

Choosing from a large tournament

Mark Fey

Received: 11 March 2003 / Accepted: 18 September 2007 / Published online: 12 October 2007
© Springer-Verlag 2007

Abstract A tournament can be viewed as a majority preference relation without ties on a set of alternatives. In this way, voting rules based on majority comparisons are equivalent to methods of choosing from a tournament. We consider the size of several of these tournament solutions in tournaments with a large but finite number of alternatives. Our main result is that with probability approaching one, the top cycle set, the uncovered set, and the Banks set are equal to the entire set of alternatives in a randomly chosen large tournament. That is to say, each of these tournament solutions almost never rules out any of the alternatives under consideration. We also discuss some implications and limitations of this result.

1 Introduction

A tournament on a set of alternatives is a complete and asymmetric binary relation. While tournaments arise in many areas, their importance in social choice theory stems from the fact that the majority preference relation of an odd number of voters with linear preference orders is always a tournament. Thus, aggregating the preferences in a society can be viewed as choosing from a tournament.

When there is no outcome that is majority preferred to every other outcome (a Condorcet winner), there is no straightforward notion of a “best” alternative. A large literature is devoted to the question of designing some principles for selecting such a set of “best” alternatives. These tournament solutions include the top cycle set (Schwartz 1972; Miller 1977), the uncovered set (Miller 1980), the Banks set (Banks 1985), the minimal covering set (Dutta 1988), the tournament equilibrium set (Schwartz 1990),

M. Fey (✉)
Department of Political Science, Harkness Hall,
University of Rochester, Rochester, NY 14627, USA
e-mail: markfey@mail.rochester.edu

and others. Axiomatizations of and connections between these sets have been established by Fishburn (1977); Moulin (1986); Laffond et al. (1995), and Laslier (1997).

In this paper, we investigate the size of several standard tournament solutions. This question is important because if a tournament solution contains many alternatives, it does not do much to narrow down the choice. Indeed, this point has driven much of the research on tournament solutions, as theorists have strived to devise ever smaller solutions. For good reason, it is taken as given in the literature that “smaller is better.”

While authors have frequently considered whether one tournament solution is always smaller or larger than another solution; they have not, for the most part, addressed the absolute size of a given tournament solution.¹ It is easy to find examples of tournaments with a large number of alternatives for which several common tournament solutions are small. And it is just as easy to give examples in which these sets are equal to the entire set of alternatives. Given that both conclusions are possible, what can be said about the “typical” large tournament? We show that with probability approaching one, the top cycle set, the uncovered set, and the Banks set are equal to the entire set of alternatives in a randomly chosen tournament. In other words, these tournament solutions almost never narrow the set of social choices.

Similar questions have been asked regarding majority preference in a continuous (multidimensional) setting with spatial individual preference, although with somewhat different results. Plott (1967) showed that the set of Condorcet winners is almost always empty. McKelvey (1976) showed that the top cycle set is almost always the whole space of alternatives and later established that is not the case for the uncovered set (McKelvey 1986). In addition, De Donder (2000) showed through simulations that several tournament solutions including the uncovered set and the bipartisan set give sharp predictions in a setting with spatial preferences. Our result on random tournaments, with no additional assumptions on the preferences of individuals, stands in contrast to these latter results for the standard spatial model of voting with spatial preferences. One way to understand this difference is that the set of large tournaments is generically inconsistent with the assumption of spatial preferences.

In the case of tournaments on a finite set of alternatives, initial results have been obtained by Bell (1981). As he discussed, if all tournaments are equiprobable then earlier results from graph theory imply that the top cycle is almost surely the whole set of alternatives. He then showed that if a random tournament is obtained by independently selecting a linear preference order for each voter and forming the resulting majority preference relation, the same result holds.

Our main result is that the probability that every alternative is in the Banks set in a random tournament goes to one as the number of alternatives goes to infinity. This implies that several other common tournament solutions such as the top cycle set and the uncovered set also fail to place any additional constraints on social choices. Moreover, by using an axiomatization of the uncovered set by Moulin (1986), we are able to show that this negative result holds for any tournament satisfying three axioms. We also point out that some tournament solutions, such as the Copeland winner, do

¹ One exception is Miller (1980), who points out that the top cycle may be the whole set of alternatives, in which case, “the set of possible or desirable decisions is not narrowed down at all.” (p. 71)

not have this negative property and conclude with some directions for further work on additional tournament solutions.

