Incentive-Compatible Voting Rules with Positively Correlated Beliefs

Mohit Bhargava [*], Dipjyoti Majumdar [†] and Arunava Sen [‡]

March 20, 2009

Abstract

We study the consequences of positive correlation of beliefs in the design of voting rules in a two-player model. 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.

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

[*] J.P. Morgan, London, United Kingdom.

[†] Concordia University, Montreal, Canada.

[‡] Indian Statistical Institute, New Delhi, India.

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. An example of the sort of situation we have in mind is the process of awarding the highly prestigious Chess Oscar every year by the Russian chess magazine *64*. A group of chess journalists and experts are asked to provide their ranking of the best chess players active during that year and these opinions are aggregated to select a recipient of the Oscar. Since the opinions of the voters are based on the performances of chess players in tournaments and match play, they are highly likely to be positively correlated.

Our goal in this paper is to explore the consequences of the assumption of positive correlation on mechanism design in this context. In doing so, we have to confront the issue of how to interpret positive correlation in distributions over linear orders. In order to avoid further complications, we consider a model with only two voters. 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 the other voter’s preference ordering being closer to her own in the Kemeny metric. The other notion of positive correlation is based on the likelihood of the other voter’s top k alternatives (for any k) agreeing with one’s own opinion of the top k alternatives. In the chess Oscar example, assume that the three players in serious

[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 easily achieved. See Maskin (1999).

contention are Anand (A), Carlsen (C) and Ivanchuk (I). Assume that a voter's opinion is A followed by C followed by I. Then she believes that it is more likely that the other voter's best alternative is A rather than either C or I. In addition she believes that it is more likely that the set of the other voter's top two players is $\{A, C\}$ rather than either $\{A, I\}$ or $\{C, I\}$. We call these two notions of positive correlation, K (or Kemeny) correlation and "top-set" or TS correlation and show that a belief that is K correlated is also TS-correlated.

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 type of the other voter and the assumption that the other voter is 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. Importantly, when beliefs are locally perturbed they must remain within the appropriate class of positively correlated beliefs. 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. More precisely, 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 provide an example of a non-dictatorial SCF satisfying unanimity which is TS-ROBIC (and hence K-ROBIC). In other words, robustness with respect to all positively correlated beliefs on the complete domain of preferences, does not lead to truth-telling being a weakly dominant strategy. 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. There are at least two significant differences between the models and consequently, the results. The first is that monetary transfers which are at 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. The permissive possibility results of (Crémer and Mclean, 1988) require only a "small" amount of correlation, either negative or positive. The same universally permissive result with small correlation does not hold in our model. We focus on characterizing SCFs which are incentive-compatible and satisfy additional robustness properties with respect to beliefs.

The paper is organized as follows. The next section introduces basic notation 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.

2 Notation and Definitions

There are two individuals or voters in the society, i.e., $N = \{1, 2\}$. The set of outcomes is the set A with $|A| = m$. Elements of A will be denoted by a, b, c, d etc. Let $\mathcal{P}$ 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 $\mathcal{P}^2$. Preference

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

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 $i = 1, 2$.

For all $P_i \in \mathcal{P}$ 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 $|\{b \neq a | bP_i a\}| = k - 1$. For all $i \in \{1, 2\}$, for any $P_i \in \mathcal{P}$ and for any $a \in A$, let $B(a, P_i) = \{b \in A | bP_i a\} \cup \{a\}$. Thus $B(a, P_i)$ is the set of alternatives that are weakly preferred to a under P_i

Definition **1** *A Social Choice Function or (SCF) f is a mapping $f : \mathcal{P}^2 \to A$.*

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 $i \in \{1, 2\}$.*

The axiom states that in any situation where both 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 both 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 $P \in \mathcal{P}^2$, there does not exist an alternative $x \in A$ such that $xP_i f(P)$ for all $i = 1, 2$.*

A SCF is anonymous if it is symmetric across voters, i.e. it does not discriminate amongst voters.

Definition **4** *A SCF is anonymous or satisfies anonymity if for all $P_1, P_2 \in \mathcal{P}$, $f(P_1, P_2) = f(P_2, P_1)$. Here (P_2, P_1) is the profile where voter 1 has preference P_2 and voter 2 has preference P_1.*

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

Definition **5** *A SCF f is dictatorial if there exists a voter i such that for all profiles $P \in \mathcal{P}^2$, $f(P) = r_1(P_i)$.*

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 6 *A SCF f is dominant strategy incentive-compatible or strategy-proof if, for all $P_i, P_j, P_i' \in \mathcal{P}$, either $f(P_i, P_j) = f(P_i', P_j)$ or $f(P_i, P_j)P_i f(P_i', P_j)$ 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 7 *A belief for voter i is a probability distribution on the set $\mathcal{P}^2$, i.e. it is a map $\mu_i : \mathcal{P}^2 \to [0,1]$ such that $\sum_{P \in \mathcal{P}^2} \mu_i(P) = 1$.*

Clearly μ_i belongs to the unit simplex of dimension $m!^2 - 1$. For all μ_i, for all P_j and P_i, we shall let $\mu_i(P_j|P_i)$ denote the conditional probability of P_j given P_i. A *belief pair* is a pair of beliefs (μ_1, μ_2), one for each voter.

Definition 8 *The utility function $u : A \to \Re$ represents $P_i \in \mathcal{P}$, 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 9 *A SCF f is Ordinally Bayesian Incentive Compatible (OBIC) with respect to the belief pair (μ_1, μ_2) if for all $i \in \{1,2\}$, for all $P_i, P_i' \in \mathcal{P}$, for all u representing P_i, we have*

$$\sum_{P_j \in \mathcal{P}} u\left(f(P_i, P_j)\right) \mu_i(P_j|P_i) \geq \sum_{P_j \in \mathcal{P}} u\left(f(P_i', P_j)\right) \mu_i(P_j|P_i) \tag{1}$$

Suppose f is a SCF which is OBIC with respect to the belief pair (μ_1, μ_2). 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 the other voter j will reveal truthfully so that a probability distribution over outcomes is induced by f and voter'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 pairs (μ_1, μ_2). 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 pairs, 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 10 *The SCF f is OBIC with respect to the belief pair (μ_1, μ_2) if for all $i \in \{1, 2\}$, for all integers $k = 1, \ldots, m$ and for all P_i and P_i',*

$$\begin{aligned} & \mu_i(\{P_j | f(P_i, P_j) \in B(r_k(P_i), P_i)\} | P_i) \\ \geq \ & \mu_i(\{P_j | f(P_i', P_j) \in B(r_k(P_i), P_i)\} | P_i) \end{aligned} \tag{2}$$

Suppose f satisfies OBIC with respect to (μ_1, μ_2). 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.

