc. We let (for instance) $\mu_1(acb, bac|abc)$ denote the conditional probability of voter 2 and 3's types being acb and bac respectively. Both K and TS correlation involve too many inequalities to be completely listed conveniently. We note, however that if $\mu_1 \in K^*$, then the following inequalities hold:

$$\mu_1 \in K^* \Rightarrow \begin{cases} \mu_1(abc, abc|abc) > \mu_1(bac, bac|abc) \\ \mu_1(abc, acb|abc) > \mu_1(bac, bca|abc) \\ \mu_1(acb, abc|abc) > \mu_1(bca, bac|abc) \\ \mu_1(acb, acb|abc) > \mu_1(bca, bca|abc) \end{cases} \tag{13}$$

On the other hand, if $\mu_1 \in TS^*$, we must have

$$\begin{aligned} &\mu_1(abc, abc|abc) + \mu_1(abc, acb|abc) + \mu_1(acb, abc|abc) + \mu_1(acb, acb|abc) \\ > &\mu_1(bac, bac|abc) + \mu_1(bac, bca|abc) + \mu_1(bca, bac|abc) + \mu_1(bca, bca|abc) \end{aligned} \tag{14}$$

Inequality 14 is the requirement that the probability that voters' 2 and 3 have the same best alternative a as voter 1, is greater than the probability that they have a common best alternative, b different from a. Note as earlier, that if 13 is satisfied, then 14 is satisfied as well.

We show below that this relationship holds generally.

Proposition 1 $K^* \subset TS^*$.

Proof: Pick $\mu_i \in K^*$. Pick orderings P_i, P_j and an integer $k \le m$ such that $B_k(P_i) = B_k(P_j)$. Let $B \subset A$ be such that $B \ne B_k(P_i)$ and $|B| = k$. There must exist two sets of distinct L alternatives ($L \le k$), say $a_{i_1}, \dots, a_{i_L}$ and $b_{i_1}, \dots, b_{i_L}$ such that $a_{i_1}, \dots, a_{i_L} \in B_k(P_i) \setminus B$ and $b_{i_1}, \dots, b_{i_L} \in B \setminus B_k(P_i)$.

Consider a bijection $\sigma : A \to A$ defined as follows:

- $\sigma(a) = a$ for all $a \in A \setminus \{a_{i_1}, \dots, a_{i_L}, b_{i_1}, \dots, b_{i_L}\}$
- $\sigma(a_{i_l}) = b_{i_l}$ for $l = 1, \dots, L$.
- $\sigma(b_{i_l}) = a_{i_l}$ for $l = 1, \dots, L$.

For an arbitrary ordering P_r and a bijection σ, we define P_r^σ to be the following ordering:

$$\text{for all } x, y \in A, \quad [xP_r y] \Leftrightarrow [\sigma(x) P_r^\sigma \sigma(y)]$$

Since $B_k(P_j) = B_k(P_i)$, $b_{i_l} \notin B_k(P_i)$, $l = 1\ldots, L$ and $a_{i_l} \in B_k(P_i)$, $l = 1\ldots, L$, we have $a_{i_l} P_j b_{i_l}$ and $b_{i_l} P_j^\sigma a_{i_l}$ for all $l = 1, \ldots, L$. Thus for every P_j such that $B_k(P_j) = B_k(P_i)$, there exists a P_j^σ such that $B_k(P_j^\sigma) = B$ and $d(P_i, P_j) < d(P_i, P_j^\sigma)$. The last inequality follows from the fact that $a_{i_l} P_i b_{i_l}$ and $a_{i_l} P_j b_{i_l}$ but $b_{i_l} P_j^\sigma a_{i_l}$ for $l = 1, \ldots, L$ and the remaining alternatives are ranked in the same way in P_j and P_j^σ. Now consider an $n-1$ preference profile $P_{-i} \equiv (P_j)$, $j \neq i$ where $B_k(P_i) = B_k(P_j)$ for all $j \neq i$. Let $P_{-i}^\sigma \equiv (P_j^\sigma)_{j\neq i}$. Since $\mu_i \in K^*$, we have $\mu_i(P_{-i}|P_i) > \mu_i(P_{-i}^\sigma|P_i)$. Since the above inequality holds for every pair (P_{-i}, P_{-i}^σ), we have,

$$\sum_{\{P_{-i}|B_k(P_j)=B_k(P_i)\ \forall j\neq i\}} \mu_i(P_{-i}|P_i) > \sum_{\{P_{-i}|B_k(P_j)=B\ \forall j\neq i\}} \mu_i(P_{-i}|P_i) \tag{15}$$

Inequality 15 establishes the Proposition. ■

Observation **2** Since the notions of K and TS correlations are defined in terms of strict inequalities, it follows that for any belief μ_i that is K (resp. TS) correlated, there will exist an ϵ neighborhood of beliefs that are also K (resp. TS) correlated.

We note that other notions of positive correlation in this model can be proposed. For instance, we can define a dual of TS-correlation where a voter believes that her k worst-ranked alternatives are most likely to be the k worst ranked alternatives of the other voter. Notions can also be built using classical concepts in statistics such as Spearman's coefficient of rank correlation. We do not pursue these lines of research any further since both K and TS offer rich and interesting possibilities.

5 Incentive-Compatibility with Local Robustness

In this section we explore incentive-compatible SCFs which satisfy an additional *local robustness* property. The latter requires the SCF to remain incentive-compatible if the belief of each voter is slightly perturbed. Successful information revelation occurs in such SCFs even if the mechanism designer makes "small mistakes" in his assessment of voter beliefs.

Definition **12** *A SCF f is K-locally robust OBIC or K-LOBIC with respect to the belief system μ if*

1. *$\mu_i \in K^*$ for all i and*
2. *there exists $\epsilon > 0$ such that f is OBIC with respect to all μ' such that $\mu' \in B_\epsilon(\mu)$.* [4]

[4] $B_\epsilon(\mu_i)$ denotes the open ball of radius ϵ centered at μ_i.

Consider a belief system μ where μ_i is K-correlated for each voter i. Then f is K-LOBIC with respect to μ if f is OBIC with respect every belief system in some neighborhood of μ. In fact, all the perturbed beliefs are also K-correlated. We say that f is K-LOBIC if there exists a belief system μ such that f is K-LOBIC with respect to μ. We can define local robustness with respect to TS-correlation analogously.

DEFINITION **13** *A SCF f is TS-locally robust OBIC or TS-LOBIC with respect to the belief system μ if*

1. *$\mu_i \in TS^*$ for all i and*

2. *there exists $\epsilon > 0$ such that f is OBIC with respect to all μ' such that $\mu' \in B_\epsilon(\mu)$.*

OBSERVATION **3** Since $K^* \subset TS^*$, the set of SCFs that are K-LOBIC with respect to K-correlation is a subset of the set of SCFs that are TS-LOBIC. Moreover the set inclusion is strict as the following example shows.

EXAMPLE **3** Let $N = \{1, 2\}$ and $A = \{a, b, c\}$. Let f^1 be the scoring rule with score vector $(2, 1.5, 0)$ and tie breaking in favor of agent 1. This SCF is described in the table below with voter 1 and 2's preference orderings represented by rows and columns respectively.

	abc	*acb*	*bac*	*bca*	*cab*	*cba*
abc	*a*	*a*	*a*	*b*	*a*	*b*
acb	*a*	*a*	*a*	*c*	*a*	*c*
bac	*b*	*a*	*b*	*b*	*a*	*b*
bca	*b*	*c*	*b*	*b*	*c*	*b*
cab	*a*	*c*	*a*	*c*	*c*	*c*
cba	*b*	*c*	*b*	*c*	*c*	*c*

(16)

We claim that f^1 is TS-LOBIC but not K-LOBIC.

We first demonstrate the latter. In fact we can show that f^1 does not satisfy OBIC with respect to *any* belief that is K-correlated. To see this, consider voter 2 with preferences *abc*. Let μ_2 be an arbitrary belief satisfying K-correlation. Consider voter 2 with preference ordering *abc*. Then OBIC with respect to the belief pair $(., \mu_2)$ requires

$$\mu_2(abc|abc) + \mu_2(acb|abc) + \mu_2(cab|abc) \geq \mu_2(abc|abc) + \mu_2(acb|abc) + \mu_2(bac|abc)$$

This is required so that voter 2 who puts a very high utility weight on a relative to b and c does not gain by misreporting *acb*. But the above inequality implies $\mu_2(cab|abc) \geq \mu_2(bac|abc)$. However, since $d(cab, abc) = 2 > d(bac, abc) = 1$, K-correlation requires $\mu_2(bac|abc) > \mu_2(cab|abc)$. Hence f^1 is not OBIC for any belief of voter 2 which is K-correlated.

We now show that f^1 is TS-LOBIC. It is easy to verify that truth-telling is weakly dominant for voter 1 of all types. In the case of voter 2, the following inequalities for μ_2 are necessary and sufficient in order that f^1 be OBIC with respect to the belief pair $(., \mu_2)$: $\mu_2(cab|abc) > \mu_2(bac|abc)$, $\mu_2(bac|acb) > \mu_2(cab|acb)$, $\mu_2(cba|bac) > \mu_2(abc|bac)$, $\mu_2(abc|bca) > \mu_2(cba|bca)$, $\mu_2(bca|cab) > \mu_2(acb|cab)$ and $\mu_2(acb|cba) > \mu_2(bca|cba)$.

These inequalities is easily satisfied by a belief μ_2 satisfying TS-correlation as the following matrix of conditional probabilities shows. In the number associated with row i and column j is the probability $\mu_2(i|j)$.

$$\begin{array}{ccccccc} & abc & acb & bac & bca & cab & cba \\ abc & 0.50 & 0.10 & 0.08 & 0.12 & 0.10 & 0.10 \\ acb & 0.10 & 0.50 & 0.10 & 0.10 & 0.08 & 0.12 \\ bac & 0.08 & 0.12 & 0.50 & 0.10 & 0.10 & 0.10 \\ bca & 0.10 & 0.10 & 0.10 & 0.50 & 0.10 & 0.08 \\ cab & 0.12 & 0.08 & 0.10 & 0.10 & 0.50 & 0.10 \\ cba & 0.10 & 0.10 & 0.12 & 0.08 & 0.10 & 0.50 \end{array} \tag{17}$$

Moreover since all the necessary inequalities (for both OBIC and TS-correlation) are satisfies strictly, they will continue to be satisfied if the conditional probabilities are perturbed slightly. Hence f^1 is TS-LOBIC.

There are SCFs which are not TS-LOBIC as the next example demonstrates.

Example 4 Let $A = \{a, b, c\}$ and $N = \{1, 2\}$. Consider the SCF f^2 as shown in the table below.

$$\begin{array}{ccccccc} & abc & acb & bac & bca & cab & cba \\ abc & a & a & a & b & a & b \\ acb & a & a & a & c & a & c \\ bac & c & a & b & b & a & b \\ bca & b & b & b & b & c & b \\ cab & a & b & a & c & c & c \\ cba & c & c & b & c & c & c \end{array} \tag{18}$$

Consider voter 2 with preference abc who considers misreporting via acb. Then she will lose by misreporting if voter 1 has preference cba by getting c instead of b; she will gain if voter 1's preference is bac by getting a instead of c. Suppose f^2 is OBIC with respect to some belief pair (μ_1, μ_2). By virtue of the robustness criterion, we can assume $\mu_2(bac|abc), \mu_2(cab|abc) > 0$. Now pick a utility representation u of abc such that $u(a) = 1, u(b) = \alpha, u(c) = 0$ where $0 < \alpha < 1$. The difference in expected utility between truth-telling and lying is $\Delta = (1 - \alpha)\mu_2(cab|abc) - \mu_2(bac|abc)$. Since $\mu_2(cab|abc), \mu_2(bac|abc) > 0$, Δ can be made strictly less than 0 by choosing α sufficiently close to 1. This contradicts the assumption that f^2 is OBIC with respect to (μ_1, μ_2).

The example above suggests a necessary condition that a TS-LOBIC SCF must satisfy. Since all conditional probabilities can be assumed to be non-zero by local robustness, expected utility for a type cannot be maximized by truth-telling if misrepresentation weakly dominates truth-telling. However in addition, the gain from truth-telling cannot be "washed out" relative to the gain from misrepresentation by picking a different utility representation. We formalize this notion below.

Definition **14** *A SCF $f : \mathbb{P}^n \to A$ satisfies Ordinal Non-Domination (OND) if for all i, for all P_i, P_i' and P_{-i} such that $f(P_i', P_{-i}) P_i f(P_i, P_{-i})$, there exists P_{-i}' such that,*

1. *Either $f(P_i, P_{-i}') = f(P_i', P_{-i})$ or $f(P_i, P_{-i}') P_i f(P_i', P_{-i})$ and*
2. *Either $f(P_i, P_{-i}) = f(P_i', P_{-i}')$ or $f(P_i, P_{-i}) P_i f(P_i', P_{-i}')$.*

Consider the SCF f^2 in Example 4. Observe that $f^2(bac, acb) = a$ is strictly preferred to $c = f^2(bac, abc)$ under abc. According to OND, there must exist another preference ordering for voter 1 where 2 does strictly better by reporting abc rather than acb. The only candidate for such an ordering for 1 is cab. However $f^2(cab, acb)$ is strictly preferred to $f^2(bac, abc)$ violating part 1 of the OND condition. The example clearly shows how OBIC will now fail: by choosing a suitable utility representation, the gain from telling the truth when 1's report is cab can be made arbitrarily small relative to the gain from lying when 1's report is bac. The necessity of part 2 of OND can be demonstrated similarly.

The OND condition is weak as the example below suggests.

Example **5** Let $A = \{a, b, c\}$ and $N = \{1, 2, 3\}$. Let f^p be the *Plurality Rule* with voter 1 as the tie-breaker. In other words, the outcome at any profile is the alternative that is ranked first by the largest number of voters. In case of a tie, voter 1's best alternative is selected.

Observe that voter 1 has a dominant strategy to be truthful. Suppose voter 2's type is abc. She can profitable deviate from truth-telling only when voter 1 and 3's best alternatives are c and b respectively. Then 2 obtains c by telling the truth and b by deviating to a type where b is the best alternative. On the other hand if voter 1 and 3's best alternatives are c and a respectively, then 2 obtains a by truth-telling and c when deviating to a type where b is the best alternative. It is easy to verify that these profiles and outcomes satisfy the requirements of OND. An identical argument holds for voter 3.

More examples of SCFs satisfying OND will be provided later in the section. We now show that OND is necessary and almost sufficient for the TS-LOBIC property to hold.