The format of the paper is straightforward. We introduce the notation and definitions in Sect. 2. We prove our main result in Sect. 3. In Sect. 4, we consider some other tournament solutions and suggest some avenues for future work.

2 Notation and definitions

2.1 Tournaments

We are interested in choosing from a large but finite tournament. So let n denote the number of *alternatives* in the set $X = \{x_1, \dots, x_n\}$.² Let T be a complete and asymmetric binary relation on X . For example, T could represent the majority preference relation of an odd number of voters with linear preferences. In any case, we say that T is a *tournament* on X , which we sometimes refer to as a tournament of order n . In particular, for any pair of distinct alternatives a and b , exactly one of aTb or bTa holds. As usual, if aTb holds, we say alternative a “beats” alternative b . For a subset $Y \subseteq X$ of alternatives, we write aTY if, for all $y \in Y$, aTy . If $V = T \cap (Y \times Y)$ for some $Y \subseteq X$, then V is a tournament on Y and V is a *subtournament* of X . We say that V is the *restriction* of T to Y , which we denote by $V = T \upharpoonright Y$.

For a fixed alternative $x \in X$, the preferred-to set is defined as

$$T(x) = \{y \in X \mid yTx\}$$

and, similarly,

$$T^{-1}(x) = \{y \in X \mid xTy\}.$$

It follows that $X = T(x) \cup \{x\} \cup T^{-1}(x)$. The *Copeland score* of x , denoted as $s(x)$, is the number of alternatives that x beats. Thus, $s(x) = |T^{-1}(x)|$. Similarly, the number of alternatives that beat x is given by $t(x) = |T(x)| = n - s(x) - 1$.

2.2 Tournament solutions

We are interested in several sets determined by the majority preference relation on X called *tournament solutions*. Formally, a tournament solution is a correspondence S that, for any tournament T , selects a nonempty subset of X .³ Some well-known tournament solutions are the following.⁴ The *top cycle* of T , denoted as $TC(T)$, is the set of alternatives that directly or indirectly beat every other alternative. The *uncovered set* of T , $UC(T)$, is the set of alternatives that are not covered by any other alternatives,

² That is, $|X| = n$, where $|\cdot|$ denotes the cardinality of a set.

³ Thus, the set of Condorcet winners is not a tournament solution because it may be empty.

⁴ Formal definitions can be found in the survey by Laslier (1997).

where x covers y if xTy and for all $z \in X$, yTz implies xTz . Further, we can consider iterations of the uncovered set. Let $UC^1(T) = UC(T)$ and define the solutions $UC^k(T)$ inductively by $UC^{k+1}(T) = UC(T \mid UC^k(T))$. The *ultimate uncovered set*, denoted by $UC^\infty(T)$, is the set for which no further reduction can occur (Miller 1980). Formally, $UC^\infty(T) = UC^k(T)$ for k such that $UC^{k+1}(T) = UC^k(T)$.

In order to define the Banks set, we first define a *chain*. In a tournament T , a chain is a (nonempty) subset H of X such that T restricted to H is transitive. That is, for all $x, y, z \in H$, xTy and yTz imply xTz . For this reason, a chain is also called a transitive subtournament of T (Moon 1968). If $|H| = k$ holds for a chain H , we say that it is a chain of order k . A *maximal chain* is a chain that is not a subset of some other chain. For a chain H , an alternative $x \in H$ is *top-ranked* in H if it beats every other alternative in H . Clearly, for each chain, such an alternative always exists. The Banks set is the set of alternatives that are top-ranked in maximal chains (Banks 1985). Formally,

$$B(T) = \{y \in X \mid y \text{ is top-ranked in some maximal chain } M\}.$$

We can also inductively define $B^k(T)$ and $B^\infty(T)$ in the same way as we defined $UC^k(T)$ and $UC^\infty(T)$. The fundamental relationship between these concepts is given by the inclusions $B(T) \subseteq UC(T) \subseteq TC(T)$ and, from this, it follows that $B^k(T) \subseteq UC^k(T)$ and $B^\infty(T) \subseteq UC^\infty(T)$.