An important special case of the incomplete information voting model that we have described above, is the *common priors* model (see, for instance...). In this case, the beliefs of the two voters μ_1 and μ_2 are constrained to be the same. The incentive-compatibility restrictions in OBIC for the two voters are then defined with respect to the same conditional beliefs $\mu(.|P_i)$. We shall consider both the general model with possibly non-identical beliefs as well as the common prior model in our analysis.

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

3 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 $P_i \in \mathcal{P}$. 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 $cP_i'aP_i'b$, then $d(P_i, P_i') = 2$. Generally, $d(P_i, P_i') \in \{0, 1, \ldots, \binom{m}{2}\}$ for any $P_i, P_i' \in \mathcal{P}$.

Definition **11** *(K-correlation) A belief μ_i for voter i is said to be positively K-correlated if, for all preferences P_i, P_j and P_j',*

$$d(P_i, P_j) < d(P_i, P_j') \Rightarrow \mu_i(P_j|P_i) > \mu_i(P_j'|P_i)$$

Thus, μ_i is positively correlated in this sense if voter i's of type P_i considers it more likely that the other voters' type is P_j relative to P_j' if the Kemeny distance between P_i and P_j is less than the Kemeny distance between P_i and P_j'. Note that there may be several orderings whose Kemeny distance from P_i is identical. K-correlation imposes no restriction on the relative conditional probabilities of realizing these orderings, given P_i.

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

We propose the following alternative notion of positive correlation. Consider voter i with preferences P_i. Then conditional on her type being P_i, she considers it most likely (amongst all sets of size k) that voter j's set of top k-alternatives for any $k = 1 \ldots, m-1$ is the set of the first k alternatives according to P_i. We call this notion of correlation, "Top-set" or TS-correlation.

Definition **12** *(TS-Correlation) A belief for voter i, μ_i is positively TS-correlated if for all P_i, P_j and for all $k = 1, ..., m-1$*

$$\sum_{\{P_j : B(r_k(P_j), P_j) = B(r_k(P_i), P_i)\}} \mu(P_j|P_i) > \sum_{\{P_j : B(r_k(P_j), P_j) \neq B(r_k(P_i), P_i)\}} \mu(P_j|P_i) \tag{3}$$

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

The following example illustrates both notions of correlation.

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

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

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

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

Observe that

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

On the other hand,

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

It is easy to verify if μ_i satisfies the system of inequalities 5, then μ_i satisfies the system of inequalities 6. 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 $\mu_i^6 = 0.3$.

The relationship illustrated in the example above holds generally as demonstrated by the Proposition below.

Proposition 1 *If $\mu \in K^*$ then $\mu \in TS^*$.*

Proof: Pick $\mu_i \in K^*$. Define $A\Delta B = (A \setminus B) \cup (B \setminus A)$, i.e. $A\Delta B =$ is the set of elements that belong to either A or B but not to both. Pick an arbitrary P_i, an integer $k \leq m$ and a set B such that $|B(r_k(P_i), P_i)| = |B|$. Let $\mathcal{B}_l = \{B : |B(r_k(P_i), P_i)\Delta B| = l\}$. We denote a generic element of $\mathcal{B}_l$ by B_l. Let $\{\mathcal{B}_l\}$, $l = 1, \ldots L$ be the collection of all possible B's such that $|B(r_k(P_i), P_i)| = |B|$. Observe that, if $l = 0$, $B(r_k(P_i), P_i) = B$.

We will prove our claim by induction on l. Observe that l can never be equal to 1 so that the minimum value of l, if $l \neq 0$ is $l = 2$. Suppose that $l = 2$. Let $B(r_k(P_i), P_i)\Delta B = \{x, y\}$. Assume without loss of generality that $x \in B(r_k(P_i), P_i)$ and $y \in B$. Consider now a bijection $\sigma : A \to A$ defined as follows:

- $\sigma(a) = a$ for all $a \in A \setminus \{x, y\}$
- $\sigma(x) = y$ and
- $\sigma(y) = x$.

Given a preference ordering P and a bijection σ, we define P^σ to be the following preference ordering:

$$xPy \Leftrightarrow \sigma(x)P^\sigma\sigma(y)$$

Since $y \notin B(r_k(P_i), P_i)$ and $x \in B(r_k(P_i), P_i)$, for any P_j such that $B(r_k(P_j), P_j) = B(r_k(P_i), P_i)$, we have xP_jy and for the corresponding P_j^σ we have $yP_j^\sigma x$. Thus for every P_j such that $B(r_k(P_j), P_j) = B(r_k(P_i), P_i)$, there exists a P_j^σ such that $B(r_k(P_j^\sigma), P_j^\sigma) = B$ and $d(P_i, P_j) < d(P_i, P_j^\sigma)$.

The last inequality follows from the fact that, xP_iy and xP_jy but $yP_j^\sigma x$ and the remaining alternatives are ranked in the same way in P_j and P_j^σ. Since $\mu_i \in K^*$, we have $\mu_i(P_j|P_i) > \mu_i(P_j^\sigma|P_i)$. Since the above inequality holds for every pair (P_j, P_j^σ), we have,

$$\sum_{\{P_j|B(r_k(P_j),P_j)=B(r_k(P_i),P_i)\}} \mu_i(P_j|P_i) > \sum_{\{P_j|B(r_k(P_j),P_j)=B\}} \mu_i(P_j|P_i) \tag{7}$$

Inequality 7 proves the claim for the case $l = 2$. Suppose now that the claim is true for all $l \leq t$. We will show that the claim is true for $l = t + 1$ Consider now a $B_t \in \mathcal{B}_t$ and a $B_{t+1} \in \mathcal{B}_{t+1}$. Observe that $|B_{t+1} \Delta B_t| = 2$. repeating the same arguments as above but now replacing $B(r_k(P_i), P_i)$ with B_t and B with B_{t+1} it follows that

$$\sum_{\{P_j|B(r_k(P_j),P_j)=B_t\}} \mu_i(P_j|P_i) > \sum_{\{P_j|B(r_k(P_j),P_j)=B_{t+1}\}} \mu_i(P_j|P_i) \tag{8}$$

By the induction hypothesis,

$$\sum_{\{P_j|B(r_k(P_j),P_j)=B(r_k(P_i),P_i)\}} \mu_i(P_j|P_i) > \sum_{\{P_j|B(r_k(P_j),P_j)=B_t\}} \mu_i(P_j|P_i) \tag{9}$$

Combining inequalities 8 and 9, we have

$$\sum_{\{P_j|B(r_k(P_j),P_j)=B(r_k(P_i),P_i)\}} \mu_i(P_j|P_i) > \sum_{\{P_j|B(r_k(P_j),P_j)=B_{t+1}\}} \mu_i(P_j|P_i) \tag{10}$$

Inequality 10 establishes that $\mu_i \in TS^*$. ■

We note that several 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.