Theorem **1** *If a SCF is TS-LOBIC, it satisfies OND. If a SCF satisfies unanimity and OND it is TS-LOBIC.*

Proof: We first prove that if a SCF is TS-LOBIC it satisfies OND.

Let f be a TS-LOBIC SCF. Then, for all i there exists $\mu_i \in TS^*$ such that for all P_i, P_i' and u representing P_i, we have,

$$\sum_{P_{-i}\in\mathbb{P}^{n-1}} \mu_i(P_{-i}|P_i)\left[u(f(P_i,P_{-i}),P_i) - u(f(P_i',P_{-i}),P_i)\right] \geq 0 \tag{19}$$

Moreover inequality 19 holds for all μ_i' in a neighborhood of μ_i. Hence we can assume without loss of generality that $\mu_i(P_{-i}|P_i) > 0$ in inequality 19. Suppose that there exists P_i, P_i' and P_{-i} such that $f(P_i', P_{-i})P_if(P_i, P_{-i})$, i.e $u(f(P_i', P_{-i})) > u(f(P_i, P_{-i}))$ for all u representing P_i. Since $\mu_i(P_{-i}|P_i) > 0$, there must exist P_{-i}' such that $u(f(P_i, P_{-i}')) > u(f(P_i', P_{-i}'))$, i.e. $f(P_i, P_{-i}')P_if(P_i', P_{-i}')$, in order for inequality 19 to hold. Let L denote the set of all such P_{-i}''s.

Now suppose $f(P_i', P_{-i})P_if(P_i, P_{-i}')$ holds for all $P_{-i}' \in L$. Then we can choose a utility representation $\hat{u}$ of P_i such that $\hat{u}(f(P_i', P_{-i}))$ is arbitrarily close to 1 and $\hat{u}(f(P_i, P_{-i}'))$, $\hat{u}(f(P_i, P_{-i}))$ and $\hat{u}(f(P_i', P_{-i}'))$ are all arbitrarily close to 0. Then, the L.H.S of 19 for the utility function $\hat{u}$ can be made arbitrarily close to $-\mu_i(P_{-i}|P_i) < 0$ violating inequality 19.

Now suppose $f(P_i', P_{-i}')P_if(P_i, P_{-i})$ holds. Then we can choose a utility representation $\tilde{u}$ of P_i such that $\tilde{u}(f(P_i', P_{-i})), \tilde{u}(f(P_i, P_{-i}'))$ and $\tilde{u}(f(P_i', P_{-i}'))$ are arbitrarily close to 1 and $\tilde{u}(f(P_i, P_{-i}))$ is arbitrarily close to 0. Once again the L.H.S of 19 for the utility function $\tilde{u}$ can be made arbitrarily close to $-\mu_i(P_{-i}|P_i) < 0$ violating inequality 19.

Hence f satisfies OND.

We now consider the proof of the second part of the Theorem.

Suppose that f satisfies unanimity and OND. We will construct an open set of beliefs for each voter satisfying TS-correlation and such that f is OBIC with respect to all beliefs in this set.

Pick a voter i and an ordering P_i. For any $k \in \{1, \cdots, m\}$ define $A_k^f(P_i) = \{P_{-i}|f(P_i, P_{-i}) = r_k(P_i)\}$. Thus $A_k^f(P_i)$ is the set of preferences for voters other than i that gives under f the k^{th} ranked alternative of voter i as outcome. Define by P_{-i}^0 the preference profile for voters other than i where each voter $j \neq i$ has the preference ordering P_i. Since f satisfies unanimity, $P_{-i}^0 \in A_1^f(P_i)$.

Let $\mathcal{C}_i^*$ denote the set of probability distributions over $\mathbb{P}^n$ such that for each $\mu_i^* \in \mathcal{C}_i^*$ and P_i, the conditional distribution $\mu_i^*(.|P_i)$ satisfies the following properties:

1. $\mu_i^*(P_{-i}|P_i) > 0$ for all P_{-i}
2. $\mu_i^*(P_{-i}^0|P_i) > \sum\limits_{P_{-i}\neq P_{-i}^0} \mu_i^*(P_{-i}|P_i)$
3. for all $P_{-i} \neq P_{-i}^0$, $\mu_i^*(P_{-i}|P_i) > \sum\limits_{P_{-i}'\in\cup_{r=k+1}^{r=m} A_r^f(P_i)} \mu_i^*(P_{-i}'|P_i)$ where $P_{-i} \in A_k^f(P_i)$.

Suppose $f(P_i, P_{-i})$ is the k^{th}-ranked alternative in P_i. Then the conditional probability $\mu_i^*(P_{-i}|P_i)$ exceeds the sum of the conditional probabilities of realizing a profile P'_{-i} where the outcome $f(P_i, P'_{-i})$ is strictly worse than the k^{th}-ranked alternative in P_i. In addition, the conditional probability of realizing the profile P^0_{-i} exceeds the sum of the conditional probabilities of realizing any other ordering. There are clearly no difficulties in defining $\mathcal{C}_i^*$. Moreover, since the restrictions on the conditional probabilities are described by strict inequalities, it follows that $\mathcal{C}_i^*$ is an open set in the unit simplex of dimension $m!^n - 1$.

We claim that $\mathcal{C}_i^* \subset TS^*$. This is easily verified by noting that the term $\mu_i^*(P^0_{-i}|P_i)$ appears in the L.H.S of every inequality in the system of inequalities 12 which define TS correlation while it does not appear on the R.H.S of any one of them. In order to complete the proof, we will show that f is OBIC with respect to all beliefs $(\mu_1^*, \cdots, \mu_n^*)$ where $\mu_i^* \in \mathcal{C}_i^*$, $i \in N$.

Pick an arbitrary voter i, orderings P_i, P'_i and a utility function u representing P_i. Let $G = \{P_{-i}|f(P'_i, P_{-i})P_if(P_i, P_{-i})\}$ and $L = \{P_{-i}|f(P_i, P_{-i})P_if(P'_i, P_{-i})\}$. Pick an arbitrary $\mu_i^* \in \mathcal{C}_i^*$. In order for OBIC to be satisfied with respect to μ_i^*, we must have

$$\sum_{P_{-i}\in L} \mu_i^*(P_{-i}|P_i)\beta(P_{-i}) - \sum_{P_{-i}\in G} \mu_i^*(P_{-i}|P_i)\gamma(P_{-i}) \geq 0 \tag{20}$$

where $\beta(P_{-i}) = [u(f(P_i, P_{-i})) - u(f(P'_i, P_{-i}))]$ and $\gamma(P_{-i}) = [u(f(P'_i, P_{-i})) - u(f(P_i, P_{-i}))]$.

If $G = \emptyset$, inequality 20 is clearly satisfied. Suppose therefore that $G \neq \emptyset$. We claim that for all $P_{-i} \in G$, there exists $P'_{-i} \in L$ satisfying,

1. $\beta(P'_{-i}) > \gamma(P_{-i})$
2. $\mu_i^*(P'_{-i}|P_i) > \sum_{\{\tilde{P}_{-i}|f(P_i,P'_{-i})P_if(P_i,\tilde{P}_{-i})\}} \mu_i^*(\tilde{P}_{-i}|P_i)$

Here 1 above follows from the assumption that f satisfies OND and 2 follows from 2 and 3 in the specification of μ_i^*.

Let $\sigma : G \to L$ be a map such that for all $P_{-i} \in G$, $\sigma(P_{-i})$ is the $P'_{-i} \in L$ satisfying 1 and 2 above. Let P'_{-i} be an arbitrary element in the range of σ and let $Q(P'_{-i}) = \{P_{-i}|\sigma(P_{-i}) = P'_{-i}\}$. A critical observation is that for all $P_{-i} \in Q$, OND implies $f(P_i, P'_{-i})P_if(P_i, P_{-i})$, i.e. $Q(P'_{-i}) \subset \{\tilde{P}_{-i}|f(P_i, P'_{-i})P_if(P_i, \tilde{P}_{-i})\}$. Hence 2 above implies $\mu_i^*(P'_{-i}|P_i) > \sum_{P_{-i}\in Q(P'_{-i})} \mu_i^*(P_{-i}|P_i)$. Moreover using 1 above, we have $\mu_i^*(P'_{-i}|P_i)\beta(P'_{-i}) > \sum_{P_{-i}\in Q(P'_{-i})} \mu_i^*(P_{-i}|P_i)\gamma(P_{-i})$. Now summing up over all P'_{-i} in L and noting that OND implies that $G \subset \cup_{P'_{-i}\in L} Q(P'_{-i})$, we obtain inequality 20.

■

The proof of the first part of Theorem 1 clearly shows that OND is a necessary condition for locally robust OBIC with respect to *any* subset of prior beliefs. It applies equally to beliefs which are restricted to lie in the set of TS or K correlated beliefs or in the set of independent beliefs of for that matter, to some subset of negative correlated beliefs, howsoever defined. It is an inescapable consequence of local robustness. The sufficiency part of Theorem 1 that TS-correlation leads to the most permissive result for incentive-compatibility subject to the very mild requirement that the SCFs under consideration satisfy unanimity.

We now show that if SCFs satisfy two additional restrictions, the are TS-LOBIC with respect to beliefs that are arbitrarily close to the *uniform prior*. These additional restrictions were introduced in Majumdar and Sen (2004).

DEFINITION **15** *Let $\sigma : A \to A$ be a permutation of A. Let P^σ denote the profile $(P_1^\sigma, \ldots P_n^\sigma)$ where for all i and $a, b \in A$,*

$$aP_ib \Rightarrow \sigma(a)P_i^\sigma\sigma(b)$$

The SCF f satisfies neutrality if, for all profiles P and for al permutation functions σ, we have

$$f(P^\sigma) = \sigma(f(P)).$$

Neutrality is a standard axiom for social choice functions which ensures that alternatives are treated symmetrically.

Let P_i be an ordering and let $a \in A$. We say that P_i' represents an *elementary a-improvement of P_i* if

- for all $x, y \in A \setminus \{a\}$, $[xP_iy] \Leftrightarrow [xP_i'y]$
- $[a = r_k(P_i)] \Rightarrow [a = r_{k-1}(P_i')]$ if $k > 1$
- $[a = r_1(P_i)] \Rightarrow [a = r_1(P_i')]$.

DEFINITION **16** *The SCF f satisfies elementary monotonicity if, for all i, P_i, P_i' and P_{-i}, such that P_i' represents an elementary a-improvement of P_i,*

$$[f(P_i, P_{-i}) = a] \Rightarrow [f(P_i', P_{-i}) = a]$$

Elementary monotonicity requires that if an alternative a is the outcome at a particular profile, then it is also the outcome at the profile where a single voter makes a single upward local switch of a. Once again, this is a weak requirement and is discussed at greater length in Majumdar and Sen (2004).

Finally let $\bar{\mu}$ denote the uniform prior system of beliefs, i.e. $\bar{\mu}_i(P_{-i}|P_i) = \frac{1}{m!^{n-1}}$ for all P_i.

According to our next result, any neutral SCF satisfying elementary monotonicity is TS-LOBIC with respect to a TS-correlated prior which can be chosen arbitrarily close to the uniform prior.

Theorem 2 *Let f be a neutral SCF satisfying elementary monotonicity and unanimity. For all $\epsilon > 0$, there exists a belief system μ such that $\mu \in B_\epsilon(\bar{\mu})$ and such that f is TS-LOBIC with respect to μ.*

Proof: Pick a voter i and an ordering P_i. As in the proof of Theorem 1 let P^0_{-i} be the preference profile for voters other than i where each voter $j \neq i$ has the preference ordering P_i. Let $K \subset \{1, \dots, m\}$ be such that (i) if $k \in K$ and $k \geq 2$, then $A^f_k(P_i) = \{P_{-i} | f(P_i, P_{-i}) = r_k(P_i)\} \neq \emptyset$ and (ii) If $k = 1$ and $k \in K$ then $A^f_1(P_i) = \{P_{-i} | f(P_i, P_{-i}) = r_1(P_i)\} \setminus P^0_{-i} \neq \emptyset$. In other words, if $k \in K$ and $k \geq 2$, there exists an $n-1$ voter profile P_{-i} such that $f(P_i, P_{-i})$ is the k^{th}-ranked alternative in P_i. The index 1 is included in K if there exists a profile P_{-i} *distinct from* P^0_{-i} such that $f(P_i, P_{-i})$ is the first-ranked alternative in P_i.

Without loss of generality, let $K = \{k_1, \dots, k_L\}$ such that $k_1 < k_2 \dots < k_L$. For each $l = 1, \dots, L$ pick $\delta_l > 0$ satisfying

1. $\frac{1}{m!^{n-1}} > \delta_0 = \sum_{l=1}^{L} \delta_{k_l}$ and
2. $\delta_1 < \delta_2 < \dots < \delta_L$.

Define the conditional beliefs $\mu^*_i(.|P_i)$ as follows:

$$\mu^*_i(P_{-i}|P_i) = \begin{cases} \frac{1}{m!^{n-1}} + \delta_0 & \text{if } P_{-i} = P^0_{-i} \\ \frac{1}{m!^{n-1}} - \frac{\delta_l}{|A^f_{k_l}(P_i)|} & \text{if } P_{-i} \in A^f_{k_l}(P_i) \text{ for some } l = 1, \dots, L \end{cases} \tag{21}$$

It is clear that by choosing the δ's sufficiently small, we can generate a belief system μ^* arbitrarily close to $\bar{\mu}$.

We claim that $\mu^*_i(.|P_i)$ is TS-correlated for all P_i. Consider an arbitrary inequality in 12, for instance for some $k = 1, \dots, m$ and a set $|B| = k$ with $B \neq B_{k_l}(P_i)$. Note that $|\{P_{-i} : B_k(P_j) = B \ \forall j \neq i\}| = |\{P_{-i} : B_k(P_j) = B_k(P_i) \ \forall j \neq i\}|$. Therefore the number of terms in the L.H.S and R.H.S of every inequality in 12 has the same number of terms. Each of these terms contains $\frac{1}{m!^{n-1}}$ which can be canceled with each other. Now consider the inequality after canceling these terms. The term δ_0 appears on the L.H.S but not in the R.H.S. Therefore a lower bound for the L.H.S is when all the terms other than P^0_{-i} belong to $A^f_{k_L}(P_i)$. Hence a lower bound for the LHS is $\delta_0 - \delta_{k_L} > 0$ by construction. On the R.H.S all terms are strictly negative (in fact the maximum value it can attain is -δ_1). Clearly the L.H.S is strictly greater than the R.H.S so that $\mu^*_i(.|P_i)$ is TS-correlated. Note that there all exists a neighborhood of μ^*_i where all beliefs are TS-correlated.