2.3 Random tournaments

In the next section, we prove results that apply to almost all large tournaments. Specifically, we show that the probability that a random tournament has a particular property goes to one as the number of alternatives goes to infinity. To do so, we must define precisely our notion of a random tournament. For each integer $n \geq 3$, let \mathcal{T}_n denote the set of possible tournaments on n alternatives. It is easy to see that this set contains $2^{\binom{n}{2}}$ distinct tournaments. We take this set to be the sample space from which we draw a random tournament.

The probability model that we investigate in this paper assigns each tournament in \mathcal{T}_n the same probability, namely, $2^{-\binom{n}{2}}$. It is useful to note that an equivalent formulation of this probability model is that a random tournament $T \in \mathcal{T}_n$ is obtained by choosing independently, for each pair of alternatives $x, y \in X$, $x \neq y$, either xTy or yTx with equal probability.

Alternatively, we could suppose that a (linear) preference order is randomly chosen for each voter. In particular, each voter is equally likely to have one of the $n!$ possible preference orders. The resulting majority preference relation determined by these assignments would then be a random tournament in the model. This is the approach investigated by Bell (1981). We conjecture that similar results would hold in this model, but we leave this to future work.

3 The main result

In this section, we present our main result and discuss some of its implications. Our main result states that with probability approaching one, the Banks set is equal to the entire set of alternatives in a randomly chosen large tournament. As is standard in the literature of random graphs (Bollobás 2001), we say a property \mathcal{Q} holds for almost all tournaments if the probability that a random tournament has property \mathcal{Q} goes to one as the number of alternatives, n , goes to infinity. In what follows, we denote this probability by $P[\mathcal{Q}]$.

Theorem 1 *In almost all tournaments, $B(T) = X$.*

It is important to note the order of quantifiers in this theorem. It does *not* state that in every tournament, almost all alternatives are in the Banks set. Rather, it says that every alternative is in the Banks set in *almost all* tournaments.

Before proceeding with the proof of the theorem, we present several useful lemmas. The first lemma is a key element in our proof. It presents the Chernoff inequality for a binomial random variable (Chernoff 1952). This inequality is a strong upper bound on the probability of obtaining a very large number of successes in a binomial random variable. Specifically, we will use a version of the Chernoff inequality due to Okamoto (1958).⁵

Lemma 1 (Okamoto 1958; Theorem 1) *If X is a random variable with a binomial distribution $B(n, p)$ and c is a positive constant, then*

$$P[X - np \geq cn] < e^{-2nc^2}.$$

Our next lemma is a simple implication of Markov’s inequality.

Lemma 2 *Let Y_1, Y_2, \dots be a sequence of nonnegative, integer-valued random variables. If $\lim_{n \rightarrow \infty} E[Y_n] = 0$, then $\lim_{n \rightarrow \infty} P[Y_n = 0] = 1$.*

Proof of Lemma 2 By Markov’s inequality, $P[Y_n \geq 1] \leq E[Y_n]$. The result follows. □

Moon (1968) proved that every tournament of order n contains a chain of order at least $\lfloor \log_2 n \rfloor + 1$, where $\lfloor x \rfloor$ is the largest integer less than or equal to x . The next lemma is an extension of this fact.

Lemma 3 *Let $k = \lfloor \log_2 n \rfloor$. Then every tournament of order n contains at least $\lfloor 2^{k-1}/k \rfloor$ disjoint chains of order k .*

Proof of Lemma 3 Let T be a tournament of order n . Let $k = \lfloor \log_2 n \rfloor$ and $r = \lfloor 2^{k-1}/k \rfloor$. Then $n \geq 2^k$ and $rk \leq 2^{k-1}$.

Moon (1968) showed that there is a chain H in T with at least $k + 1$ elements. Pick any k element subset of H and call it H_1 . Now consider the tournament $T_1 = T \setminus H_1$

⁵ See also Johnson et al. (1992).

of order $n - k$. Depending on whether $\lfloor \log_2(n - k) \rfloor$ equals k or $k - 1$, there is a chain H in T_1 with $k + 1$ or k elements. Again, select any k element subset and call it H_2 . By construction, H_1 and H_2 are disjoint. Now let $T_2 = T_1 \setminus H_2$ and repeat this selection to form H_3 .