We also note that the extension of these notions of correlation to the more than 2 voters case presents additional difficulties. Consider the case of 3 voters and $A = \{a, b, c\}$. Suppose we wish to find a generalization of K-correlation to this setting. Suppose voter 1 has ordering abc while voters 2 and 3 have (abc, cba) respectively in one case and (cab, bac) respectively in another case. Observe that the profile of Kemeny distances in the case of (abc, cba) is $(0, 3)$ while it is $(2, 1)$ in the case of (cab, bac). If μ_i is K-correlated, what relationship if any, is to be assumed between $\mu_i((abc, cba)|abc)$ and $\mu_i((cab, bac)|abc)$? It is clear that one amongst several plausible assumptions can be made.

4 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 **13** *A SCF f is K-locally robust OBIC or K-LOBIC if, for each voter i, there exists $\mu_i \in K^*$ and an $\epsilon > 0$ such that f is OBIC with respect to all $\mu_i' \in B_\epsilon(\mu_i) \cap K^*$.* [4]

Local robustness with respect to TS-correlation can be analogously defined.

Definition **14** *A SCF f is TS-locally robust OBIC or TS-LOBIC if, for each voter i, there exists $\mu_i \in TS^*$ and an $\epsilon > 0$ such that f is OBIC with respect to all $\mu_i' \in B_\epsilon(\mu_i) \cap TS^*$.*

An important observation at this point is that in both definitions 13 and 14 we are allowing for voters to have non-identical beliefs.

Observation **2** 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 **2** Let $A = \{a, b, c\}$; 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*

(11)

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

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

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 $\mu_2(i|j)$.

	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

(12)

Moreover since all the necessary inequalities (for both OBIC and TS-correlation) are satisfies strictly, they will continued 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 **3** Let $A = \{a, b, c\}$. Consider the SCF f^2 as shown in the table below.

	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

(13)

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 **15** *A SCF f satisfies Ordinal Non-Domination (OND) if for all P_i, P_i' and P_j such that $f(P_i', P_j)P_i f(P_i, P_j)$, there exists P_j' such that,*

1. *Either $f(P_i, P_j') = f(P_i', P_j)$ or $f(P_i, P_j')P_i f(P_i', P_j)$ and*
2. *Either $f(P_i, P_j) = f(P_i', P_j')$ or $f(P_i, P_j)P_i f(P_i', P_j')$.*

Consider the SCF f^2 in Example 3. Observe that $f^2(bac, abc) = a$ 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 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.

Our main result in this section is 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

$$\sum_{P_j \in \mathcal{P}} \mu_i(P_j|P_i)\left[u(f(P_i, P_j), P_i) - u(f(P_i', P_j), P_i)\right] \geq 0 \tag{14}$$

Moreover inequality 14 holds for all μ_i' in a neighborhood of μ_i. Hence we can assume without loss of generality that $\mu_i(P_j|P_i) > 0$ in inequality 14. Suppose that there exists P_i, P_j and P_i' such that $f(P_i', P_j)P_i f(P_i, P_j)$, i.e $u(f(P_i', P_j)) > u(f(P_i, P_j))$ for all u representing P_i. Since $\mu_i(P_j|P_i) > 0$, there must exist P_j' such that $u(f(P_i, P_j')) > u(f(P_i', P_j'))$, i.e. $f(P_i, P_j')P_i f(P_i', P_j')$, in order for inequality 14 to hold. Let L denote the set of all such P_j''s.

Now suppose $f(P_i', P_j)P_i f(P_i, P_j')$ holds for all $P_j' \in L$. Then we can choose a utility representation $\hat{u}$ of P_i such that $\hat{u}(f(P_i', P_j))$ is arbitrarily close to 1 and $\hat{u}(f(P_i, P_j')), \hat{u}(f(P_i, P_j))$ and $\hat{u}(f(P_i', P_j'))$ are all arbitrarily close to 0. Then, the L.H.S of 14 for the utility function $\hat{u}$ can be made arbitrarily close to $-\mu_i(P_j|P_i) < 0$ violating inequality 14.

Now suppose $f(P_i', P_j')P_i f(P_i, P_j)$ holds. Then we can choose a utility representation $\tilde{u}$ of P_i such that $\tilde{u}(f(P_i', P_j)), \tilde{u}(f(P_i, P_j'))$ and $\tilde{u}(f(P_i', P_j'))$ are arbitrarily close to 1 and $\tilde{u}(f(P_i, P_j))$ is arbitrarily close to 0. Once again the L.H.S of 14 for the utility function $\tilde{u}$ can be made arbitrarily close to $-\mu_i(P_j|P_i) < 0$ violating inequality 14.

Thus 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 a 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 $k \in \{1, \dots, m\}$ define $A_k^f(P_i) = \{P_j | f(P_i, P_j) = r_k(P_i)\}$. Thus $A_k^f(P_i)$ is the set of preferences for voter j that gives under f the k^{th} ranked alternative of voter i as outcome. Since f satisfies unanimity, $P_i \in A_1^f(P_i)$.

Let $\mathcal{C}_i^*$ denote the set of probability distributions over $\mathcal{P}$ 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. $\mu_i^*(P_j|P_i) > 0$ for all P_j
2. $\mu_i^*(P_i|P_i) > \sum\limits_{P_j \neq P_i} \mu_i^*(P_j|P_i)$
3. for all $P_j \neq P_i$, $\mu_i^*(P_j|P_i) > \sum\limits_{P_i' \in \cup_{r=k+1}^{r=m} A_r^f(P_i)} \mu_i^*(P_i'|P_i)$ where $P_j \in A_k^f(P_i)$.

Suppose $f(P_i, P_j)$ is the k^{th}-ranked alternative in P_i. Then the conditional probability $\mu_i^*(P_j|P_i)$ exceeds the sum of the conditional probabilities of realizing an ordering P_j' where the outcome $f(P_i, P_j')$ is strictly worse than the k^{th}-ranked alternative in P_i. In addition, the conditional probability of realizing P_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 $m!^2 - 1$.