We now show that for any voter i and type P_i, f satisfies incentive-compatibility with respect all priors chosen in a suitable neighborhood of $\mu^*_i(.|P_i)$. Pick an arbitrary P'_i and an integer $k \in \{1, \dots, m-1\}$. Let $\sigma : A \rightarrow A$ be a permutation such that $r_l(P_i) = r_{\sigma(l)}(P'_i)$ for all $l = 1, \dots, m$. Let $\mathbb{P}_{-i} = \{P_{-i} : f(P) \in B_k(P_i)\}$ and $\mathbb{P}^\sigma_{-i} = \{P^\sigma_{-i} : f(P) \in \sigma^{-1}(B_k(P_i))\}$ [5]. For each $P_{-i} \in \mathbb{P}_{-i}$, let $s(P_{-i}) \in \{k_1, \dots, k_L\}$ be such that $f(P_i, P_{-i}) = r_{s(P_{-i})}(P_i)$. Let

[5] $\sigma^{-1}(B_k(P_i) = \{a \in A : \sigma(a) \in B_k(P_i)\}$.

$\Delta(\mathbb{P}_{-i}) = \sum_{P_{-i}\in\mathbb{P}_{-i}} \frac{\delta_{s(P_{-i})}}{|A^f_{s(P_{-i})}(P_i)|}$. Similarly, let $\Delta(\mathbb{P}^\sigma_{-i}) = \sum_{P_{-i}\in\mathbb{P}^\sigma_{-i}} \frac{\delta_{s(P_{-i})}}{|A^f_{s(P_{-i})}(P_i)|}$.

Therefore we have

$$\mu^*_i(\{P_{-i}|f(P_i,P_{-i})\in B_k(P_i)\}|P_i) = \frac{1}{m!^{n-1}}|\mathbb{P}_{-i}| + \delta_0 - \Delta(\mathbb{P}_{-i}) \tag{22}$$

$$\mu^*_i(\{P_{-i}|f(P^\sigma_i,P_{-i})\in B_k(P_i)\}|P_i) = \frac{1}{m!^{n-1}}|\mathbb{P}^\sigma{}_{-i}| + \delta_0\mathbb{I}_{\{r_1(P_i)\in\sigma^{-1}(B_k(P_i))\}} - \Delta(\mathbb{P}^\sigma_{-i}) \tag{23}$$

Here $\mathbb{I}$ is the indicator function, i.e. $\mathbb{I}_{\{r_1(P_i)\in\sigma^{-1}(B_k(P_i))\}} = 1$ if $r_1(P_i)\in\sigma^{-1}(B_k(P_i))$ and 0 otherwise.

Majumdar and Sen (2004) prove that if f satisfies elementary monotonicity and neutrality, then

1. $|A^f_k(P_i)| \geq |A^f_t(P_i)|$ whenever $k<t$
2. $|\mathbb{P}_{-i}| \geq |\mathbb{P}^\sigma_{-i}|$.

Fix an arbitrary $k\in\{1,\ldots,m-1\}$. We wish to compare the R.H.S of equations 22 and 23. Consider the following cases.

Case I: $m!^{n-1} > |\mathbb{P}_{-i}| = |\mathbb{P}^\sigma_{-i}|$.

Let $T^0 = \{P_{-i} : P_{-i}\in\mathbb{P}_{-i}\setminus\mathbb{P}^\sigma_{-i}\}$ and $T^1 = \{P_{-i} : P_{-i}\in\mathbb{P}^\sigma_{-i}\setminus\mathbb{P}_{-i}\}$. In view of 2 above, $|T^0|\geq|T^1|\neq 0$. Pick an arbitrary $P_{-i}\in T^1$. Since $P_{-i}\notin\mathbb{P}_{-i}$, it follows that $s(P_{-i})>k$. On the other hand, for all $P_{-i}\in T^0$, we have $s(P_{-i})\leq k$. Note that for all P_{-i},P'_{-i} if $s(P_i)>s(P'_{-i})$, then $\delta_{s(P_{-i})}>\delta_{s(P'_{-i})}$ by construction and $|A^f_{s(P_{-i})}|\leq|A^f_{s(P'_{-i})}|$, so that $\frac{\delta_{s(P_{-i})}}{|A^f_{s(P_{-i})}|} > \frac{\delta_{s(P'_{-i})}}{|A^f_{s(P'_{-i})}|}$. Therefore,

$$\Delta(\mathbb{P}^\sigma_{-i}) - \Delta(\mathbb{P}_{-i}) = \sum_{P_{-i}\in T^1}\frac{\delta_{s(P_{-i})}}{|A^f_{s(P_{-i})}|} - \sum_{P_{-i}\in T^0}\frac{\delta_{s(P_{-i})}}{|A^f_{s(P_{-i})}|} > 0 \tag{24}$$

Hence,

$$\begin{aligned}&\mu^*_i(\{P_{-i}|f(P_i,P_{-i})\in B_k(P_i)\}|P_i) - \mu^*_i(\{P_{-i}|f(P^\sigma_i,P_{-i})\in B_k(P_i)\}|P_i)\\ &\geq \Delta(\mathbb{P}^\sigma_{-i}) - \Delta(\mathbb{P}_{-i})\\ &> 0\end{aligned} \tag{25}$$

Case II: $m!^{n-1} > |\mathbb{P}_{-i}| > |\mathbb{P}^\sigma_{-i}|$.

We claim that $\Delta(\mathbb{P}^\sigma_{-i}) - \Delta(\mathbb{P}_{-i}) < \delta_0$. Note that $\Delta(\mathbb{P}^\sigma_{-i}) - \Delta(\mathbb{P}_{-i}) = \sum_{l=1}^L \delta_{k_l} = \delta_0$ only if either $\mathbb{P}_{-i}$ or $\mathbb{P}^\sigma_{-i}$ is the set of all $n-1$ voter profiles, i.e. either $|\mathbb{P}_{-i}| = m!^{n-1}$ or $|\mathbb{P}^\sigma_{-i}| = m!^{n-1}$. However both cases contradict underlying assumptions for Case II to hold. Consequently $\Delta(\mathbb{P}^\sigma_{-i}) - \Delta(\mathbb{P}_{-i}) < \delta_0$. Thus,

$$
\begin{aligned}
&\mu_i^*(\{P_{-i}|f(P_i,P_{-i})\in B_k(P_i)\}|P_i)-\mu_i^*(\{P_{-i}|f(P_i^\sigma,P_{-i})\in B_k(P_i)\}|P_i)\\
&\geq \tfrac{1}{m!^{n-1}}(|\mathbb{P}_{-i}|-|\mathbb{P}^\sigma_{-i}|)+\Delta(\mathbb{P}^\sigma_{-i})-\Delta(\mathbb{P}_{-i})\\
&> \tfrac{1}{m!^{n-1}}-\delta_0\\
&> 0
\end{aligned}
\tag{26}
$$

Observe that if either Case I or II hold, then incentive-compatibility conditions hold with *strict* inequality with respect to μ_i^*. Therefore they will continue to hold in a neighborhood of μ_i^*. The only remaining case (in view of 2) is when $|\mathbb{P}_{-i}|=m!^{n-1}$. In this case $\mu_i^*(\{P_{-i}|f(P_i,P_{-i})\in B_k(P_i)\}|P_i)=1$ so that incentive-compatibility conditions will continue to hold for all beliefs.

We have established that f is TS-LOBIC at μ^* as required. ■

Observation 4 Majumdar and Sen (2004) show that a large class of "well-behaved" SCFs satisfy neutrality and elementary monotonicity. These include all *scoring rules* with neutral tie-breaking rules (for instance, by picking the maximal element amongst all alternatives with the highest score according to a fixed voter's ordering). Note that these SCFs must also satisfy the OND condition.

We now discuss our results on local robustness at greater length.

5.1 Discussion and Interpretation

Theorems 1 and 2 stand in sharp contrast to results in Majumdar and Sen (2004) for the independent beliefs case. Their main result says the following: if beliefs are independent, there exists a set which is generic such that OBIC with respect to *any* beliefs in this set implies that truth-telling must be a dominant strategy. It may be possible to find non-dictatorial SCFs for very special beliefs such as the uniform prior. However, if beliefs are picked from a slightly perturbed set, the class of incentive-compatible compatible SCFs immediately shrinks to the dictatorial class. In contrast, if beliefs are TS or K-correlated, it is possible to find SCFs that are incentively-compatible with respect to *all* beliefs in some neighborhood of beliefs.

Theorem 1 provides a very general answer to the question of what SCFs are TS-LOBIC. The proof of the second part of the Theorem constructs a class of conditional beliefs for each voter with respect to which a SCF satisfying OND and unanimity, is TS-LOBIC. These beliefs *depend* on the SCF. This should not be surprising; in the next section we show that imposing stronger notions of robustness lead to a drastic reduction in the class of incentive-compatible SCFs. The beliefs constructed are as follows: a voter i with type i puts "high" weight on all other voters types being P_i (i.e. coinciding with her own); in addition she puts higher weight on voters types being P_{-i} instead of P'_{-i} if $f(P_i, P_{-i})$ is strictly better than $f(P_i, P'_{-i})$ according to P_i. In general, one may say that voters are "optimistic" in their beliefs in the sense that they assign "much higher" probabilities to more favorable events. In this case, these events are realizations of the other voter's types which lead to better outcomes through the SCF. Loosely speaking, this is in accordance with the general intuition regarding why positive correlation may ameliorate the problems of designing incentive compatible SCFs.

Theorem 2 shows that SCFs under consideration which satisfy neutrality and elementary monotonicity are TS-LOBIC. Moreover, the neighborhood of beliefs with respect to which the SCFs are LOBIC, can be chosen to be arbitrarily close to the uniform prior, i.e. at the center of the simplex. In contrast, the neighborhood of beliefs constructed in the proof of Theorem 1 was near the vertex of the simplex where an agent believes that all other agents have the same type that she does. In the case of Theorem 2 as earlier, beliefs are constructed assuming that voters are "optimistic" about their beliefs about the types of other voters. However, it suffices for their optimism to be "very small" relative to uniform beliefs.

Do Theorems 1 and 2 go through for K-LOBIC? The answer is no; in fact Example 3 demonstrates that OND is not sufficient for K-LOBIC. The OND condition guarantees that if misrepresentation is more profitable than truth-telling for voter i if the type profile of the other voters is, say P_{-i}, then there is another type profile of the other voters, say P'_{-i} where the misrepresentation is "ordinally costlier" than truth-telling, relative to the situation at P_{-i}. In order to strengthen the condition to make it K-LOBIC necessary, additional restrictions on $d(P_i, P'_{-i})$ relative to $d(P_i, P_{-i})$ must also hold. These restrictions may be quite subtle and we do not pursue this question further.

In the next section, we consider the consequences of strengthening the robustness requirement.

6 Incentive Compatibility with Global Robustness

In this section we analyze the issue of *Global Robustness* with positively correlated beliefs.

Definition **17** *A SCF $f : \mathbb{P}^n \to A$ is K-Globally Robust OBIC (K-ROBIC) if it is OBIC with respect to all belief systems μ where $\mu_i \in K^*$ for all i.*

We have an analogous definition for TS-correlation.

Definition **18** *A SCF $f : \mathbb{P}^n \to A$ is TS-Globally Robust OBIC (TS-ROBIC) if it is OBIC with respect to all belief systems μ where $\mu_i \in TS^*$.*

Observation **5** Since $K^* \subset TS^*$, (Proposition 1), a SCF which is TS-ROBIC, is also K-ROBIC.

Our goal is to investigate the class of K and TS- ROBIC SCFs. We first focus our attention on SCFs which are K-ROBIC. Since the K-ROBIC property is clearly a strong requirement, it is reasonable to conjecture that a SCF which satisfies it, is strategy-proof. Stating it differently, it may seem plausible that the consequences of imposing incentive-compatibility with respect to all positively correlated beliefs is equivalent to imposing incentive-compatibility with respect to all beliefs. Rather surprisingly this is false as we show below.

Definition **19** *The SCF f^{us} is the unanimity with status-quo rule if there exists an alternative x such that for all profiles P,*

$$f(P) = \begin{cases} r_1(P_1) & \text{if } r_1(P_1) = r_1(P_2) = \cdots = r_1(P_n) \\ x & \text{otherwise.} \end{cases} \tag{27}$$

In other words, f^{us} picks the status quo alternative x unless both voters have a common best ranked alternative. It is clear that f^{us} is not strategy-proof. For instance suppose $A = \{a, b, x\}$ and let P be the profile where aP_1bP_1x and bP_jaP_jx for all $j \neq 1$. The outcome of f^{us} in this profile is x (the status quo alternative). But voter 1 can mis-report $bP'_1aP'_1x$ and obtain b which is better than x according to P_1. We show however that f^{us} is TS- ROBIC and therefore K-ROBIC as well.

Proposition **2** *f^{us} is TS-ROBIC.*

Proof: As before, we denote the status quo alternative by x. Pick an arbitrary voter i with ordering P_i. If $r_1(P_i) = x$, then $f^{us}(P_i, P_{-i}) = x$ for all P_{-i}. Truth-telling is a weakly dominant strategy in this case and will lead to a (weakly) higher expected payoff irrespective of the representation u of P_i and beliefs $\mu(P_{-i}|P_i)$.

Assume therefore that $r_1(P_i) = a \neq x$. Let P'_i be such that either $b = x$ or xP_ib where $r_1(P'_i) = b \neq a$. Since $f^{us}(P_i, P_{-i})$ is either a or x for all P_{-i} and $f^{us}(P'_i, P_{-i})$ is either b or x for all P_{-i}, it follows again that truth-telling will weakly dominate the strategy of misreporting via P'_i.