How many times can we repeat this procedure and be assured of finding a chain of order at least k ? After r times, the remaining tournament T_r has $n - rk$ elements. From the above,

$$n - rk \geq 2^k - 2^{k-1} = 2^{k-1}.$$

Therefore, T_r (and all earlier ones) have a k element chain. Thus, this procedure yields H_1, H_2, \dots, H_r , which are disjoint chains of order k . □

Our final lemma gives a useful characterization of alternatives contained in the Banks set.

Lemma 4 $x \in B(T)$ if and only if there exists some chain $H \subseteq T^{-1}(x)$ such that, for every $y \in T(x)$, hTy holds for some $h \in H$.

Proof of Lemma 4 (\Rightarrow) Let $H = M \setminus \{x\}$, where M is a maximal chain with x top-ranked. Then $H \subseteq T^{-1}(x)$ and the remainder of the condition follows from maximality of M .

(\Leftarrow) Assume that a chain H exists with the above property. Then $H' = H \cup \{x\}$ is a chain in X . Thus, it must be contained in some maximal chain M . Suppose that M contains some $y \in T(x)$. Then transitivity of M implies yTh for all $h \in H$. This is a contradiction. So $M \subseteq T^{-1}(x) \cup \{x\}$. This implies that x is top-ranked in M , so $x \in B(T)$. □

We are now ready to prove our main result.

Proof of Theorem 1 For each tournament T in \mathcal{T}_n , let

$$Y(T) = |\{y \in X \mid y \in X \setminus B(T)\}|.$$

Then $P[Y(T) = 0]$ is the probability that a random tournament has the property that $B(T) = X$. By Lemma 2, in order to show that almost all tournaments have $B(T) = X$, it suffices to show that $E(Y) \rightarrow 0$ as n becomes large.

We can write Y as $Y = \sum_{i=1}^n Y_i$, where

$$Y_i(T) = \begin{cases} 1 & \text{if } x_i \notin B(T), \\ 0 & \text{if } x_i \in B(T). \end{cases}$$

In this case, $E[Y_i] = E[Y_j]$ and therefore $E[Y] = n E[Y_i]$, where $E[Y_i]$ is just the probability that x_i is not in the Banks set of a random tournament. We will now construct an upper bound for this probability, which we denote as $P[x_i \notin B(T)]$.

As noted in the final part of Sect. 2, choosing a random tournament from \mathcal{T}_n is equivalent to constructing a tournament by independently choosing xTy or yTx with

probability $1/2$, for each pair of alternatives x and y . As these choices are independent, the number of alternatives in a random tournament that beat x_i , $t(x_i)$ is a binomial random variable with distribution $B(n - 1, 1/2)$. Now fix a constant $0 < c < 1/2$ and let $q_n = 1 - e^{-2(n-1)c^2}$. Then by Lemma 1, with probability greater than q_n , $t(x_i) \leq (1/2 + c)(n - 1)$.

We begin our calculation of $P[x_i \notin B(T)]$ with the following trivial bound:

$$P[x_i \notin B(T)] \leq (1 - q_n)(1) + q_n P[x_i \notin B(T) \mid t(x_i) \leq (1/2 + c)(n - 1)].$$

In order to evaluate the last term in this expression, we suppose that $t(x_i) \leq (1/2 + c)(n - 1)$ in what follows. Let $L = \log_2(1/2 - c)(n - 1)$ and $k = \lfloor L \rfloor$. Viewing $T^{-1}(x_i)$ as a tournament by itself, we know by Lemma 3 that $T^{-1}(x_i)$ contains at least $\lfloor 2^{k-1}/k \rfloor$ disjoint chains of order k . In particular, there are at least $2^{k-2}/k$ such chains. For a given alternative $x \in T(x_i)$ and a given chain $H \subseteq T^{-1}(x_i)$ of order k , the probability that xTH does not hold is $1 - 2^{-k}$. Thus, the probability that xTH does not hold for any $x \in T(x_i)$ is

$$\begin{aligned} (1 - 2^{-k})^{t(x_i)} &> (1 - 2^{-(L-1)})^{(1/2+c)(n-1)} \\ &= \left[\left(1 - \frac{2/(1/2 - c)}{n - 1} \right)^{n-1} \right]^{1/2+c} \\ &\rightarrow e^{-2\frac{1+2c}{1-2c}} \quad \text{as } n \rightarrow \infty. \end{aligned}$$

As $0 < c < 1/2$, there is a constant $0 < d < 1$ such that, for sufficiently large n , the probability that xTH does not hold for any $x \in T(x_i)$ is greater than $1 - d$. Thus, the probability that at least one $x \in T(x_i)$ beats a given H is less than d .