We claim that $\mathcal{C}_i^* \subset TS^*$. This is easily verified by noting that the term $\mu_i^*(P_i|P_i)$ appears in the L.H.S of every inequality in the system of inequalities 6 which define TS correlation while it does not appear on the R.H.S of none of them. In order to complete the proof, we will show that f is OBIC with respect to all beliefs (μ_1^*, μ_2^*) where $\mu_i^* \in \mathcal{C}_i^*$, $i = 1, 2$.

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

$$\sum_{P_j \in L} \mu_i^*(P_j|P_i)\beta(P_j) - \sum_{P_j \in G} \mu_i^*(P_j|P_i)\gamma(P_j) \geq 0 \tag{15}$$

where $\beta(P_j) = [u(f(P_i, P_j)) - u(f(P_i', P_j))]$ and $\gamma(P_j) = [u(f(P_i', P_j)) - u(f(P_i, P_j))]$.

If $G = \emptyset$, inequality 15 is clearly satisfied. Suppose therefore that $G \neq \emptyset$. We claim that for all $P_j \in G$, there exists $P_j' \in L$ satisfying

1. $\beta(P_j') > \gamma(P_j)$
2. $\mu_i^*(P_j'|P_i) > \sum_{\{\tilde{P}_j | f(P_i, P_j') P_i f(P_i, \tilde{P}_j)\}} \mu_i^*(\tilde{P}_j|P_i)$

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 $P_j \in G$, $\sigma(P_j)$ is the $P_j' \in L$ satisfying 1 and 2 above. Let P_j' be an arbitrary element in the range of σ and let $Q(P_j') = \{P_j | \sigma(P_j) = P_j'\}$. A critical observation is that for all $P_j \in Q$, OND implies $f(P_i, P_j') P_i f(P_i, P_j)$, i.e. $Q(P_j') \subset \{\tilde{P}_j | f(P_i, P_j') P_i f(P_i, \tilde{P}_j)\}$. Hence 2 above implies $\mu_i^*(P_j'|P_i) > \sum_{P_j \in Q(P_j')} \mu_i^*(P_j|P_i)$. Moreover using 1 above, we have $\mu_i^*(P_j'|P_i)\beta(P_j') > \sum_{P_j \in Q(P_i')} \mu_i^*(P_j|P_i)\gamma(P_j)$. Now summing up over all P_j' in L and noting that OND implies that $G \subset \cup_{P_j' \in L} Q(P_j')$, we obtain inequality 15.

■

We now make a series of observations regarding Theorem 1 and its implications.

Observation **3** 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. In fact, it can be checked that the following condition which is weaker than unanimity, suffices: for all profiles P such that $P_i = P_j$, $f(P_i, P_j) = r_1(P_i)$.

Observation **4** The proof of the second part of Theorem 1 explicitly 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 voter j's type being P_i (i.e. coinciding with her own); in addition she puts "significantly higher" weight on voter j's type being P_j instead of P_j' if $f(P_i, P_j)$ is strictly better than $f(P_i, P_j')$ 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.

Observation **5** Example 2 demonstrates that OND is not sufficient for the K-LOBIC property to hold. The OND condition guarantees that if misrepresentation is more profitable than truth-telling for some type of voter j, say P_j, then there is another type of j, P_j' where the misrepresentation is "ordinally costlier" than truth-telling, relative to the situation at P_j. In order to strengthen the condition to make it K-LOBIC necessary, additional restrictions on $d(P_i, P_j')$ relative $d(P_i, P_j)$ must also hold. These restrictions may be quite subtle and we do not pursue this question further.

Observation **6** Theorem 1 stands sharply in contrast to results in Majumdar and Sen (2004) for the independent beliefs case. In the latter case, OBIC and local robustness imply that the SCF is dictatorial (if $m \geq 3$), i.e. truth-telling must be dominant.

Observation **7** Does the argument for the sufficiency part of Theorem 1 hold if beliefs are restricted to be common? Observe that our argument pinned downed conditional beliefs for each voter. We can represent them by two non-negative matrices $m! \times m!$, X and Y denoting conditional beliefs for 1 and 2 respectively. If i^{th} row and j^{th} column refer to preferences P_i and P_j respectively in both matrices, then the $(i,j)^{th}$ element in X and Y are $\mu_1(P_j|P_i)$ and $\mu_2(P_i|P_j)$ respectively. All the row sums of X and column sums of Y add up to 1. In order for X and Y to be derived from a common joint distribution, we need to find appropriate marginal probabilities. Specifically, we need to find $m! \times m!$ non-negative diagonal matrices P and Q with trace P = trace $Q = 1$ such that $PX = QY$. Here the $(i,i)^{th}$ elements of

P and Q are the marginal probabilities of voter i' and j's types being P_i respectively. A solution P, Q for general X and Y will not exist because the equation system $PX = QY$ involves $m! \times m!$ equations in $2m$ unknowns (neglecting the constraints of non-negativity and adding upto 1). However we are to find a sufficient condition for TS-LOBIC with respect to common priors as we show below.

Definition **16** *A SCF f is TS-LOBIC with respect to common priors, if there exists a belief $\mu \in TS^*$ and an $\epsilon > 0$ such that f is OBIC with respect to the belief pair (μ', μ') all $\mu' \in B_\epsilon(\mu) \cap TS^*$.*

Our next result states that OND is sufficient in the common priors model if the SCF satisfies the additional hypothesis of anonymity.

Theorem **2** *If a SCF satisfies unanimity, anonymity and OND, then it is TS-LOBIC with respect to common priors.*

Proof: We use the same construction for conditional beliefs as in the proof of the second part of Theorem 1. By virtue of the assumption that f satisfies anonymity, we can find *identical* conditional beliefs satisfying properties 1, 2 and 3 in the definition of the set $\mathcal{C}_i^*$. More precisely, there exists conditional beliefs μ^* such that for all P_i, P_j, we have

$$\mu^*(P_j|P_i) = \mu^*(P_2 = P_j|P_1 = P_i) = \mu^*(P_1 = P_j|P_2 = P_i)$$

Moreover

1. $\mu^*(P_j|P_i) > 0$ for all P_j
2. $\mu^*(P_i|P_i) > \sum_{P_j \neq P_i} \mu^*(P_j|P_i)$
3. for all $P_j \neq P_i$, $\mu^*(P_j|P_i) > \sum_{P'_j \in \cup_{r=k+1}^{r=m} A_r^f(P_i)} \mu^*(P'_j|P_i)$ where $P_j \in A_k^f(P_i)$.