It follows that the only case which needs to be considered is the one where $r_1(P'_i) = b$ and bP_ix. Here voter i will gain by misreporting P'_i for all P_{-i} such that for all $j \neq i$, $r_1(P_j) = b$. Denote the set of such P_{-i}'s by G. On the other hand, i loses by misreporting P'_i for all P_{-i} such that for all $j \neq i$ $r_1(P_j) = a$. Denote the set of such P_{-i}'s by L. In particular observe that

- $f^{us}(P_i, P_{-i}) = x$, $f^{us}(P'_i, P_{-i}) = b$ for all $P_{-i} \in G$
- $f^{us}(P_i, P_{-i}) = a$, $f^{us}(P'_i, P_{-i}) = x$ for all $P_{-i} \in L$
- $f^{us}(P_i, P_{-i}) = f^{us}(P'_i, P_{-i}) = x$ for all $P_{-i} \notin G \cup L$

Let u be an arbitrary utility function that represents P_i and let $\mu_i \in TS^*$. The expected utility from truth-telling is

$$\sum_{P_{-i} \in L} u(a)\mu_i(P_{-i}|P_i) + \sum_{P_{-i} \in G} u(x)\mu_i(P_{-i}|P_i) + \sum_{P_{-i} \notin G \cup L} u(x)\mu_i(P_{-i}|P_i) \tag{28}$$

The expected utility from misreporting via P_i' is

$$\sum_{P_{-i} \in L} u(x)\mu_i(P_{-i}|P_i) + \sum_{P_{-i} \in G} u(b)\mu_i(P_{-i}|P_i) + \sum_{P_{-i} \notin G \cup L} u(x)\mu_i(P_{-i}|P_i) \tag{29}$$

Let Δ denote the gain from truth-telling. The two equations above imply

$$\Delta = [u(a) - u(x)] \sum_{P_{-i} \in L} \mu_i(P_{-i}|P_i) - [u(b) - u(x)] \sum_{P_{-i} \in G} \mu_i(P_{-i}|P_i) \tag{30}$$

Since u represents P_i, we have $u(a) > u(b) > u(x)$. Also TS correlation implies $\sum_{P_{-i} \in L} \mu_i(P_{-i}|P_i) > \sum_{P_{-i} \in G} \mu_i(P_{-i}|P_i)$ (since voter i of type P_i considers it more likely that the probability that top-ranked alternative of *all* the other voters agrees with her own (i.e. it is a) than the case that *all* the other voters have the same top-ranked alternative different from her own (i.e., b in this case)). Hence $\Delta \geq 0$ and f^{us} is TS-ROBIC.

■

The unanimity with status quo rule has some nice features. It is both anonymous [6] and neutral. However it has a serious drawback: the rule picks the status quo in many situations where both voters prefer other alternatives. It violates efficiency.

Our main result shows that imposing efficiency together with global robustness leads to dictatorial SCFs. In other words, efficiency and global robustness can be satisfied only if truth-telling is a weakly dominant strategy. Observe that robustness does not directly imply weak dominance of truth-telling because robustness is imposed only with respect to positively correlated beliefs.

Our main result in this section is:

THEOREM **3** *Assume $m \geq 3$. A SCF is efficient and K-ROBIC if and only if it is dictatorial.*

The proof of the Theorem is complicated and relegated to the Appendix. Later in the section we provide a sketch of the proof. Before doing so, we state and prove an auxiliary result which we use repeatedly in the proof. We believe that the result is also of some independent interest because it illuminates the restrictions that the K-ROBIC assumption imposes on a SCF.

[6] A SCF is anonymous if it does not discriminate amongst voters.

Proposition 3 *Let f be a K-ROBIC SCF. Let (P_i, P_{-i}) and (P_i', P_{-i}) be two profiles such that $f(P_i', P_{-i})P_i f(P_i, P_{-i})$. Then there exists P_{-i}' satisfying the following properties:*
(i) for all $j \neq i$, $d(P_i, P_j') < d(P_i, P_j)$
(ii) either $f(P_i, P_{-i}') = f(P_i', P_{-i})$ or $f(P_i, P_{-i}')P_i f(P_i', P_{-i})$
(iii) either $f(P_i, P_{-i}) = f(P_i', P_{-i}')$ or $f(P_i, P_{-i})P_i f(P_i', P_{-i}')$

Proof: If f is K-ROBIC, it must also be K-LOBIC. From Theorem 1, it follows that f must satisfy OND. Suppose that there exist profiles (P_i, P_{-i}) and (P_i', P_{-i}) such that $f(P_i', P_{-i})P_i f(P_i, P_{-i})$. Then OND implies that there exists P_{-i}' satisfying (ii) and (iii). It only remains to show (i).

Let $L = \{P_{-i}' | f(P_i, P_{-i}')P_i f(P_i', P_{-i}')\}$. Suppose that for all $P_{-i}' \in L$, there exists a $j \neq i$ such that $d(P_i, P_j') \geq d(P_i, P_j)$. Note that for any $\delta_1, \delta_2 > 0$ we can always choose a utility function u representing P_i such that $u(f(P_i', P_{-i})) - u(f(P_i, P_{-i})) = \delta_1$ and

$$\Big| \max_{P_{-i}' \in L} [u(f(P_i, P_{-i}')) - u(f(P_i', P_{-i}'))] - [u(f(P_i', P_{-i})) - u(f(P_i, P_{-i}))] \Big| < \delta_2$$

Also for any ϵ_1, ϵ_2 such that $1 > \epsilon_1 > \epsilon_2 > 0$, there exists a belief system $(\mu_1, \cdots, \mu_n) \in K^*$ such that (i) $\mu(P_{-i}'|P_i) > \epsilon_1$ if either (a) $P_{-i}' = P_{-i}$ or (b) for all $j \neq i$, $d(P_i, P_j') < d(P_i, P_j)$ and (ii) $\mu(P_{-i}'|P_i) < \epsilon_2$ for all other P_{-i}'. Let $\Delta = \sum_{P_{-i}' \in \mathbb{P}^{n-1}} [u(f(P_i', P_{-i}')) - u(f(P_i, P_{-i}'))]\mu(P_{-i}'|P_i)$. It follows that$\Delta \geq \epsilon_1\delta_1 - (m! - 1)^{N-1}(\delta_1 + \delta_2)\epsilon_2$. It is clear that by choosing ϵ_2 sufficiently close to zero, the R.H.S of the inequality above can be made strictly positive, i.e. $\Delta > 0$. But this violates the assumption that f is K-ROBIC. ■

The extra strengthening of OND for K-ROBIC is natural. As we have discussed earlier, OND (parts (ii) and (iii) above) ensures that the gain from truthful reporting at P_i instead of P_i' at P_{-i}' is "ordinally" greater than the loss from truthful reporting at P_{-i}. In addition, the Kemeny distance between the P_i and P_{-i}' must be strictly smaller than that between P_i and P_{-i}. If this were not true, the expected utility from lying could be made to exceed that of truth-telling by choosing a conditional probability distribution such that $\mu_i(P_{-i}'|P_i)$ is made arbitrarily small relative to $\mu_i(P_{-i}|P_i)$.

Observation 6 The condition described in the statement of Proposition 3 is *not* sufficient for a SCF to be K-ROBIC. A further condition is required which as follows. For all P_i and P_i', let $G = \{P_{-i} | f(P_i', P_{-i})P_i f(P_i, P_{-i})\}$ and let $L = \{P_{-i}' | f(P_i, P_{-i}')P_i f(P_i', P_{-i}')\}$. According to Proposition 3, there exists a map $\sigma : G \to L$ such that $P_{-i} \in G$ there exists a $P_{-i}' \in L$ satisfying conditions (i), (ii) and (iii). The additional requirement which is necessary and together with Proposition 3 is also sufficient, is that the map σ must be *injective*. We do not include a proof of this result in the paper because it is not required for the proof of Theorem 3.

We now provide the sketch of the proof of Theorem 3.

SKETCH OF PROOF: For any profile P we define $D(P)$ to be the maximal Kemeny distance between any two preference orderings in the profile P. The result is proved by an induction on $D(P)$. The proof consists of two claims. In Claim 1 we show that there exists voter i such that for all profiles, P with $D(P) = 1$, we have $f(P) = r_1(P_i)$. There are three steps to prove Claim 1. In Step 1 we consider a subclass of profiles where all voters other than say voter k have the same preference ordering that is different from voter k's. We denote such profiles by $\mathbb{D}(k)$. Let us consider a profile $\tilde{P} \in \mathbb{D}(k)$ where voter k has preference $\tilde{P}_j$ and all other voters have preference $\tilde{P}_i$. By Pareto efficiency $f(\tilde{P}) \in \{r_1(\tilde{P}_j), r_1(\tilde{P}_i)\}$. In Step 1 we show that if in such a profile $\tilde{P} \in \mathbb{D}(k)$, the social choice function picks the common top-ranked alternative of all voters other than k, $r_1(\tilde{P}_i)$ as the outcome, then for all profiles $P \in \mathbb{D}(k)$, $f(P) = r_1(P_i)$ where P_i is the common preference ordering of all voters other than k in the profile P. By varying $k = 1, \cdots, n$ we get the collections of profiles $\mathbb{D}(1), \mathbb{D}(2), \cdots, \mathbb{D}(n)$. In any profile $P \in \mathbb{D}(1)$ for example, voter 1 has preference ordering P_j and all voters other than 1 has a common preference ordering P_i. Since k is arbitrarily chosen, Step 1 holds for all such collection of profiles, $\mathbb{D}(1), \mathbb{D}(2), \cdots, \mathbb{D}(n)$. In Step 2 we show that there exists a k and profile $P^k = (P_i, \cdots, P_i, P_j, P_i, \cdots, P_i) \in \mathbb{D}(k)$ such that $f(P^k) = r_1(P_j)$. We can describe Step 2 as "identifying the dictator step". Step 3 completes Claim 1 by showing that the voter identified as the dictator in Step 2 indeed dictates over all profiles P with $D(P) = 1$.

We then move to Claim 2, where we use induction on $D(P)$ to complete the proof. Specifically, we assume without loss of generality that k is an integer greater than 1 and that $f(P') = r_1(P'_1)$ whenever $D(P') \leq k$. Let P be a profile such that $D(P) = k + 1$. In Claim 2 we show that $f(P) = r_1(P_1)$. The proof of Claim 2 is divided in two parts. First we consider profiles where voter 1 has preference ordering P_1 and all other voters have common preference ordering P_i and $d(P_1, P_i) = k + 1$. Let P^1 be such a profile. In the first part we show that $f(P^1) = r_1(P_1)$. In the next part we show that if $f(P^1) = r_1(P_1)$, then for all P such that $D(P) = k + 1$, $f(P) = r_1(P_1)$. we show this by sequentially changing the preference orderings of all voters other than 1 and ensuring that the social choice outcome remains unchanged in all these profiles.

An obvious implication of Theorem 3 is the following result:

COROLLARY **1** *Assume $m \geq 3$. A SCF is efficient and TS-ROBIC if and only if it is dictatorial.*

We have seen that efficiency cannot be weakened to the assumption of unanimity because the unanimity with status quo rule clearly satisfies unanimity. However, are there other K-ROBIC SCFs which satisfy unanimity? We have an answer in the special case where $m = 3$ but not to the general question.

Proposition 4 *Assume $m = 3$. A SCF is K-ROBIC and satisfies unanimity if and only if it is either dictatorial or the unanimity with status quo rule.*

The proof of this result is omitted. It is available from the authors on request.

7 Conclusion

In this paper we have explored the problem of mechanism design in a voting environment with an arbitrary number of voters where a voter's belief about the type of the other voters are positively correlated with her own type. Our general conclusion is that the prospects for constructing incentive-compatible social choice functions in this environment are significantly improved relative to the independent case. In this respect, our results parallel those in environments with transfers and quasi-linear utility such as Crémer and Mclean (1988). However the reasons behind the enhanced possibilities in the voting environment are quite different from the quasi-linear context.

In future research we hope to extend our analysis to other notions of correlation and to understand better, the relationship between the structure of beliefs and incentive-compatible social choice functions.

References

Baigent, N. (1987): "Preference Proximity and Anonymous Social Choice," *Quarterly Journal of Economics*, 102, 161–169.

Bossert, W. and T. Storcken (1992): "Strategy-proofness of Social Welfare Functions: The Use of The Kemeny Distance Between Preference Orderings," *Social Choice and Welfare*, 9, 345–360.

Crémer, J. and R. Mclean (1988): "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56, 1247–1257.

d'Aspremont, C. and B. Peleg (1988): "Ordinal Bayesian Incentive Compatible Representation of Committees," *Social Choice and Welfare*, 5, 261–280.

Gibbard, A. (1973): "The Manipulation of Voting Schemes: A General Result," *Econometrica*, 41, 587–601.

Kemeny, J. and J. Snell (1962): *Mathematical Methods in Social Sciences*, Ginn, New York.

Kendall, M. (1970): *Rank Correlation Methods*, Griffin, London.

Majumdar, D. and A. Sen (2004): "Ordinal Bayesian Incentive Compatible Voting Rules," *Econometrica*, 72, 523–540.

Maskin, E. (1999): "Nash Equilibrium and Welfare Optimality," *Review of Economic Studies*, 66, 39–56.

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.

Wilson, R. (1987): *Advances in Econometric Theory: Fifth World Congress*, Cambridge University Press, chap. Game Theoretic Approaches to Trading Processes, 33–77, editor: T. Bewley.

8 Appendix

We provide a proof of Theorem 3 below.

Proof: The sufficiency part of the theorem is trivial since dictatorial SCFs are strategy-proof and efficient. We shall therefore only prove necessity. In what follows, we assume that f is efficient and K-ROBIC.

For any profile P, let $D(P) = \max_{i,j} d(P_i, P_j)$. Thus $D(P)$ is the maximal Kemeny distance between any two preference orderings in the profile P. We shall prove the result by induction on $D(P)$. In particular we shall prove the following claims.

CLAIM 1: There exists a voter i such that for all profiles P such that $D(P) = 1$, we have $f(P) = r_1(P_i)$.

CLAIM 2: Let k be an integer with $k > 1$. Suppose that there exists a voter i such that for all profiles P' with $D(P') \leq k$, we have $f(P') = r_1(P'_i)$. Let P be a profile such that $D(P) = k + 1$. Then, $f(P) = r_1(P_i)$.

It is evident that Claims 1 and 2 establish that f is dictatorial and voter i is the dictator. We begin our proof of Claim 1 with two useful Lemmas.

Lemma 1 *Let P_i and P'_i be orderings for voter i such that $B(r_t(P_i), P_i) = B(r_t(P'_i), P'_i)$ for some integer t lying between 1 and $m - 1$. Let P_{-i} be a profile of orderings for voters other than i such that $D(P_i, P_{-i}) = 2$. Then*

$$[f(P'_i, P_{-i}) \in B(r_t(P'_i), P'_i)] \Rightarrow [f(P_i, P_{-i}) \in B(r_t(P_i), P_i)]$$

Proof: Suppose that the Lemma is false. Let $f(P_i', P_{-i}) = y$ and $f(P_i, P_{-i}) = x \notin B(r_t(P_i), P_i)$. Following Proposition 3 there exists a P'_{-i} such that,

1. $d(P_i, P_j') < d(P_i, P_j)$ for all $j \neq i$
2. $f(P_i, P'_{-i}) \in B(y, P_i)$ and
3. $f(P_i', P'_{-i}) \in W(x, P_i)$. [7]

Observe that since $B(r_t(P_i), P_i) = B(r_t(P_i'), P_i')$, $W(x, P_i) \cap B(r_t(P_i'), P_i') = \emptyset$. Since $D(P_i, P'_{-i}) = 1$, for any $j \neq i$, either,

(a.) $r_1(P_i) = r_1(P_j')$ or

(b.) $r_1(P_j') = r_2(P_i)$, $r_2(P_j') = r_1(P_i)$ and for all $l \notin \{1, 2\}$, $r_l(P_i) = r_l(P_j')$.

