A taxonomy of runoff methods

Bernard Grofman*

Department of Political Science and Center for the Study of Democracy, University of California, Irvine, CA 92697-5100, USA

Keywords:
Electoral system
Runoff
Two round ballot
Alternative vote

ABSTRACT

We look at ways of classifying runoff methods in terms of characteristics such as number of rounds, rules used to determine which candidates advance to the next round, and rules which determine the final winner. We also compare runoffs and so-called instant runoffs such as the alternative vote.

© 2008 Published by Elsevier Ltd.

1. Introduction

Most voting rules can (barring ties) always be completed in a single round. By a runoff method we will simply mean an election which may require more than a single round of balloting. How many ballots will be required will depend upon the results of the first (and subsequent) rounds and on the specific runoff rules.

In this essay we begin by providing a typology of runoffs that is based on the conditions that will trigger a \((k+1)\)th ballot. We restrict ourselves to rules that require voters simply to list their first preferences, and we have limited ourselves to runoff elections designed to select a single candidate from among some set. We discuss five distinctions among such runoff rules: (1) the maximum number of rounds of voting that might take place; (2) how many candidates can be expected to advance to the second round (and subsequent rounds, if any); (3) the threshold of exclusion which determines the worst-case scenario for what it takes for a candidate to advance to the next round; (4) whether or not advancement procedures are invariant (constant) across all rounds; and (5) whether or not there is some kind of distribution requirement that opens up the possibility that the popular vote winner may not automatically be selected (such as rules for regional balance in vote distribution (as in Nigeria, Indonesia, or Kenya) or rules such as the weighted voting rule in use in the U.S Electoral College).

In the second section of the essay we consider parallels between different runoff rules eliciting a single \(x\) vote at each stage of the balloting and various voting procedures that require voters to provide a (full) rank-ordering of their preferences, or which ask voters for information beyond their first choice. Here we offer an important theorem, namely that, for every runoff rule where voters give first preferences, we can construct a voting rule using rank-ordered ballots to which it is mathematically isomorphic. If we assume votes vote ‘sincerely’, and if we neglect issues of strategic incentives for candidates to present themselves for election, the runoff rule and its rank-ordered counterpart will give the same results. However, we also show that the converse is not true; i.e., there exist voting rules using rank-ordered ballots which have no counterpart among the class of runoff rules eliciting a single \(x\) for first choice at each ballot stage.

In our conclusion we argue for the neglected importance of distinguishing among the properties of the many different types of rules that select a single candidate from a larger set, and argue that the collapsing of the choice continuum into one between majority/plurality systems on the one hand, and PR systems, on the other, misses much of importance about the incentives structuring political competition.

---

* An earlier version of this essay was given at the Conference on ‘Plurality and Multi-round Elections’, University of Montreal, June 16–19, 2006. I am indebted to Clover Behrend-Gethard for bibliographic assistance. Some ideas discussed here are presented in a much more preliminary fashion in Grofman and Lewis-Beck (2005).

* Tel.: +1 949 824 5143; fax: +1 949 824 2094.

E-mail address: bgrofman@uci.edu

0261-3794/$ – see front matter © 2008 Published by Elsevier Ltd.
doi:10.1016/j.electstud.2008.04.007
2. Distinguishing types of runoffs

Perhaps the single most important distinction among voting rules that allows us to distinguish runoff rules from others is in terms of the maximum number of rounds of voting that might be required. For example, the French double ballot system at both the presidential and legislative level never requires more than two rounds of election, and this is true for most other runoff methods in common use, including the runoff rules used for partisan primary elections in the many states in the southern United States. However, there are runoff methods that may require more than two rounds, such as the multi-round sequential elimination procedure (MRSE) that is used for selecting party leaders in the U.S. House of Representatives (McGann et al., 2002a,b; Grofman et al., 2002) and in party elections elsewhere in the world (see e.g., McSweeney, 1999). Similarly, while exact details have varied, the methods that have used for electing the Pope for the past centuries may, in principle, require multiple ballots, and usually have done so (Tobin, 2003).