The matrices of conditional probabilities X and Y (Observation 7) are such that $Y = X^T$. Clearly any arbitrary diagonal matrix of priors P (with trace equal to 1) will satisfy the equation $PX = PY$ generating a common prior μ^*. Let $\mathcal{C}^*$ be the set of beliefs which generate conditional beliefs satisfying 1, 2 and 3 above. It is clearly an open set. Replicating the arguments in Theorem 1, it follows that f is TS-LOBIC with respect to any belief (μ^*, μ^*) where $\mu^* \in \mathcal{C}^*$. ■

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

5 Incentive Compatibility with Global Robustness

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

Definition **17** *A SCF $f : \mathcal{P}^2 \to A$ is K-Globally Robust OBIC (K-ROBIC) if it is OBIC with respect to all (μ_1, μ_2) where $\mu_1, \mu_2 \in K^*$.*

We have an analogous definition for TS-correlation.

Definition **18** *A SCF $f : \mathcal{P}^2 \to A$ is TS-Globally Robust OBIC (TS-ROBIC) if it is OBIC with respect to all (μ_1, μ_2) where $\mu_1, \mu_2 \in TS^*$.*

Observation **8** 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,*

$$f(P) = \begin{cases} r_1(P_1), & \text{if } r_1(P_1) = r_1(P_2); \\ x, & \text{otherwise.} \end{cases} \tag{16}$$

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_2aP_2x. The outcome of f^{us} in this profile is x (the status quo alternative). But voter 1 can misreport $bP_1'aP_1'x$ 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 $f^{us}(P_i, P_j) = x$ for all P_j. 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 $\mu(P_j|P_i)$.

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 $f^{us}(P_i, P_j)$ is either a or x for all P_j and $f^{us}(P_i', P_j)$ is either b or x for all P_j, 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_j such that $r_1(P_j) = b$. Denote the set of such P_j's by G. On the other hand, i loses by misreporting P_i' for all P_j such that $r_1(P_j) = a$. Denote the set of such P_j's by L. In particular observe that

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

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

$$\sum_{P_j \in L} u(a)\mu_i(P_j|P_i) + \sum_{P_j \in G} u(x)\mu_i(P_j|P_i) + \sum_{P_j \notin G \cup L} u(x)\mu_i(P_j|P_i) \tag{17}$$

The expected utility from misreporting via P_i' is

$$\sum_{P_j \in L} u(x)\mu_i(P_j|P_i) + \sum_{P_j \in G} u(b)\mu_i(P_j|P_i) + \sum_{P_j \notin G \cup L} u(x)\mu_i(P_j|P_i) \tag{18}$$

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

$$\Delta = [u(a) - u(x)] \sum_{P_j \in L} \mu_i(P_j|P_i) - [u(b) - u(x)] \sum_{P_j \in G} \mu_i(P_j|P_i) \tag{19}$$

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

■

The unanimity with status quo rule has some nice features. It is both anonymous and neutral [5]. 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

[5] An SCF is neutral if it does not discriminate amongst alternatives.

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.*

Before proving the result, we state an prove an auxiliary result which we use repeatedly in the proof of Theorem 3. 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.

PROPOSITION **3** *Let f be a K-ROBIC SCF. Let P_i, P_i' and P_j be such that $f(P_i', P_j)P_i f(P_i, P_j)$. Then there exists P_j' satisfying the following properties:*
(i) $d(P_i, P_j') < d(P_i, P_j)$
(ii) either $f(P_i, P_j') = f(P_i', P_j)$ or $f(P_i, P_j')P_i f(P_i', P_j)$
(iii) either $f(P_i, P_j) = f(P_i', P_j')$ or $f(P_i, P_j)P_i f(P_i', P_j)$

Proof: If f is K-ROBIC, it must also be K-LOBIC. From Theorem 1 and Observation 3 it follows that f must satisfy OND. Suppose that P_i, P_i' and P_j are such that $f(P_i', P_j)P_i f(P_i, P_j)$. Then OND implies that there exists P_j' satisfying (ii) and (iii). It only remains to show (i).

Let $d(P_i, P_j) = k$. Suppose that $d(P_i, P_j') \geq k$ for all $P_j' \in G$. Note that for any $\delta_1, \delta_2 > 0$ we can always choose a utility function u representing P_i such that $u(f(P_i', P_j)) - u(f(P_i, P_j)) = \delta_1$ and

$$\max_{P_j' \in G} \left| [u(f(P_i, P_j')) - u(f(P_i', P_j'))] - [u(f(P_i', P_j)) - u(f(P_i, P_j))] \right| < \delta_2$$

Also for any ϵ_1, ϵ_2 such that $1 > \epsilon_1 > \epsilon_2 > 0$, there exists $\mu \in K^*$ such that (i) $\mu(\hat{P}_j|P_i) > \epsilon_1$ if $d(P_i, \hat{P}_j) < k$ or $\hat{P}_j = P_j$ and (ii) $\mu(\hat{P}_j|P_i) < \epsilon_2$ for all other $\hat{P}_j$. Let $\Delta = \sum_{\hat{P}_j \in \mathcal{P}} [u(f(P_i', \hat{P}_j)) - u(f(P_i, \hat{P}_j))]\mu(\hat{P}_j|P_i)$. It follows that$\Delta \geq \epsilon_1\delta_1 - (m! - 1)(\delta_1 + \delta_2)\epsilon_2$. 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_j' is "ordinally" greater than the loss from truthful reporting at P_j. In addition, the Kemeny distance between the P_i and P_j' must be strictly smaller than that between P_i and P_j. 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 $\mu_i(P_j'|P_i)$ is made arbitrarily small relative to $\mu_i(P_j|P_i)$.

Observation **9** 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 $G = \{P_j | f(P_i', P_j) P_i f(P_i, P_j)\}$ and let $L = \{P_j' | f(P_i, P_j') P_i f(P_i', P_j')\}$. According to Proposition 3, there exists a map $\sigma : G \to L$ such that $P_j \in G$ there exists a $P_j' \in L$ 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 return to the proof of Theorem 3.

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. We shall prove the result by induction on the cardinality of the distance between two profiles. In particular we shall prove the following claims.

CLAIM 1: There exists a voter i such that for all profiles P such that $d(P_1, P_2) = 1$, 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 $d(P_1', P_2') \leq k$, we have $f(P') = r_1(P_i')$. Let P be a profile such that $d(P_1, P_2) = k + 1$. 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.