Consider any $z \in W(x, P_i)$. Since, $B(r_t(P_i), P_i) = B(r_t(P_i'), P_i')$, (a) and (b) implies that, there exists a $w \in B(r_t(P_i), P_i)$, such that for all $j \in N$, $wP_j'z$. But then, $z \notin PE(P_i', P'_{-i})$. Therefore, Pareto-efficiency of f implies that $f(P_i', P'_{-i}) \neq z$. Since the above statement is true for any $z \in W(x, P_i)$, part 3. in the above requirement is violated. ■

Lemma 2 *Fix an integer $t \in \{1, \cdots, m-1\}$. Consider a profile (P_i, P_{-i}) such that*

1. *there exists a $j \neq i$ such that $B(r_t(P_i), P_i) = B(r_t(P_j), P_j)$ and*
2. *$f(P_i, P_{-i}) \in B(r_t(P_i), P_i)$*

Then, $f(P_j, P_{-i}) \in B(r_t(P_j), P_j)$

Proof: Suppose the claim is false. Then voter i with preference ordering P_j gains by misreporting ordering P_i at the profile (P_j, P_{-i}) . However in the preference profile (P_{-i}) for the voters other than i, there exists voter j with preference ordering P_j. Since $d(P_j, P_j) = 0$, the condition in Proposition 3 will be violated. ■

Proof of Claim 1: We prove Claim 1 in three steps. Some of the steps have several substeps.

Proof: Let $\mathbb{H}$ denote the set of pairs of orderings P_i, P_j satisfying the following conditions.

1. $r_1(P_i) = r_2(P_j)$

[7] The set $W(x, P_i)$ is the set of alternatives that are weakly worse than x according to P_i, i.e. $W(x, P_i) = \{z \in A | xP_iz\} \cup \{x\}$.

2. $r_2(P_i) = r_1(P_j)$

3. $r_t(P_1) = r_t(P_2)$ for all $t > 2$

In other words, $\mathbb{H}$ consists of pairs of orderings which agree on the rankings of all alternatives except those which are ranked first and second. Let k be an integer lying between 1 and n. Then $\mathbb{D}(k)$ is the set of preference profiles where voter k has ordering P_j and all voters other than k have the ordering P_i and $(P_i, P_j) \in \mathbb{H}$. A typical element of $\mathbb{D}(k)$ will be denoted by P^k where

$$P^k \equiv (\underbrace{P_i, \cdots, P_i}_{k-1}, P_j, \underbrace{P_i, \cdots, P_i}_{N-k})$$

and $(P_i, P_j) \in \mathbb{H}$.

Step 1: We will show the following. Fix an integer k and let $\tilde{P}^k \in \mathbb{D}(k)$ be the profile where voter k has the ordering $\tilde{P}_j$ and all other voters have the ordering $\tilde{P}_i$. Suppose $f(\tilde{P}^k) = r_1(\tilde{P}_i)$. Then $f(P^k) = r_1(P_i)$ for all $P^k \in \mathbb{D}(k)$. Here P^k is the profile where voter k has the ordering P_j and all other voters have the ordering P_i and $(P_i, P_j) \in \mathbb{H}$. Without loss of generality let $r_1(\tilde{P}_i) = b$ and $r_1(\tilde{P}_j) = a$. Since f is efficient $f(\tilde{P}^k)$ is either a or b. Under our assumption $f(\tilde{P}^k) = b$. Without loss of generality let $k = 1$. That is voter 1 has the preference P_j and all the other voters have the preference P_i. In order to simplify the notation (so that there is no confusion) we will set $P_j = P_1$ and $P_i = P_2$. In other words, we say voter 1 has the preference ordering P_1 and all voters other than 1 has the preference ordering P_2. This is just to simplify the notation and does not, in any way, affect the analysis that follows. We will prove Step 1 in a series of sub-steps. Step 1(i): Let $\{c, d\}$ be an arbitrary pair of alternatives and let $P'^1, \hat{P}^1 \in \mathbb{D}(1)$ be such that $r_1(P'_1) = r_1(\hat{P}_1) = c$ and $r_1(P'_2) = r_1(\hat{P}_2) = d$. We show that if $f(P'^1) = d$, then $f(\hat{P}^1) = d$.

Suppose $f(P'^1) = d$. Consider the case where $d(P'_1, \hat{P}_1) = d(P'_2, \hat{P}_2) = 1$. In other words, $\hat{P}_1$ and $\hat{P}_2$ are obtained from P'_1 and P'_2 respectively by the transposition of some common pair $\{x, y\}$ of alternatives. Consider the profile$(\hat{P}_1, P'_2, \cdots, P'_2)$. Suppose $f(\hat{P}_1, P'_2, \cdots, P'_2) \neq d$. Then Pareto efficiency implies $f(\hat{P}_1, P'_2, \cdots, P'_2) = c$. Since cP'_1d, Proposition 3 implies that there exists a $\tilde{P}_{-1}$ such that,

(i.) for all $j \neq 1$, $d(P'_1, \tilde{P}_j) < d(P'_1, P'_2)$

(ii.) $f(P'_1, \tilde{P}_{-1}) = c$ and

(iii.) $f(\hat{P}_1, \tilde{P}_{-1}) = d$ or $f(\hat{P}_1, \tilde{P}_{-1}) = x$ where dP'_1x.

Note that $d(P'_1, P'_2) = 1$. Therefore the only candidate $\tilde{P}_{-1}$ is where for all $j \neq 1$, $d(P'_1, \tilde{P}_j) = 0$. In other words, for all $j \neq 1$, $\tilde{P}_j = P'_1$. Since $r_1(P'_1) = r_1(\hat{P}_1) = c$,

$f(P_1', \tilde{P}_{-1}) = c \Rightarrow f(\hat{P}_1, \tilde{P}_{-1}) = c$. Hence conditions (ii.) and (iii.) above are not simultaneously satisfied. Hence $f(\hat{P}_1, P_2', \cdots, P_2') = d$. Consider now the profile $(\hat{P}_1, \hat{P}_2, P_2', \cdots, P_2')$. In other words, in the profile $(\hat{P}_1, \hat{P}_2, P_2', \cdots, P_2')$, voter 1 has preference $\hat{P}_1$, voter 2 has preference $\hat{P}_2$ and all other voters have the preference ordering P_2'. By Pareto efficiency, $f(\hat{P}_1, \hat{P}_2, P_2', \cdots, P_2') \in \{c, d\}$. Suppose $f(\hat{P}_1, \hat{P}_2, P_2', \cdots, P_2') \neq d$. Then the previous statement implies $f(\hat{P}_1, \hat{P}_2, P_2', \cdots, P_2') = c$. Since $d\hat{P}_2c$, Proposition 3 implies that there exists a $\tilde{P}_{-2}$ such that,

(i.) $d(\hat{P}_2, \tilde{P}_1) < d(\hat{P}_2, \hat{P}_1)$

(ii.) for all $j \neq 1, 2$, $d(\hat{P}_2, \tilde{P}_j) < d(\hat{P}_2, P_2')$

(iii.) $f(\hat{P}_2, \tilde{P}_{-2}) = d$ and

(iv.) $f(P_2', \tilde{P}_{-2}) = c$ or $f(P_2', \tilde{P}_{-2}) = x$ where $c\hat{P}_2x$

Now $d(\hat{P}_2, \hat{P}_1) = 1$ and for all $j \neq 1, 2$, $d(\hat{P}_2, P_2') = 1$. This implies the only candidate $\tilde{P}_{-2}$ is one where for all $j \neq 2$, $\tilde{P}_j = \hat{P}_2$. Since $\hat{P}_2$ and P_2' have the same top-ranked element d, $[f(\hat{P}_2, \tilde{P}_{-2}) = d] \Rightarrow [f(P_2', \tilde{P}_{-2}) = d]$. Hence conditions (iii.) and (iv.) cannot hold together. Therefore, $f(\hat{P}_1, \hat{P}_2, P_2', \cdots, P_2') = d$. We now consider the profile $(\hat{P}^1)$. Observe that since $r_1(P_1') = r_1(\hat{P}_1) = c$, $r_1(P_2') = r_1(\hat{P}_2) = d$ and $P'^1, \hat{P}^1 \in \mathbb{D}(1)$, it has to be the case that $r_1(\hat{P}_1) = r_2(\hat{P}_2) = c$ and $r_2(\hat{P}_1) = r_1(\hat{P}_2) = d$. We can find a sequence of profiles $P(r)$, $r = 2, \cdots, N$ such that (i) $P(2) = (\hat{P}_1, \hat{P}_2, P_2', \cdots, P_2')$, (ii) $P(N) = \hat{P}^1$ and (iii) $P(r) = (\hat{P}_1, \hat{P}_2, \cdots, \hat{P}_2, P_2', \cdots, P_2')$, $r = 3, \cdots, N$. In other words, the profile $\hat{P}^1$ can be obtained from the profile $(\hat{P}_1\hat{P}_2, P_2', \cdots, P_2')$ by sequentially changing the preference of one voter at a time from P_2' to $\hat{P}_2$. The profile $P(r)$ is the one where voter 1 has preference $\hat{P}_1$, voters 2 to r have preference $\hat{P}_2$ and all other voters have preference P_2'. Using the arguments in the precious two paragraph we can conclude that $f(P(r)) = d$ implies $f(P(r+1)) = d$, $r = 2, \cdots, N-1$. Hence $f(\hat{P}^1) = d$ which establishes Step $1(i)$.

Since we will be looking at profiles where all voters other than voter 1 has the same preference ordering, we will abuse notation slightly and denote this block of voters other than 1 (the 'coalesced voter') by $\bar{\mathbf{1}}$. Let $\{c, d\}$ be an arbitrary ordered pair of alternatives. We will say that a voter $i \in \{1, \bar{\mathbf{1}}\}$ dictates over $\{c, d\}$ if for all $P^1 = (P_1, P_2, \cdots, P_2)$ such that (i) $P^1 \in \mathbb{D}(1)$, (ii)$r_1(P_1) = r_2(P_2) = c$ and (iii) $r_2(P_1) = r_1(P_2) = d$, we have $f(P^1) = r_1(P_i)$. In view of Step $1(i)$ some voter $i \in \{1, \bar{\mathbf{1}}\}$ dictates over each pair of alternatives. In particular we can infer that voter $\bar{\mathbf{1}}$ dictates over $\{a, b\}$.

Step $1(ii)$: Let c be an alternative distinct from a and b. Then voter $\bar{\mathbf{1}}$ dictates over $\{a, c\}$.

Let $\bar{\mathbb{D}}$ be the set of preference orderings where the top three alternatives belong to the set $\{a, b, c\}$ while the rankings of all other alternatives are fixed. Formally $P_i \in \bar{\mathbb{D}}$ if

1. $\cup_{\{k=1,2,3\}} r_k(P_i) = \{a, b, c\}$

2. for all $d \neq a, b, c$, there exists an integer $q \geq 4$, such that $d = r_q(P_i)$. Moreover q does not depend on P_i.

For notational convenience, we will denote elements of $\bar{\mathbb{D}}$ by $abc\ldots, acb\ldots, bac\ldots, bca\ldots, cab\ldots$ and $cba\ldots$. Here $abc\ldots$ denotes the ordering where a, b and c are ranked first, second and third respectively. In view of Step 1(i), Step 1(ii) is complete if we can show that

$$f(P^1) = f\begin{pmatrix} a & c & c\cdots c \\ c & a & a\cdots a \\ b & b & b\cdots b \\ . & . & .\cdots . \end{pmatrix} = c$$

We proceed in a sequence of sub-steps. Let us consider the profile,

$$\hat{P}^1 = \begin{pmatrix} a & b & b\cdots & b \\ b & a & a\cdots & a \\ c & c & c\cdots & c \\ . & . & .\cdots . & \end{pmatrix}$$

Step 1(ii)a: $f(acb..., \hat{P}^1_{-1}) = b$. Since $\bar{\mathbf{1}}$ dictates over $\{a, b\}$ by assumption, we have $f(\hat{P}^1) = b$. We have also shown in Step 1(i) that a transposition in voter 1's ordering which does not involve her first-ranked alternative in this profile, does not change the outcome. Hence $f(acb..., \hat{P}_{-1}) = b$.
Step 1(ii)b: $f(\hat{P}^1_{-2}, bca...) = b$. Observe that $f(\hat{P}^1) = b$. Since the top-ranked element of $bac...$ and $bca...$ are the same and $D(P^1_{-2}, bca...) = 2$, using Lemma 1 we get the desired result. Now repeatedly using Lemma 1 we get that,

$$f(\tilde{P}) = f\begin{pmatrix} a & b & b\cdots b \\ b & c & c\cdots c \\ c & a & a\cdots a \\ . & . & .\cdots . \end{pmatrix} = b \tag{31}$$

Step 1(ii)c: From equation 31 we claim that,

$$f(\bar{P}) = f\begin{pmatrix} a & b & b\cdots b \\ c & c & c\cdots c \\ b & a & a\cdots a \\ . & . & .\cdots . \end{pmatrix} \in \{b, c\} \tag{32}$$

Suppose equation 32 is false. Since f is efficient the only possibility is that,

$$f\begin{pmatrix} a & b & b\cdots b \\ c & c & c\cdots c \\ b & a & a\cdots a \\ . & . & .\cdots . \end{pmatrix} = a$$

Then voter 1 in the profile $\tilde{P}$ gains by reporting $acb...$ (we know from the previous step that $f(\tilde{P}) = b$). Since $D(\tilde{P}) = 2$ and the top ranked elements of $abc...$ and $acb...$ are the same, Lemma 1 implies that

$$f\begin{pmatrix} a & b & b\cdots b \\ c & c & c\cdots c \\ b & a & a\cdots a \\ . & . & .\cdots . \end{pmatrix} \neq a$$

Hence the claim.