So to complete the calculation of $P[x_i \notin B(T) \mid t(x_i) \leq (1/2 + c)(n - 1)]$, we observe that by Lemma 4, this probability is less than or equal to the probability that the above holds for each of the $2^{k-2}/k$ disjoint chains of order k in $T^{-1}(x_i)$. That is,

$$P[x_i \notin B(T) \mid t(x_i) \leq (1/2 + c)(n - 1)] \leq d^{2^{k-2}/k} \leq d^{\frac{(1/2-c)(n-1)}{8 \log_2(1/2-c)(n-1)}},$$

using the fact that $2^{k-2}/k \geq 2^{L-3}/L$.

We conclude that

$$\begin{aligned} E[Y] &= n P[x_i \notin B(T)] \\ &\leq n(1 - q_n) + nq_n P[x_i \notin B(T) \mid t(x_i) \leq (1/2 + c)(n - 1)] \\ &\leq ne^{-2(n-1)c^2} + n \left(1 - e^{-2(n-1)c^2} \right) d^{\frac{(1/2-c)(n-1)}{8 \log_2(1/2-c)(n-1)}}. \end{aligned}$$

As this bound goes to zero as n goes to infinity, the proof is complete. □

Using the nestedness of the top cycle, the uncovered set, and the Banks set, the following corollary is an immediate consequence of Theorem 1.

Corollary 1 *In almost all tournaments, $TC(T) = UC(T) = UC^\infty(T) = B^\infty(T) = X$.*

The last claim follows from the fact that $B^k(T) = X$ implies $B^{k+1}(T) = X$.

More generally, we can use Theorem 1 to evaluate the size of any tournament solution that satisfies certain axioms. Moulin (1986) proved that the uncovered set is the finest tournament solution that satisfies Condorcet consistency, neutrality, and expansion.⁶ It then follows that any tournament solution satisfying these three axioms will equal the entire set of alternatives in almost all tournaments.

4 Other tournament solutions

In the previous section, we have shown that a number of common tournament solutions almost surely fail to reduce the number of choices. But the literature contains many other tournament solutions that are not covered by our main result. In this section, we briefly address these other tournament solutions.

In order to show that our main result does not hold for all tournament solutions, consider the set of *Copeland winners*, $C(T)$. This is the set of alternatives with the highest Copeland score. If $C(T) = X$ were to hold, then every alternative must have the same Copeland score. It is easy to see that in a tournament of order n , the sum of the Copeland scores must equal $\binom{n}{2} = n(n-1)/2$. So if $C(T) = X$, then $s(x) = (n-1)/2$ for every $x \in X$. This is obviously impossible if n is even. If n is odd, as the Copeland score of a given alternative is binomially distributed with mean $(n-1)/2$, it follows that the probability that it occurs in a random tournament of odd order goes to zero as n gets large. This simple analysis shows that in almost all tournaments, $C(T) \neq X$.⁷ Although this analysis is useful to demonstrate that Theorem 1 does not apply to all tournament solutions, we must caution against viewing this result as claiming supremacy for the Copeland solution. Indeed, it is easy to show that the Copeland score of a Copeland winner in a large tournament will almost always be “close” to the Copeland score of a nonwinner. To put this more formally, if $s_{(1)}(T)$ is the largest Copeland score of T and $s_{(2)}(T)$ is the second largest Copeland score of T , then for every constant $c > 0$, $P[s_{(1)} - s_{(2)} < cn]$ goes to zero as n goes to infinity. Thus, any claim that an alternative with maximal Copeland score is significantly “better” than an alternative with the second highest Copeland score is unpersuasive in large tournaments.⁸

More generally, the size of other tournament solutions in large tournaments remains an open question. Some initial investigations suggest that the minimal covering set (Dutta 1988) is almost always equal to the whole set of alternatives, but we have not yet proved this conjecture. A resolution of this conjecture could aid in investigating whether refinements of the minimal covering set such as the bipartisan set (Laffond et al. 1993) are almost always equal to the whole set of alternatives. Likewise, further work remains to be done on whether refinements of the Banks set such as the tournament equilibrium set (Schwartz 1990) also have this property.