Proof Claim 1: Pick an arbitrary pair of alternatives $\{a, b\}$ and let (P_1, P_2) be a profile where

1. $r_1(P_1) = a$ and $r_2(P_1) = b$
2. $r_1(P_2) = b$ and $r_2(P_2) = a$
3. $r_k(P_1) = r_k(P_2)$ for all $k > 2$

Observe that $d(P_1, P_2) = 1$. Also a and b are the only efficient alternatives at this profile. Since f is efficient, $f(P_1, P_2)$ is either a or b. Assume w.l.o.g that $f(P_1, P_2) = b$. We will prove Claim 1 by showing that for all profiles (P_1', P_2') such that $d(P_1', P_2') = 1$, we must have $f(P_1', P_2') = r_1(P_2')$. If $r_1(P_1') = r_1(P_2')$, the required conclusion follows trivially from the efficiency of f. We will assume therefore that $r_1(P_1') \neq r_1(P_2')$. Since $d(P_1', P_2') = 1$ by assumption, it must be the case that $r_2(P_1') = r_1(P_2')$, $r_1(P_1') = r_2(P_2')$ and $r_k(P_1') = r_k(P_2')$ for all $k > 2$. We will denote the class of such profiles by $\mathcal{D}^2(1)$. We will prove the claim in a series of steps.

Step 1: Let $\{c, d\}$ be an arbitrary pair of alternatives and let $P', \hat{P} \in \mathcal{D}^2(1)$ 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 $f(P') = d$, then $f(\hat{P}) = d$.

Suppose $f(P') = d$. 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. Suppose $f(\hat{P}_1, P_2') \neq d$. Then the efficiency of f implies that $f(\hat{P}_1, P_2') = c$. Note that $d(\hat{P}_1, P_2') = 2$. Since $cP_1'd$, Proposition 3 implies that there exists $\tilde{P}_2$ such that (i) $d(P_1', \tilde{P}_2) \leq 1$ (ii) $f(P_1', \tilde{P}_2) = c$ (iii) if $f(\hat{P}_1, \tilde{P}_2) = z$, then either $d = z$ or $dP_1'z$. If $r_1(\tilde{P}_2) = c$, efficiency of f implies that $f(\hat{P}_1, \tilde{P}_2) = c$ contradicting (iii). But if $r_1(\tilde{P}_2) \neq c$, then (i) implies that $\tilde{P}_2 = P_2'$ which would in turn would imply that $f(P_1', \tilde{P}_2) = d$ contradicting (ii). Hence $f(\hat{P}_1, P_2') = d$.

We now show that $f(\hat{P}) = d$. Suppose this is false. Then efficiency of f implies that $f(\hat{P}) = c$. Since $d\hat{P}_2c$, Proposition 3 implies that there exists $\tilde{P}_1$ such that (i) $d(\tilde{P}_1, \hat{P}_2) < 1$ (ii) $f(\tilde{P}_1, \hat{P}_2) = d$ and (iii) if $f(\tilde{P}_1, P_2') = z$, then either $z = c$ or $c\hat{P}_2z$. But (i) implies that $\tilde{P}_1 = \hat{P}_2$. In that case efficiency implies $f(\tilde{P}_1, P_2') = d$ contradicting (iii). Hence $f(\hat{P}) = d$.

Now consider the general case where $\hat{P} \in \mathcal{D}^2(1)$ is such that $r_1(\hat{P}_1) = r_2(\hat{P}_2) = c$ and $r_1(\hat{P}_2) = r_2(\hat{P}_1) = d$. We can find a sequence of profiles P^r, $r = 0, 1, .., T$ such that (i) $P^0 = P'$ (ii) $P^T = \hat{P}$ (iii) $P^r \in \mathcal{D}^2(1)$ with $r_1(P_1^r) = r_2(P_2^r) = c$ and $r_1(P_2^r) = r_2(P_1^r) = d$ for all r and (iii) $d(P_i^r, d_i^{r+1}) = 1$, $r = 0, .., T-1$, $i = 1, 2$. In other words, the profile $\hat{P}$ can be obtained from P' be a sequence of transpositions not involving c or d. Using the arguments in the two previous paragraph, we can conclude that $f(P^r) = d$ implies $f(P^{r+1}) = d$, $r = 0, ..., T-1$. Hence $f(\hat{P}) = d$ which establishes Step 1.

Let $\{c, d\}$ be an arbitrary, ordered pair of alternatives. We will say that voter i, $i = 1, 2$ dictates over $\{c, d\}$ if for all $P \in \mathcal{D}^2(1)$ such that $r_1(P_1) = r_2(P_2) = c$ and $r_1(P_2) = r_2(P_1) = d$, we have $f(P) = r_1(P_i)$. According to Step 1, some voter i will dictate over each pair of alternatives. In particular, we can infer that voter 2 dictates over $\{a, b\}$.

Step 2: Let c be an alternative distinct from a and b. Then voter 2 dictates over $\{a, c\}$.

Let $\bar{\mathcal{D}}$ 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 $P_i \in \bar{\mathcal{D}}$ 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 $\bar{\mathcal{D}}$ by $abc...$, $acb....$, $bac...$, $bca....$, $cab...$, and $cba.....$ Here $abc...$ denotes the ordering where a, b and c are ranked first, second and third respectively. In view of Step 1, Step 2 is complete if we can show that $f(acb..., ca...) = c$. We proceed in a sequence of sub-steps.

Step 2(i): $f(acb..., bac...) = b$. Since 2 dictates over $\{a, b\}$ by assumption, we have $f(abc..., bac...) = b$. We have also shown in the proof of Step 1 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 $f(acb..., bac...) = b$.

Step 2(ii): $f(abc..., bca....) = b$. Suppose this is false. The efficiency of f implies $f(abc..., bca...) = a$. Observe that voter 2 with preference $bca...$ obtains b by reporting $bac...$ which is better than a according to $bca....$ Since $d(abc..., bca...) = 2$, Proposition 3 implies that there exists P_1 such that (i) $d(P_1, bca...) \leq 1$ (ii) $f(P_1, bca..) = b$ (iii) $f(P_1, bac...) = x$ implies either $x = a$ or x is worse than a according to $bca....$ The only candidates for P_1 are $abc...$ or $bac...$ or a transposition of some pair of alternatives in $bca...$ not involving b or c. If $P_1 = abc...$, then requirement (ii) is violated. If either $P_1 = bac...$, or P_1 is a transposition of a pair of alternatives in $bca..$ not involving b or c, then efficiency of f implies $f(P_1, bac...) = b$ violating requirement (iii). Therefore P_1 satisfying requirements (i), (ii) and (iii) does not exist. Hence $f(abc..., bca...) = b$.