Step 1(ii)d: Observe that the two top-ranked elements in $cba...$ and $bca...$ are the same. Moreover for any $j \in N$, $d(\bar{P}_j, cba...) \leq 2$. From equation 32 and using Lemma 1 we get that

$$f\begin{pmatrix} a & b & c\cdots b \\ c & c & b\cdots c \\ b & a & a\cdots a \\ . & . & .\cdots . \end{pmatrix} \in \{b, c\} \tag{33}$$

Repeatedly using the above argument, we obtain,

$$f(P^\star) = f\begin{pmatrix} a & b & c\cdots c \\ c & c & b\cdots b \\ b & a & a\cdots a \\ . & . & .\cdots . \end{pmatrix} \in \{b, c\} \tag{34}$$

Let us now consider the profile

$$P^1 = \begin{pmatrix} a & c & c\cdots c \\ c & a & a\cdots a \\ b & b & b\cdots b \\ . & . & .\cdots . \end{pmatrix}$$

By Pareto efficiency, $f(P^1) \in \{a, c\}$. If $f(P^1) = a$ then $f(P^1_{-2}, cba...) = a$. The argument to establish this step is identical to the one in Step 1(i) and Step 1(ii)a. Repeatedly using the same argument we get,

$$f(\breve{P}) = f\begin{pmatrix} a & c & c\cdots c \\ c & b & b\cdots b \\ b & a & a\cdots a \\ . & . & .\cdots . \end{pmatrix} = a \tag{35}$$

Step 1(ii)e: $f(\breve{P}) \neq a$. Since $f(P^\star) \in \{b, c\}$ and $(\breve{P}_{-2}, bca...) = (P^\star)$, voter 2 in the profile $\breve{P}$ gains by reporting $bca...$ instead of $cba...$. Observe that $D(\breve{P}) = 2$. Since the two top-ranked alternatives in *cba*... and *bca*... are the same, we apply Lemma 1 to arrive at the conclusion.

We now establish Step 1(ii). Efficiency and Steps 1$(ii)d$ and 1$(ii)e$ imply $f(P^1) = c$. Hence voter $\bar{\mathbf{1}}$ dictates over $\{a, c\}$.

STEP 1(iii): Let c be an alternative distinct from a and b. Then voter $\bar{\mathbf{1}}$ dictates over $\{c, b\}$. In view of our earlier arguments it is suffice to show that,

$$f(P^1) = f\begin{pmatrix} c & b & b\cdots b \\ b & c & c\cdots c \\ a & a & a\cdots a \\ . & . & .\cdots . \end{pmatrix} = b \tag{36}$$

Suppose equation 36 is false. By efficiency $f(P^1) = c$. By replicating the arguments of Steps 1(i), 1(ii) and 1(iii) with the roles of voters and alternatives interchanged, we can conclude that,

$$f\begin{pmatrix} c & b & b\cdots b \\ b & a & c\cdots c \\ a & c & a\cdots a \\ . & . & .\cdots . \end{pmatrix} = f\begin{pmatrix} c & b & b\cdots b \\ b & c & c\cdots c \\ a & a & a\cdots a \\ . & . & .\cdots . \end{pmatrix} = b \tag{37}$$

Using Lemma 1 we get from equation 37 that,

$$f\begin{pmatrix} c & b & b\cdots b \\ b & a & c\cdots c \\ a & c & a\cdots a \\ . & . & .\cdots . \end{pmatrix} \in \{a, c\} \tag{38}$$

Consider now the profile

$$\hat{P} = \begin{pmatrix} c & b & b\cdots b \\ b & a & a\cdots a \\ a & c & c\cdots c \\ . & . & .\cdots . \end{pmatrix}$$

Repeatedly using Lemma 2 we get, from equation 38 that,

$$f(\hat{P}) = f\begin{pmatrix} c & b & b\cdots b \\ b & a & a\cdots a \\ a & c & c\cdots c \\ . & . & .\cdots . \end{pmatrix} \in \{a, c\} \tag{39}$$

We have already established in Step 2(i) that $f(acb, \hat{P}_{-1}) = b$. Therefore voter 1 gains in profile $(acb..., \hat{P}_{-1})$ by reporting *cab*... instead of *acb*.... Observe that $D(acb..., \hat{P}_{-1}) = 2$.

Moreover, the two top-ranked elements in *acb*... and *cab*... are the same. Hence by applying Lemma 1 we arrive at a contradiction. Therefore, voter $\bar{\mathbf{1}}$ dictates over $\{c, b\}$.

Step 1(iv): Let c, d be a pair of alternatives such that a, b, c, d are all distinct. Then voter $\bar{\mathbf{1}}$ dictates over the pair $\{c, d\}$. In order to verify this claim, note that Step 1(ii) implies that voter $\bar{\mathbf{1}}$ dictates over $\{a, d\}$. Now applying Step 1(iii), we conclude that voter $\bar{\mathbf{1}}$ dictates over $\{c, d\}$

Step 1(v): Voter $\bar{\mathbf{1}}$ dictates over the pair $\{b, a\}$. Let $c \neq a, b$ (we are using the assumption that $|A| \geq 3$). According to Step 1(ii), voter $\bar{\mathbf{1}}$ dictates over $\{a, c\}$. Applying Step 1(iii), voter $\bar{\mathbf{1}}$ dictates over $\{b, c\}$ and applying Step 1(ii) again, we conclude that voter $\bar{\mathbf{1}}$ dictates over $\{b, a\}$.

Step 2: We show that there exists voter k and $P^k \in \mathbb{D}(k)$ (where voter k's ordering is P_j and all other voters have ordering P_i) such that $f(P^k) = r_1(P_j)$.

Proof of Step 2: Without loss of generality, let $P_i = bac...$ and $P_j = abc....$ Then the claim made in Step 2 implies the following:

$$f\begin{pmatrix} a & b & b\cdots b \\ b & a & a\cdots a \\ c & c & c\cdots c \\ . & . & .\cdots . \end{pmatrix} = f\begin{pmatrix} b & a & b\cdots b \\ a & b & a\cdots a \\ c & c & c\cdots c \\ . & . & .\cdots . \end{pmatrix} = f\begin{pmatrix} b & b & a\cdots b \\ a & a & b\cdots a \\ c & c & c\cdots c \\ . & . & .\cdots . \end{pmatrix} = \cdots = f\begin{pmatrix} b & b & b\cdots a \\ a & a & a\cdots b \\ c & c & c\cdots c \\ . & . & .\cdots . \end{pmatrix} \neq b \tag{40}$$

Suppose that the claim is not true.

Step 2(i): Consider now the preference orderings $P_i = cba...$ and $P_j = bca....$ Then,

$$f(P^2) = f\begin{pmatrix} c & b & c\cdots c \\ b & c & b\cdots b \\ a & a & a\cdots a \\ . & . & .\cdots . \end{pmatrix} = c \tag{41}$$

Consider now the preference ordering $cab...$ for voter 1. From equation 41 we claim that,

$$f\begin{pmatrix} c & b & c\cdots c \\ a & c & b\cdots b \\ b & a & a\cdots a \\ . & . & .\cdots . \end{pmatrix} = c \tag{42}$$

Otherwise voter 1 of type $cab...$ is losing to type $cba...$ on a profile P with $D(P) = 2$. Since $B(r_1(cab), cab) = B(r_1(cba), cba)$ the result follows using Lemma 1 . By repeatedly making use of the above argument we can show that,

$$f\begin{pmatrix} c & b & c\cdots c \\ a & c & a\cdots a \\ b & a & b\cdots b \\ . & . & .\cdots . \end{pmatrix} = c \tag{43}$$

Consider now the preference ordering $acb...$ for voter 3. Since, $B(r_2(acb...), acb...) = B(r_2(cab...), cab...) = \{a, c\}$ and

$$f\begin{pmatrix} c & b & c\cdots c \\ a & c & a\cdots a \\ b & a & b\cdots b \\ . & . & .\cdots . \end{pmatrix} \in \{a, c\}$$

we have,

$$f\begin{pmatrix} c & b & a\cdots c \\ a & c & c\cdots a \\ b & a & b\cdots b \\ . & . & .\cdots . \end{pmatrix} \in \{a, c\} \tag{44}$$

Let us now consider preference orderings $P_i = cab...$ and $P_j = acb....$ Then,

$$f(P^3) = f\begin{pmatrix} c & c & a\cdots c \\ a & a & c\cdots a \\ b & b & b\cdots b \\ . & . & .\cdots . \end{pmatrix} = c \tag{45}$$

Using an argument similar to above, but now changing the preference ordering for voter 2 to $bca...$ we get

$$f\begin{pmatrix} c & b & a\cdots c \\ a & c & c\cdots a \\ b & a & b\cdots b \\ . & . & .\cdots . \end{pmatrix} \in \{b, c\} \tag{46}$$

Combining equations 45 and 46 we get that,

$$f\begin{pmatrix} c & b & a\cdots c \\ a & c & c\cdots a \\ b & a & b\cdots b \\ . & . & .\cdots . \end{pmatrix} = c \tag{47}$$

Using similar logic, but changing the identity of the voter with preference ordering *acb...* we get,

$$f\begin{pmatrix} c & b & a\cdots c \\ a & c & c\cdots a \\ b & a & b\cdots b \\ . & . & .\cdots . \end{pmatrix} = f\begin{pmatrix} c & b & c & a\cdots c \\ a & c & a & c\cdots a \\ b & a & b & b\cdots b \\ . & . & . & .\cdots . \end{pmatrix} = \cdots = f\begin{pmatrix} c & b & c\cdots a\cdots c \\ a & c & a\cdots c\cdots a \\ b & a & b\cdots b\cdots b \\ . & . & .\cdots .\cdots . \end{pmatrix} =$$

$$f\begin{pmatrix} c & b & c\cdots a \\ a & c & a\cdots c \\ b & a & b\cdots b \\ . & . & .\cdots . \end{pmatrix} = c \tag{48}$$

Since we have assumed that the claim in Step 2 is false, we have,

$$f(P^1) = f\begin{pmatrix} a & b & b\cdots b \\ b & a & a\cdots a \\ c & c & c\cdots c \\ . & . & .\cdots . \end{pmatrix} = b \tag{49}$$

In Step 1 we have already established that if $f(P^1) = b$ (as in equation 49) then, following the arguments in Step 1, we have,

$$f\begin{pmatrix} a & b & b\cdots b \\ c & a & a\cdots a \\ b & c & c\cdots c \\ . & . & .\cdots . \end{pmatrix} = b \tag{50}$$

Consider now the preference ordering *cab...* for voter 1. Observe that, $B(r_2(acb...), acb...) = B(r_2(cab...), cab...)$. Then from equation 50 and Lemma 1 it follows that,

$$f\begin{pmatrix} c & b & b\cdots b \\ a & a & a\cdots a \\ b & c & c\cdots c \\ . & . & .\cdots . \end{pmatrix} \notin \{a, c\} \tag{51}$$

Imposing Pareto efficiency we get from equation 51 that,

$$f\begin{pmatrix} c & b & b\cdots b \\ a & a & a\cdots a \\ b & c & c\cdots c \\ . & . & .\cdots . \end{pmatrix} = b \tag{52}$$

Now take any voter $i \in \{3, \cdots, n\}$. Consider the preference ordering $abc...$ for agent i. Applying Lemma 1 to equation 52 we get that,

$$f\begin{pmatrix} c & b & b\cdots a\cdots b \\ a & a & a\cdots b\cdots a \\ b & c & c\cdots c\cdots c \\ . & . & .\cdots .\cdots . \end{pmatrix} \in \{a, b\} \tag{53}$$

By varying i from 3 to n we actually get $n-2$ such conditions. Without loss of generality take $i = 3$. Then equation 53 implies,

$$f\begin{pmatrix} c & b & a\cdots b \\ a & a & b\cdots a \\ b & c & c\cdots c \\ . & . & .\cdots . \end{pmatrix} \in \{a, b\} \tag{54}$$

Repeating the argument above we can conclude that,

$$f(\hat{P}) = f\begin{pmatrix} c & b & a & b & a\cdots a \\ a & a & b & a & b\cdots b \\ b & c & c & c & c\cdots c \\ . & . & . & . & .\cdots . \end{pmatrix} \in \{a, b\} \tag{55}$$

There are two cases to consider:

Case I: Let,

$$f(\hat{P}) = f\begin{pmatrix} c & b & a & b & a\cdots a \\ a & a & b & a & b\cdots b \\ b & c & c & c & c\cdots c \\ . & . & . & . & .\cdots . \end{pmatrix} = b \tag{56}$$

Consider now a preference ordering $acb...$ for voter 1. Using an argument similar to the on made before equation 51 we can say that

$$f\begin{pmatrix} a & b & a & b & a\cdots a \\ c & a & b & a & b\cdots b \\ b & c & c & c & c\cdots c \\ . & . & . & . & .\cdots . \end{pmatrix} = b \tag{57}$$

Observe that if we now change voter 3's preference ordering to $acb...$ ROBIC and Pareto efficiency would imply that

$$f\begin{pmatrix} a & b & a & b & a\cdots a \\ c & a & c & a & b\cdots b \\ b & c & b & c & c\cdots c \\ . & . & . & . & .\cdots . \end{pmatrix} = b \tag{58}$$

Repeating the argument above we get,

$$f\begin{pmatrix} a & b & a & b & a\cdots a \\ c & a & c & a & c\cdots c \\ b & c & b & c & b\cdots b \\ . & . & . & . & .\cdots . \end{pmatrix} = b \tag{59}$$

Let,

$$P = (acb..., P_{-1}) = \begin{pmatrix} a & b & a & b & a\cdots a \\ c & a & c & a & c\cdots c \\ b & c & b & c & b\cdots b \\ . & . & . & . & .\cdots . \end{pmatrix}$$

For any P_j in the profile P, $d(acb..., P_j) \leq 2$. Moreover we have already noted that the two top-ranked alternatives in $acb...$ and $cab...$ are the same. Therefore using Lemma 1 repeatedly we get,

$$f\begin{pmatrix} c & b & c & b & c\cdots c \\ a & a & a & a & a\cdots a \\ b & c & b & c & b\cdots b \\ . & . & . & . & .\cdots . \end{pmatrix} = b \tag{60}$$

Consider now the preference ordering $bca...$ for voter 2. The top ranked elements of $bac...$ and $bca...$ are the same. Then from equation 50 and using Lemma 1, it follows that,

$$f\begin{pmatrix} c & b & c & b & c\cdots c \\ a & c & a & a & a\cdots a \\ b & a & b & c & b\cdots b \\ . & . & . & . & .\cdots . \end{pmatrix} = b \tag{61}$$