⁶ See Laslier (1997) for the formal definition and discussion of these axioms.

⁷ This result casts a new light on some claims made about the attributes of the Copeland tournament solution. In particular, Moulin (1986) argued that Copeland winners may be poor choices because, for some tournaments, $C(T)$ is outside $TC(UC(T))$. However, it is immediate from our results and the fact that $UC^\infty(T) \subseteq TC(UC(T))$ that the former is almost always a proper subset of the latter.

⁸ However, see Grofman et al. (1987) for arguments in favor of the Copeland rule in a spatial setting.

As a final avenues for future work, it would be interesting to modify our conception of a random tournament to that of a random preference order, as mentioned at the end of Sect. 2. We conjecture that our results will still hold in this model. The complication raised by this alternative assumption is that we lose independence of the majority preference across pairs of alternatives. To deal with this problem, Bell (1981) used a result from Niemi and Weisberg (1968) that approximates the votes from a random preference order with a multivariate normal. It may be possible to apply a similar technique to address the question of choosing from a large tournament raised here.

Acknowledgments My thanks to John Duggan for many helpful discussions and Steve Callander and Ken Shotts for useful feedback.

References

- Banks JS (1985) Sophisticated voting outcomes and covering relation. *Soc Choice Welfare* 1:295–306
- Bell CE (1981) A random voting graph almost surely has a Hamiltonian cycle when the number of alternatives is large. *Econometrica* 49(6):1597–1603
- Bollobás B (2001) *Random graphs*, 2nd edn. Cambridge University Press, Cambridge
- Chernoff H (1952) A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann Math Stat* 23:493–507
- De Donder P (2000) Majority voting solution concepts and redistributive taxation. *Soc Choice Welfare* 17(4):601–627
- Dutta B (1988) Covering sets and a new Condorcet choice correspondence. *J Econ Theory* 44(1):63–80
- Fishburn PC (1977) Condorcet social choice functions. *SIAM J Appl Math* 33(3):469–489
- Grofman B, Owen G, Noviello N, Glazer A (1987) Stability and centrality of legislative choice in the spatial context. *Am Pol Sci Rev* 81(2):539–553
- Johnson NL, Kotz S, Kemp AW (1992) *Univariate discrete distributions*, 2nd edn. Wiley, New York
- Laffond G, Laslier J, Le Breton M (1993) The bipartisan set of a tournament game. *Games Econ Behav* 5(1):182–201
- Laffond G, Laslier JF, Le Breton M (1995) Condorcet choice correspondences: a set-theoretical comparison. *Math Soc Sci* 30(1):23–35
- Laslier JF (1997) *Tournament solutions and majority voting*. Springer, Berlin
- McKelvey RD (1976) Intransitivities in multidimensional voting models and some implications for agenda control. *J Econ Theory* 12:472–482
- McKelvey RD (1986) Covering, dominance, and institution-free properties of social choice. *Am J Pol Sci* 30:283–314
- Miller NR (1977) Graph-theoretical approaches to the theory of voting. *Am J Pol Sci* 21(4):769–803
- Miller NR (1980) A new solution set for tournaments and majority voting. *Am J Pol Sci* 24(1):68–96
- Moon JW (1968) *Topics on tournaments*. Winston, New York
- Moulin H (1986) Choosing from a tournament. *Soc Choice Welfare* 3(4):271–291
- Niemi RG, Weisberg HF (1968) A mathematical solution for the probability of the paradox of voting. *Behav Sci* 13(4):317–323
- Okamoto M (1958) Some inequalities relating to the partial sum of binomial probabilities. *Ann Inst Stat Math* 10:29–35
- Plott CR (1967) A notion of equilibrium and its possibility under majority rule. *Am Econ Rev* 57:787–806
- Schwartz T (1972) Rationality and the myth of the maximum. *Nous* 6:97–117
- Schwartz T (1990) Cyclic tournaments and cooperative majority voting: a solution. *Soc Choice Welfare* 7(1):19–29