Step 2(iii): $f(acb..., bca...) \in \{b, c\}$. Suppose this is false. Since f is efficient, the only possibility is $f(acb..., bca...) = a$. Then voter 1 in the profile $(abc..., bca...)$ gains by reporting $acb...$ (we know from Step 2(ii) that $f(abc..., bca...) = b$). Since $d(abc..., bca...) = 1$, Proposition 3 implies that there exists P_2 such that (i) $d(abc..., P_2) < 1$ (ii) $f(abc..., P_2) = a$ and (iii) if $f(acb..., P_2) = x$, then either $x = b$ or x is worse than b according to $abc....$ However (i) implies $P_2 = abc....$ Hence efficiency forces $x = a$ and (iii) is violated. Therefore P_2 satisfying requirements (i), (ii) and (iii) does not exist implying $f(acb..., bca...) \in \{b, c\}$.

Step 2(iv): If $f(acb..., cab...) = a$, then $f(acb..., cba...) = a$. The argument to establish this step is identical to the one in Step 1 and Step 2(i), viz. if the outcome in a profile whose distance is 1 is one voter's first-ranked alternative, then this remains the outcome when there is a transposition of a pair of alternatives not involving her first-ranked alternative of the other voter.

Step 2(v): $f(acb..., cba...) \neq a$. Suppose this is false. Since $f(acb..., bca...) \in \{b, c\}$, voter 2 in the profile $(acb..., cba...)$ gains by reporting $bca...$ instead of $cba....$ Observe $d(acb..., cba...) = 2$. Applying Proposition 3, there must exist P_1 such that (i) $d(P_1, cba...) < 2$ (ii) $f(P_1, cba...) \in \{b, c\}$ and (iii) if $f(P_1, bca...) = x$ then either $x = a$ or x is worse than a according to $cba...$, i.e $x \in A - \{a, b, c\}$. The only candidates for P_1 are (I) $cba...$ (II) $cab...$ (III) $bca...$ and (IV) a transposition of a pair of alternatives in $cba...$ not involving c or b. However the efficiency of f implies that in each of the cases (I)-(IV), $f(P_1, bca...) \in \{b, c\}$. Hence (iii) is violated in each case. Hence P_1 satisfying requirements (i), (ii) and (iii) do not exist. Hence $f(acb..., cba...) \neq a$.

We now establish Step 2. Efficiency and Steps $2(iv)$ and $2(v)$ imply $f(acb..., cab...) = c$. Therefore voter 2 dictates over the pair $\{a, c\}$.

Step 3: Let c be an alternative distinct from a and b. Then voter 2 dictates over $\{c, b\}$. In view of our earlier arguments, it will suffice to prove that $f(cba..., bca...) = b$. Suppose this is false. By efficiency, $f(cba..., bca...) = c$. By replicating the arguments of Steps $2(i)$, $2(ii)$ and $2(iii)$ with the roles of voters and alternatives interchanged, we can conclude that $f(cba..., bac...) = f(cab..., bca...) = c$ and $f(cab..., bac...) \in \{a, c\}$. We have already established in Step $2(i)$ that $f(acb..., bac...) = b$. Therefore voter 1 gains in profile $(acb..., bac...)$ by reporting $cab...$ instead of $acb...$. Observe that $d(acb..., bac...) = 2$. Applying Proposition 3, we conclude that there exists P_2 such that (i) $d(acb..., P_2) < 2$ (ii) $f(acb..., P_2) \in \{a, c\}$ and (iii) $f(cab..., P_2) = x$ implies either $x = b$ or x in worse than b according to $acb...$, i.e. $x \in A - \{a, b, c\}$. The only candidates for P_2 are (I) $acb...$ (II) $cab...$ (III) $abc...$ and (IV) a transposition of a pair of alternatives in $acb...$ not involving a or c. Notice that efficiency implies that in each case $f(acb..., P_2) \in \{a, c\}$ contradicting requirement (iii) for P_2. Therefore it is impossible to find P_2 satisfying (i), (ii) and (iii). Consequently $f(cba..., bca...) = b$ establishing Step 3.

Step 4: Let c, d be a pair of alternatives such that a, b, c, d are all distinct. Then voter 2 dictates over the pair $\{c, d\}$. In order to verify this claim, note that Step 2 implies that voter 2 dictates over $\{a, d\}$. Now applying Step 3, we conclude that voter 2 dictates over $\{c, d\}$.

Step 5: Voter 2 dictates over the pair $\{b, a\}$. Let $c \neq a, b$ (we are using the assumption that $|A| \geq 3$). According to Step 2, voter 2 dictates over $\{a, c\}$. Applying Step 3, voter 2 dictates over $\{b, c\}$ and applying Step 2 again, we conclude that voter 2 dictates over $\{b, a\}$.

Steps 1-5 establish Claim 1. ■

Proof of Claim 2: We assume without loss of generality that k is an integer strictly greater than 1 and that $f(P') = r_1(P'_2)$ whenever $d(P'_1, P'_2) \leq k$. Let P be a profile such that $d(P_1, P_2) = k + 1$. We will show that $f(P) = r_1(P_2)$.

Suppose $f(P) = x$. Let $\hat{P}_1$ be the ordering obtained by lifting x to the top of P_2 leaving all other alternatives unchanged. Formally,

1. $r_1(\hat{P}_2) = x$ and
2. for all $y, z \neq x$, $y\hat{P}_1 z \Leftrightarrow yP_2 z$.

Observe that $d(\hat{P}_1, P_2) = t$ where x is $t + 1^{th}$ ranked under P_2, i.e. $r_t(P_2) = x$. We claim that exactly one of the following two cases must hold.

Case A: $t < k + 1$

Case B: $P_1 = \hat{P}_1$.

Suppose $x \neq r_1(P_1)$. Let $w = r_1(P_1)$ so that wP_1x. Since x is efficient in the profile P, we must have xP_2w, i.e. the rank of w in P_2 is at least $t+2$. In order to transform P_1 to P_2, 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 $d(P_1, P_2) = k+1 \geq t+1$. This implies that Case A holds.