Using a similar argument we get,

$$f(\tilde{P}) = f\begin{pmatrix} c & b & c & a & c\cdots c \\ a & c & a & b & a\cdots a \\ b & a & b & c & b\cdots b \\ . & . & . & . & .\cdots . \end{pmatrix} \in \{a, b\} \tag{62}$$

Case Ia: Let $f(\tilde{P}) = a$. Then we claim that,

$$f(\bar{P}) = f\begin{pmatrix} c & b & c & a & c\cdots c \\ a & c & a & c & a\cdots a \\ b & a & b & b & b\cdots b \\ . & . & . & . & .\cdots . \end{pmatrix} = a \tag{63}$$

Suppose not. Then

$$f(\bar{P}) = f\begin{pmatrix} c & b & c & a & c\cdots c \\ a & c & a & c & a\cdots a \\ b & a & b & b & b\cdots b \\ . & . & . & . & .\cdots . \end{pmatrix} \in \{b, c\} \tag{64}$$

If $f(\bar{P}) = b$, and given that the top two elements in $cab...$ and $acb...$ are the same, a repeat of the arguments earlier shows that,

$$f\begin{pmatrix} a & b & a & a & a\cdots a \\ c & c & c & c & c\cdots c \\ b & a & b & b & b\cdots b \\ . & . & . & . & .\cdots . \end{pmatrix} = b \tag{65}$$

Given that the top ranked elements of $bac...$ and $bca...$ are the same and that for any P_i in the profile in equation 55 $d(bac..., P_i) = 2$, we can use Lemma 1 to say,

$$f(\check{P}) = f\begin{pmatrix} a & b & a & a & a\cdots a \\ c & a & c & c & c\cdots c \\ b & c & b & b & b\cdots b \\ . & . & . & . & .\cdots . \end{pmatrix} = b \tag{66}$$

Given that the top-ranked elements of acb and abc are the same and the profile $\check{P}$ is such that $D(\check{P}) = 2$, by repeatedly using Lemma 1 we get,

$$f(\check{P}) = f\begin{pmatrix} a & b & a & a & a\cdots a \\ b & a & b & b & b\cdots b \\ c & c & c & c & c\cdots c \\ . & . & . & . & .\cdots . \end{pmatrix} = b \tag{67}$$

Now in the profile $\check{P}$, voter 2 has preference ordering $(bac...)$ and all other voters have preference ordering $(abc...)$. Given that we have assumed that the claim in Step 2 is false, equation 67 cannot hold. Therefore we have $f(\bar{P}) = c$. According to Case IA $f(\tilde{P}) = f(\bar{P}_{-4}, abc...) = a$. Therefore, voter 4 of type $acb...$ can gain by mis-reporting her type as $abc...$ at the profile $\bar{P}$. Observe that, $D(\bar{P}) = 3$. Specifically $d(\bar{P}_2, acb...) = 3$ and for any $i \neq 2, 4$, $d(\bar{P}_i, acb...) = 1$. Proposition 3 then implies that there must exist a P'_{-4} such that, Proposition 3 then implies that there must exist a P'_{-4} such that,

(i) for all $i \neq 2,4$, $d(P'_i, acb...) = 0$

(ii) $d(P'_2, acb...) \leq 2$

(iii) $f(acb..., P'_{-4}) = a$ and

(iv) $f(abc..., P'_{-4}) = c$

For any $i \neq 2,4$ the candidates for P'_i is $acb...$. For P'_2 the candidates are $acb...$, $abc...$, $bac...$, $cab...$, and $cba...$ For the first 3 candidates of P'_2,

$$[f(acb..., P'_{-4}) = a] \Rightarrow [f(abc..., P'_{-4}) = a]$$

If $P_2 = cab...$ then Lemma 1 implies that,

$$[f(acb..., P'_{-4}) = a] \Rightarrow [f(abc..., P'_{-4}) = a]$$

So for $P'_2 \in \{acb..., abc..., bac..., cab...\}$ conditions (iii) and (iv) above cannot be satisfied simultaneously. The only other case to consider is when $P'_2 = cba...$. Now we have seen that $[f(acb..., P'_{-4}) = a] \Rightarrow [f(abc..., P'_{-4}) = a]$. Now the top ranked element in the orderings $cab...$ and $cba...$ are the same. Moreover $D(abc..., cab..., P'_{-\{3,4\}}) = 2$. Therefore using Lemma 1 again, we get that $f(abc..., cab..., P'_{-\{3,4\}}) \in \{a, b\}$. But that contradicts requirement (iv). Therefore we get that, $f(\bar{P}) \neq c$. Hence,

$$f(\bar{P}) = a \tag{68}$$

But equation 68 contradicts equation 48. Therefore, $f(\tilde{P}) \neq a$.

Case Ib: Let $f(\tilde{P}) = b$. Since the top two elements in $cab...$ and $acb...$ are the same and $D(cab..., \tilde{P}_{-1}) = 2$, using Lemma 1 we can show that $f(acb, \tilde{P}_{-1}) = b$. repeating the argument above we see that,

$$f(\breve{P}) = f\begin{pmatrix} a & b & a & a & a \cdots a \\ c & c & c & b & c \cdots c \\ b & a & b & c & b \cdots b \\ . & . & . & . & . \cdots . \end{pmatrix} = b \tag{69}$$

Since the top-ranked elements of $bca...$ and $bac...$ are the same and $D(bac..., \breve{P}_{-2}) = 2$, we can use Lemma 1 to say that $f(bac..., \breve{P}_{-2}) = b$. Now repeatedly using the arguments of Proposition 3 we get

$$f\begin{pmatrix} a & b & a & a & a \cdots a \\ b & a & b & b & b \cdots b \\ c & c & c & c & c \cdots c \\ . & . & . & . & . \cdots . \end{pmatrix} = b \tag{70}$$

But equation 70 contradicts our supposition that the claim made in Step 2 is false. Hence, $f(\tilde{P}) \neq b$ and as a consequence, $f(\hat{P}) \neq b$. This completes Case I.

Case II: Let,

$$f(\hat{P}) = f\begin{pmatrix} c & b & a & b & a\cdots a \\ a & a & b & a & b\cdots b \\ b & c & c & c & c\cdots c \\ . & . & . & . & .\cdots . \end{pmatrix} = a \tag{71}$$

Given a preference profile P and a preference ordering P_i let $S(P, P_i) = \{j \in N | P_j = P_i\}$. There are two sub-cases to consider:

Case IIa: Let $|S(\hat{P}, abc...)| > 1$. That is there are more than one agent with preference $abc...$ in $\hat{P}$. Now pick a $j \in S(\hat{P}, abc...)$ and change her preference to $bac...$. Applying proposition 3 it follows that $f(\hat{P}_{-j}, abc...) \in \{a, b\}$. Let $f(P^\star) = f(\hat{P}_{-j}, abc...) = b$. Then proceed to Case I and we will arrive at the contradiction. So suppose

$$f(P^\star) = a \tag{72}$$

If $|S(P^\star, abc...)| > 1$ we go back to the beginning of Case IIa with $P^\star$ as the relevant profile.

Case IIb: Since N is finite the process will reach a stage i.e., a $\tilde{P}$ such that $|S(\tilde{P}, abc...)| = 1$. Let $j = 3$ be the individual with $\tilde{P}_3 = abc...$. For agent 1 $\tilde{P}_1 = cab...$. For all other agents $\tilde{P}_i = bac...$ and $f(\tilde{P}) = a$. i.e.,

$$f(\tilde{P}) = f\begin{pmatrix} c & b & a & b & b\cdots b \\ a & a & b & a & a\cdots a \\ b & c & c & c & c\cdots c \\ . & . & . & . & .\cdots . \end{pmatrix} = a \tag{73}$$

But then using Lemma 1 we get that,

$$f(\tilde{P}_{-n}, bca...) = f\begin{pmatrix} c & b & a & b & b\cdots b \\ a & a & b & a & a\cdots c \\ b & c & c & c & c\cdots a \\ . & . & . & . & .\cdots . \end{pmatrix} \in \{a, c\} \tag{74}$$

If $f(\tilde{P}_{-n}, bca...) = c$ then Lemma 1 implies that,

$$f(\tilde{P}_{-\{3,n\}}, bca..., bac...) = f\begin{pmatrix} c & b & b & b & b\cdots b \\ a & a & a & a & a\cdots c \\ b & c & c & c & c\cdots a \\ . & . & . & . & .\cdots . \end{pmatrix} = c \tag{75}$$

Repeatedly using the above argument, we get,

$$f(\bar{P}) = f\begin{pmatrix} c & b & b & b & b\cdots b \\ a & c & c & c & c\cdots c \\ b & a & a & a & a\cdots a \\ . & . & . & . & .\cdots . \end{pmatrix} = c \tag{76}$$

If $f(\bar{P})$ then $f(cba..., \bar{P}_{-1}) = c$ which contradicts our assumption that the 'coalesced' voter $\bar{\mathbf{1}}$ dictates over the pair $\{c, b\}$. So

$$f(\tilde{P}_{-n}, bca...) = a \tag{77}$$

Using the same argument as above we can say that,

$$f(\breve{P}) = f\begin{pmatrix} c & b & a & b & b\cdots b \\ a & c & b & c & c\cdots c \\ b & a & c & a & a\cdots a \\ . & . & . & . & .\cdots . \end{pmatrix} = a \tag{78}$$

But then since the top two elements of *cba*... and *bca*... are the same, equation 78 implies that, $f(\breve{P}_{-4}, cba...) = a$. And repeatedly using the above argument for all agents $j \in \{4, ..., n\}$ we get that,

$$f\begin{pmatrix} c & b & a & c & c\cdots c \\ a & c & b & b & b\cdots b \\ b & a & c & a & a\cdots a \\ . & . & . & . & .\cdots . \end{pmatrix} = a \tag{79}$$

But 79 implies that,

$$f(\grave{P}) = f\begin{pmatrix} c & b & a & c & c\cdots c \\ a & c & b & a & a\cdots a \\ b & a & c & b & b\cdots b \\ . & . & . & . & .\cdots . \end{pmatrix} = a \tag{80}$$

But then using the arguments made in Case I we can see that $f(\grave{P}) = f(\grave{P}_{-3}, acb...) = a$ which contradicts equation 48.

Step 3. We complete the proof of Claim 1 by showing that $f(P) = r_1(P_k)$ for all profiles P such that $D(P) = 1$. We have already established in Step 2 that for any profile for any profile $P^k = (P_i, \cdots, P_i, P_k, P_i, \cdots, P_i)$ such that $D(P^k) = 1$, $f(P^k) = r_1(P_k)$. In the preference profile P^k the k-th voter has the preference P_k and all the other voters have the preference ordering P_i. Let us now consider a preference profile P such that (i) the k-th agent has the preference ordering P_k and (ii) $D(P) = 1$. We will show that $f(P) = r_1(P_k)$. Since $D(P) = 1$, for any $i \in N \setminus \{k\}$ P_i has to be of the following form:

(i) either $[r_1(P_i) = r_2(P_k)$ and $r_2(P_i) = r_1(P_k)$ and for all $t \neq 1, 2$, $r_t(P_i) = r_t(P_k)]$ or

(ii) $r_1(P_i) = r_1(P_k)$

Define $S(P, P_k) = \{j \in N \setminus \{k\} | P_j = P_k\}$. If $S(P, P_k) = \emptyset$ and $D(P) = 1$, then $P = P^k$ and by Step 2, $f(P) = r_1(P_k)$. Consider now the case where $D(P) = 1$ and $|S(P, P_k)| = 1$. In other words there exists only one voter $j \in N \setminus \{k\}$ such that $P_j = P_k$. We claim that $f(P) = r_1(P_k)$. Without loss of generality assume $r_1(P_k) = b$ and $r_2(P_k) = a$. From the discussion above it follows that for all the voters $i \neq k, j$ the top-two alternatives in their preferences are a and b respectively. For voter j the top ranked alternative is b and the second-ranked alternative is a. The rankings of all the other alternatives are the same across all voters. Suppose that $f(P) \neq r_1(P_k) = b$. From Pareto efficiency it follows then that $f(P) = a$. Now in the profile P^k, for voter j we have $aP_j^k bP_j^k x$ for all $x \in A \setminus \{a, b\}$ and in the profile P, for voter j we have $bP_j aP_j x$ for all $x \in A \setminus \{a, b\}$. So in the profile P, voter j can gain by mis-reporting his preference ordering as P_j^k instead of P_j. From Proposition 3 it then follows that there must exist a profile $\tilde{P}_{-j}$ such that

(i.) for all $i \neq j$ $d(\tilde{P}_i, P_j) < d(P_i, P_j)$,

(ii.) $f(\tilde{P}_{-j}, P_j) = b$ and

(iii.) $f(\tilde{P}_{-j}, P_j^k) = z$ where either $z = a$ or $aP_j z$.

However, $P_k = P_j$ i.e., $d(P_k, P_j) = 0$. Hence there does not exist a profile $\tilde{P}_{-j}$ satisfying condition (i.). Therefore, $f(P) = b$. Now we consider the general case where the profile P is such that $D(P) = 1$ and $|S(P, P_k)| > 1$. We can find a sequence of profiles P^r, $r = 0, 1, \cdots, T$ such that (i) $P^0 = P^k$ (ii) $P^T = P$ (iii) for all r, P^r is such that $D(P^r) = 1$ and (iv) $|S(P^{r+1}, P_k)| = |S(P^r, P_k)| + 1$, $r = 1, \cdots, T - 1$. In other words, the profile P can be obtained from P^k by sequentially changing the preferences of the voter j's from P_j^k to P_j. Using the arguments in the previous two paragraphs we can conclude that $f(P^r) = b$ implies $f(P^{r+1}) = b$, $r = 0, \cdots, T$. Hence $f(P) = b = r_1(P_k)$ which establishes Step 3. Steps 1, 2 and 3 complete the proof of Claim 1. ■

Proof of Claim 2: We now complete the proof of Claim 2.

Proof: We assume without loss of generality that k is an integer strictly greater than 1 and that $f(P') = r_1(P'_1)$ whenever $D(P') \leq k$. Let P be a profile such that $D(P) = k + 1$. We will show that $f(P) = r_1(P_1)$.

Let P^1 be a profile such that

1. $P^1 = (P_j, \underbrace{P_i, \cdots, P_i}_{n-1})$

2. $d(P_i, P_j) = k + 1$

That is all voters other than voter 1 has preference ordering P_i and voter 1 has preference ordering P_j. We first show that $f(P^1) = r_1(P_j)$. Suppose $f(P^1) = x$. Let $\hat{P}_i$ be an ordering obtained by lifting x to the top of P_j leaving the relative rankings of all other alternatives unchanged. Formally,

1. $r_1(\hat{P}_i) = x$ and

2. for all $y, z \neq x$, $y\hat{P}_i z \Leftrightarrow yP_j z$.

Observe that $d(\hat{P}_i, P_j) = t$ where x is $t + 1^{th}$ ranked under P_j, i.e. $r_{t+1}(P_j) = x$. We claim that exactly one of the following two cases must hold.

Case A: $t < k + 1$

Case B: $P_i = \hat{P}_i$.

Suppose $x \neq r_1(P_i)$. Let $w = r_1(P_i)$ so that $wP_i x$. Since x is efficient in the profile P, we must have $xP_j w$, i.e. the rank of w in P_j is at least $t+2$. In order to transform P_i to P_j, the minimal number of transpositions required are (i) at least one to make x first ranked and (ii) at least t to make w, $t+2$ ranked starting from rank 2. Hence $d(P_i, P_j) = k+1 \geq t+1$. This implies that Case A holds.

Now suppose $x = r_1(P_i)$. If the ranking of any pair of alternatives y, z distinct from x differs between P_i and P_j (i.e. $yP_i z$ and $zP_j y$), then $d(\hat{P}_i, P_j) < d(P_i, P_j)$ and Case A holds again. The only remaining possibility is that P_i and P_j agree on all alternatives distinct from x. In this case $P_i = \hat{P}_i$ and Case B holds. Note that if Case B holds, $k + 1 = t$ so that Case A does not hold. Summarizing, we have shown that Cases A and B are mutually exclusive and exhaustive. We now deal with each case in turn.

Case A: We begin by claiming claim that,

$$d(\hat{P}_i, P_i) = d(P_i, P_j) - t$$

First observe that by triangle inequality $d(\hat{P}_i, P_i) \geq d(P_i, P_j) - t$. We will show that the equality is exact. Observe that for any $y, z \neq x$, any one of the following three is true:

$$\begin{aligned}
&[yP_i z \text{ and } yP_j z] \Leftrightarrow [yP_i z \text{ and } y\hat{P}_i z] \\
&[yP_i z \text{ and } zP_j y] \Leftrightarrow [yP_i z \text{ and } z\hat{P}_i y] \\
&[zP_i y \text{ and } yP_j z] \Leftrightarrow [zP_i y \text{ and } y\hat{P}_j z]
\end{aligned} \tag{81}$$

In other words if for any pair of alternatives $y, z \neq x$, if the relative rankings of y and z agree (disagree) in P_i and P_j, then they also agree (disagree) in P_i and $\hat{P}_i$. Also observe that any alternative that is ahead of x in P_j will be below x in P_i and vice versa; otherwise $x \notin PE(P^n)$. Moreover, for any such alternative z, the relative ranking of x and z is the same in P_i and $\hat{P}_i$. Summarizing we have, $[xP_iz$ and $zP_jx] \Rightarrow [xP_iz$ and $x\hat{P}_iz]$. Since $|\{z|zP_2x\}| = t$, equation 81together with the last argument imply that $d(P_i, \hat{P}_i)$ can at most be $(k+1)-t$. Hence $d(\hat{P}_i, P_i) \leq d(P_i, P_j) - t$. Combining the last inequality with the in equality above we have $d(\hat{P}_i, P_i) = d(P_i, P_j) - t$. This implies $d(\hat{P}_i, P_i) \leq k$. We have already seen that $d(\hat{P}_i, P_j) = t \leq k$. Now consider the profile $\hat{P} = (P_j, \hat{P}_i, P_i, \cdots, P_i)$. In other words in the profile $\hat{P}$ voter 1 has the ordering P_j, voter 2 has the ordering $\hat{P}_i$ and all other voters have the preference ordering P_i. Given the above argument $D(\hat{P}) \leq k$. Therefore the induction hypothesis applies to the profile $\hat{P}$. Hence, $f(\hat{P}) = r_i(P_j)$. Suppose $r_1(P_j) = y \neq x$. Observe that voter 2 in profile $\hat{P}$ gains by mis-reporting P_i instead of $\hat{P}_i$. (If voter 2 mis-reports his preference ordering as P_i, the outcome is $f(P_j, P_i, \cdots, P_i) = x = r_1(\hat{P}_i)$). Applying Proposition 3, we conclude that there exists a $\tilde{P}_{-2}$ (a preference profile for voters other than 2) such that,

(i.) for all $l \neq 2$, $d(\hat{P}_i, \tilde{P}_l) < d(\hat{P}_i, P_l)$

(ii.) $f(\hat{P}_i, \tilde{P}_{-2}) = x$

(iii.) $f(P_i, \tilde{P}_{-1}) \in W(y, \hat{P}_i)$

The hypothesis of Case A. and (i) implies, that $f(\hat{P}_i, \tilde{P}_{-2}) = r_1(\tilde{P}_1)$. Then (ii) implies that $r_1(\tilde{P}_1) = x$.

Since, for all $l \notin \{1, 2\}$, $P_l = P_i$, from (i) we get that, for all $l \notin \{1, 2\}$, $d(\hat{P}_i, \tilde{P}_l) < d(\hat{P}_i, P_i) = k + 1 - t < k + 1$. For $l = 1$, (notice $P_1 = P_j$), by triangle inequality, $d(P_i, \tilde{P}_n) \leq d(\hat{P}_i, P_i) + d(\hat{P}_i, \tilde{P}_n)$. Now $d(\hat{P}_i, \tilde{P}_n) < d(\hat{P}_i, P_n) = d(\hat{P}_i, P_j) = t$, by assumption. Since $d(\hat{P}_i, P_i) = d(P_i, P_j) - t$ we have, $d(P_i, \tilde{P}_n) < d(P_i, P_j) = k + 1$. The induction hypothesis therefore applies to the profile $(P_i, \tilde{P}_{-2})$, i.e., $f(\tilde{P}) = r_1(\tilde{P}_1)$. We have already established that $r_1(\tilde{P}_1) = x$. But in order for requirement (iii) for $\hat{P}_i$ to hold, we must either have $x = y$ or $y\hat{P}_ix$. Since $x \neq y$ by assumption and $x = r_1(\hat{P}_i)$ by construction, neither can hold and we have a contradiction. Therefore $x = y$ must hold, so that $f(P^1) = r_1(P_1) = r_1(P_j)$. This completes the argument for Case A.

Case B: Suppose that $P^1 = (P_j, \underbrace{P_i, \cdots, P_i}_{n-1})$ is such that $d(P_i, P_j) = k+1$, $r_1(P_i) = x$, P_i and P_j agree on all alternatives other than x and $f(P^1) = x \neq r_1(P_j) = y$. It is clear that $x = r_{k+2}(P_j)$. Using Claim 1, we can also assume that $k \geq 1$; otherwise $d(P_i, P_j) = 1$ which has been dealt with in Claim 1. Construct P'_j by transposing x with the alternative immediately above it in P_j. Since $k \geq 1$ and $x = r_{k+2}(P_2)$, we still have $r_1(P'_j) = r_1(P_j) = y$. We must also have $d(P_i, P'_j) = k$ and hence $D(P^1_{-1}, P'_j) = k$. Hence, $f(P^1_{-1}, P'_j) = y$ by the induction hypothesis. Since yP_jx, it follows from Proposition 3, that there exists a P'_{-1} such that,

(i.) for all $i \neq 1$, $d(P'_i, P_j) < k+1$,

(ii.) $f(P'_{-1}, P_j) = y$ and

(iii.) $f(P'_{-1}, P'_j) = w$ implies either $w = x$ or xP_jw

We first claim that for any $i \neq 1$, $r_1(P'_i) \neq x$. If this were true then $d(P'_i, P_j) > k+1$. To see this observe that P_i and P_j agree on the rankings of all alternatives other than x. Hence, for any P'_i with $r_1(P'_i) = x$, $d(P'_i, P_j) \geq k+1$ with the equality holding only if $P_i = P'_i$. Let l be such that $d(P'_l, P_j) = \max_{i \in N\setminus 1} d(P'_i, P_j)$, that is, in the profile P'_{-1}, the l-th voter has the ordering P'_l that has the maximum distance from P_j. Given the argument above $r_1(P'_l) \neq x$. So let $z = r_1(P'_l)$. If xP_jz then it would require at least $k+1$ transpositions from P_j for z to be first ranked, i.e., $d(P'_l, P_j) \geq k+1$. Therefore, zP_jx. Since by construction $d(P'_i, P_j) \leq d(P'_l, P_j)$ for all $i \neq 1$, either, (i) $z = r_1(P'_i)$ or (ii) zP'_ix. Otherwise, $d(P'_i, P_j) > d(P'_l, P_j)$. The same is true for any w such that xP_jw, i.e., zP'_iw. The construction of P'_j implies zP'_jw. There are two cases to consider. First, let $w \neq x$. Since for all $i \neq 1$, zP'_iw, and for voter 1 with preference P'_j, zP'_jw, $f(P'_{-1}, P'_j) = w$ contradicts the assumption that f is efficient at profile (P'_{-1}, P'_j). Hence P'_{-1} satisfying requirements (i), (ii) and (iii) cannot exist. The other case to consider is $w = x$ and z is ranked immediately above x in P_j. This means that x is ranked immediately above z in P'_j. If there exists an element between z and x under P_j the first case applies. If z is ranked immediately above x in P_j, then $z = r_{k+1}(P_j)$. Since z is the top-ranked element in P'_l, the minimum distance between P'_l and P_j is k, i.e., $d(P'_l, P_j) \geq k$. But as mentioned above $d(P'_l, P_j)$ has to be less than $k+1$. Therefore the only allowable case is $d(P'_l, P_j) = k$. Since $z = r_{k+1}(P_j)$ and $d(P'_l, P_j) = k$, it must be the case that for all $x, y \neq z$,

$$[xP'_ly \Leftrightarrow xP_jy]. \tag{82}$$

Otherwise, $d(P'_l, P_j) > k$. Since for all $i \neq 1$, $d(P'_i, P_j) \leq d(P'_l, P_j)$, equation 82 together with the fact that $k \geq 1$ implies that there exists a $v \in A \setminus z$ such that, for all $i \neq 1$

$$vP'_ix \text{ and } vP_jx \tag{83}$$

Since P_j and P'_j agree on the ranking of all alternatives other than x and z, $vP_jx \Rightarrow vP'_jx$. But then $x \notin PE(P'_{-1}, P'_j)$. Hence P'_{-1} satisfying requirements (i), (ii) and (iii) cannot exist. Therefore $f(P^1) = r_1(P_j)$ completing the argument for Case B.

We now complete the proof of Claim 2. We show that for any preference profile P such that, $D(P) = k+1$, $f(P) = r_1(P_1)$.

We know from the previous steps that for any $i \in N \setminus \{1\}$, $f(P_1, \underbrace{P_i, \cdots, P_i}_{n-1}) = r_1(P_1) = y$ (say). Therefore,

$$f(P_1, P_2, \cdots, P_2) = f(P_1, P_3, \cdots, P_3) = y$$

Consider now the profile $(P_1, P_2, P_3, \cdots, P_3)$. We claim $f(P_1, P_2, P_3, \cdots, P_3) = y$. Suppose not. Suppose $f(P_1, P_2, P_3, \cdots, P_3) = x \neq y$. We claim that yP_3x. To see this observe that $f(P_1, P_3, \cdots, P_3) = y$. If xP_3y, it implies voter 2 gains by reporting P_2 at the profile $(P_1, P_3, \cdots, P_3)$ where there is at least another voter with the same preference P_3 as voter 2. But according to Lemma 2 that is not possible. Therefore, yP_3x. Now suppose xP_2y. From Lemma 2 then it follows that $f(P_1, P_2, P_2, P_3, \cdots, P_3) \in B(x, P_2)$. Repeatedly using the same argument we get that $f(P_1, P_2, \cdots, P_2) \in B(x, P_2) = z$ (say). Since $z \in B(x, P_2)$, and xP_2y by assumption, zP_2y which contradicts the fact that $f(P_1, P_2, \cdots, P_2) = y$. Therefore, yP_2x. Since $y = r_1(P_1)$, $x \notin PE(P_1, P_2, P_3, \cdots, P_3)$. Therefore we have a contradiction. Hence $f(P_1, P_2, P_3, \cdots, P_3) = y$.

Let us now consider the profile, $(P_1, P_2, P_3, P_4, P_3, \cdots, P_3)$, that is voters 1, 2, 3 and 4 have preferences P_1, P_2, P_3 and P_4 respectively, and the remaining voters have preferences P_3. We claim $f(P_1, P_2, P_3, P_4, P_3, \cdots, P_3) = y$. Suppose not. Let $f(P_1, P_2, P_3, P_4, P_3, \cdots, P_3) = w \neq y$. Using the same argument as before we get yP_3w. Now we claim that yP_2w. To see this observe that if wP_2y, then repeatedly using a similar argument as above we get $f(P_1, P_2, \cdots, P_2) = z \in B(w, P_2)$. In other words zP_2y. But this contradicts the fact that $f(P_1, P_2, \cdots, P_2) = y$. Now suppose that wP_4y. Then again using Lemma 2 repeatedly we get $f(P_1, P_4, \cdots, P_4) = a \in B(y, P_4)$. This implies aP_4y. But this contradicts the fact that $f(P_4, P_4, P_4, P_4, P_4, \cdots, P_4) = y$. Therefore, yP_4w. Now $y = r_1(P_1)$. This implies, $w \notin PE(P_1, P_2, P_3, P_4, P_3, \cdots, P_3)$. Thus, we have arrived at a contradiction. Therefore, $f(P_1, P_2, P_3, P_4, P_3, \cdots, P_3) = y$. Now suppose that for some $1 < t \leq n-1$ $f(P_1, P_2, \cdots, P_t, P_t, \cdots, P_t) = y$ We will show that $f(P_1, P_2, \cdots, P_t, P_{t+1}, \cdots, P_t) = y$. Suppose not. Let $f(P_1, P_2, \cdots, P_t, P_{t+1}, \cdots, P_t) = v \neq y$. Using the same arguments as before we can say that for all $l \leq t$, yP_lv. Let $vP_{t+1}y$. Then using Lemma 2 repeatedly we get that, $f(P_1, P_{t+1}, \cdots, P_{t+1}) = b \in B(v, P_{t+1})$. But that contradicts the fact that $f(P_1, P_{t+1}, \cdots, P_{t+1}) = y$. Therefore $yP_{t+1}v$. But then $v \notin PE(P_1, P_2, \cdots, P_t, P_{t+1}, \cdots, P_t)$. This completes the proof of claim 2. ■

Claims 1 and 2 together complete the proof of Theorem 3.

■