Now suppose $x = r_1(P_1)$. If the ranking of any pair of alternatives y, z distinct from x differs between P_1 and P_2 (i.e. yP_1z and zP_2y), then $d(\hat{P}_1, P_2) < d(P_1, P_2)$ and Case A holds again. The only remaining possibility is that P_1 and P_2 agree on all alternatives distinct from x. In this case $P_1 = \hat{P}_1$ 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: Since $d(\hat{P}_1, P_2) < k$, the induction hypothesis applies to the profile $(\hat{P}_1, P_2)$. Therefore $f(\hat{P}_1, P_2) = r_1(P_2)$. Suppose $r_1(P_2) = y \neq x$. Observe that in voter 1 in $(\hat{P}_1, P_2)$ gains by announcing P_1. Applying Proposition 3, we conclude that there exists $\hat{P}_2$ such that (i) $d(\hat{P}_1, \hat{P}_2) < d(\hat{P}_1, P_2)$ (ii) $f(\hat{P}_1, \hat{P}_2) = x$ and (iii) either $f(P_1, \hat{P}_2) = y$ or $y\hat{P}_1 f(P_1, \hat{P}_2)$. Now the hypothesis of Case A and (i) implies that $f(\hat{P}_1, \hat{P}_2) = r_1(\hat{P}_2)$. Then the induction hypothesis and (ii) implies $r_1(\hat{P}_2) = x$.

We claim that $d(\hat{P}_1, P_1) = d(P_1, P_2) - t$. First observe that by triangle inequality, $d(\hat{P}_1, P_) \geq d(P_1, P_2) - t$. We will show that the equality is exact. Observe that for any $y, z \neq x$, any one of the following three is true:

$$[yP_1z \text{ and } yP_2z] \Leftrightarrow [yP_1z \text{ and } y\hat{P}_1z] \tag{20}$$

$$[yP_1z \text{ and } zP_2y] \Leftrightarrow [yP_1z \text{ and } z\hat{P}_1y] \tag{21}$$

e

$$[zP_1y \text{ and } yP_2z] \Leftrightarrow [zP_1y \text{ and } y\hat{P}_1z] \tag{22}$$

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_1 and P_2, then they also agree (disagree) in P_1 and $\hat{P}_1$. Also observe that any alternative that is ahead of x in P_2 will be below x in P_1 and vice versa; otherwise $x \notin PE(P_1, P_2)$. Moreover, for any such alternative z, the relative ranking of x and z is the same in P_1 and $\hat{P}_1$. Summarizing we have, $[xP_1z \text{ and } zP_2x] \Rightarrow [xP_1z \text{ and } x\hat{P}_1z]$. Since $|\{z|zP_2x\}| = t$, equation (1), (2) and (3) together with the last argument imply that $d(\hat{P}_1, P_1)$ can atmost be $(k+1) - t$. Hence $d(\hat{P}_1, P_) \leq d(P_1, P_2) - t$. Combining this last inequality with the inequality above we have $d(\hat{P}_1, P_) = d(P_1, P_2) - t$.

By the triangle inequality, $d(P_1, \hat{P}_2) \leq d(P_1, \hat{P}_1) + d(\hat{P}_1, \hat{P}_2)$. Since we have established that $d(P_1, \hat{P}_1) = d(P_1, P_2) - t$ and $d(\hat{P}_1, \hat{P}_2) < t$ by assumption, we have $d(P_1, \hat{P}_2) < d(P_1, P_2) = k+1$. The induction hypothesis therefore applies to the profile $(P_1, \hat{P}_2)$, i.e

$f(P_1, \hat{P}_2) = r_1(\hat{P}_2)$. But we have already established that $r_1(\hat{P}_2) = x$. But in order for requirement (iii) for $\hat{P}_2$ to hold, we must either have $x = y$ or $y\hat{P}_1 x$. Since $x \neq y$ by assumption and $x = r_1(\hat{P}_1)$ by construction, neither can hold and we have a contradiction. Therefore $x = y$ must hold, so that $f(P) = r_1(P_2)$. This completes the argument for Case A.

Case B: Suppose that (P_1, P_2) are such that $d(P_1, P_2) = k+1$, $r_1(P_1) = x$, P_1 and P_2 agree on all alternatives other than x and $f(P_1, P_2) = x \neq r_1(P_2) = y$. It is clear that $x = r_{k+2}(P_2)$. Using Claim 1, we can also assume that $k \geq 1$; otherwise $d(P_1, P_2) = 1$ which has been dealt with in Claim 1. Construct P_2' by transposing x with the alternative immediately above it in P_2. Since $k \geq 1$ and $x = r_{k+2}(P_2)$, $r_1(P_2') = y$. We must also have $d(P_1, P_2') = k$ and $f(P_1, P_2') = y$ by the induction hypothesis. Since yP_2x, it follows from Proposition 3, that there exists P_1' such that (i) $d(P_1', P_2) < k + 1$ (ii) $f(P_1', P_2) = y$ and (iii) $f(P_1', P_2') = w$ implies either $w = x$ or xP_2w.

We first claim that $r_1(P_1') \neq x$. If this were true, then $d(P_1', P_2) > k + 1$. To see this observe that P_1' and P_2 agree on the rankings of all alternatives other than x. Hence, for all P_1 such that $r_1(P_1) = x$, $d(P_1, P_2) \geq k + 1 = d(P_1', P_2)$. So let $z = r_1(P_1')$. If xP_2z, then, it would require at least $k + 1$ transpositions from P_2 for z to be first ranked, i.e. $d(P', P_2) \geq k + 1$. This implies that zP_2x. The construction of P_2' implies that $zP_2'w$. There are two cases to consider. First, let $w \neq x$. Since $zP_1'w$, $f(P_1', P_2') = w$ contradicts the assumption that f is efficient at profile (P_1', P_2'). 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_2. This means that x is ranked immediately above z in P_2'. If there exists an element between z and x under P_2 the first case applies. If z is ranked immediately above x in P_2, then $z = r_{k+1}(P_2)$. In P_1' z is the top-ranked element. Therefore, the minimum distance between P_1' and P_2 is k, i.e., $d(P_1', P_2) \geq k$. But as mentioned above $d(P_1', P_2)$ has to be less than $k + 1$. Therefore the only allowable case is $d(P_1', P_2) = k$. Since $z = r_{k+1}(P_2)$ and $d(P_1', P_2) = k$, it must be the case that for all $x, y \neq z$,

$$[xP_1'y \Leftrightarrow xP_2y] \tag{23}$$

. Otherwise, $d(P_1', P_2) > k$. Equation (4) together with the fact that $k \geq 1$ implies that there exists a $v \in A$ such that,

$$vP_1'x \text{ and } vP_2x \tag{24}$$

But then $x \notin PE(P_1', P_2)$. Hence P_1' satisfying requirements (i), (ii) and (iii) cannot exist. Therefore $f(P_1, P_2) = r_1(P_2)$ completing the argument for Case B. ■

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.

6 Conclusion

In this paper we have explored the problem of mechanism design in a voting environment with two voters where a voter's belief about the type of the other voter 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).

In future research we hope to extend our analysis to an arbitrary number of voters, to other notions of correlation and to understand further, 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.