Another important (and related) way to distinguish among runoff methods is in terms of the conditions that will trigger a \((k + 1)\)th ballot round. For any such condition, we can determine its threshold of exclusion; i.e., the smallest share of the vote received by a candidate on the \(k\)th round sufficient to guarantee that the \(k\)th round be the last round. The most common form of runoff, we believe, is the majority runoff. In a simple majority runoff each round has a threshold of exclusion of \(1/2\). In other words, a candidate has to have a majority of the votes cast in order to win. Another common runoff rule requires a 40% vote share for a candidate to be declared victorious (e.g., Costa Rica for presidential elections). One intriguing rule for deciding when a runoff is needed was proposed over a decade ago by Shugart and Taagepera (1994). Let \(v_1\) and \(v_2\) represent the share of the vote received by the plurality winner and the next most popular candidate. Their double complement rule proposes that a runoff be held if and only if \(2(v_1 - 50) > v_2 - 50\). Of course, if \(v_1 > 50\), no runoff is needed since this condition will never be met under those circumstances. If, say \(v_1 = 40\), then we would have a runoff if \(v_2 > 30\).

However, even among runoff methods that have the same threshold of exclusion and have only two rounds, a further dimension of difference is in the rule for determining how many and which candidates advance to the second round. Here a variety of different rules apply in majority runoff elections around the world, including the ‘top three’ rule used in the U.S. Electoral College in the absence of a candidate winning an Electoral College majority, and the ‘top two’ rule used in runoff elections in the U.S. to select political party nominees in party primaries in many Southern states. In France, in the parliamentary elections, only those candidates who have received at least one-eighth of the vote are eligible to advance; while in French presidential elections, as in most southern primaries in the U.S., only the ‘top two’ candidates may advance. We can also have quite complex rules for advancement, such as advance the set of candidates with the most votes who, between them, have, say, 75% of the total vote. In this advancement rule, if there is a \(k\)th round, the number of candidates who advance to that round is not predetermined, but is contingent on the actual distribution of voter preferences.

For runoffs which might under certain conditions continue past a second round, another useful distinction is between what we will call constant rules and non-constant rules. For runoff methods which potentially might extend past a second round, a constant runoff rule is one where both the threshold of exclusion and the rule for determining which candidates advance to the next round does not vary by round of balloting. A non-constant rule is one where one or both of these statements is not true.

It is, we think, relatively rare for runoff procedures to have constant rules. For double ballot rules it is common to have the second ballot decided by simple plurality if the rule for advancement does not restrict the field to two candidates. But we can have various kinds of dependencies of rule on ballot round. Consider the recent change in the voting rules used in France, rather than the double ballot system in France.

---

1 In France, double ballot methods have also been used at the regional or local level (Jérome and Jérome-Speziari, 2005).
2 In MRSE elections, a majority of the vote is required for election. If no candidate receives a majority of the votes, and a new ballot is conducted. This process continues until some candidate gets a majority of votes cast. Since we ultimately come down to two candidates this method is what social choice theorists call pairwise; i.e., it always results in the selection of a winner and cannot result in deadlock (harring, of course, a tie in the penultimate round among the second place and one or more lower-placed candidates – but even such a situation can be resolved by some lottery-like tie breaking rule).
3 When a unanimous vote was required to elect a Pope, this does not give us what we previously referred to as a decisive rule, so in principle there might be endless numbers of rounds of balloting. The fact that the cardinals were locked in until they reached agreement tended to foster an eventual consensus – albeit it did not guarantee it; the Middle Ages saw schisms within the Roman Catholic Church and competing claimants to the papacy (Tobin, 2003).
4 The threshold of exclusion (Loosemore and Hanby, 1971) can be taken to be a measure of how proportional a voting rule is when viewed from a purely theoretical perspective. For a national legislature, the threshold of exclusion varies from a low of roughly 1/(5S + 1), where S is the number of seats in the legislature, which obtains in list PR systems that elect nationwide, to a maximum value of 1/2, which applies to a national election under plurality bloc voting. However, calculating a threshold of exclusion is not always straightforward, since, in some countries, there are complex tiering rules that establish different degrees of proportionality at different levels of aggregation. In others there are special thresholds of required nationwide support that affect eligibility to win seats locally.
5 In some former Soviet bloc countries, a majority of eligible voters was required. Where that rule is still in place, it is a leftover of the days of one-party dominance.
6 Note that plurality also has this same threshold of exclusion.
7 Because the double ballot system for the presidency and for the National Assembly in France differ in their rules for advancing to the second round, we believe it makes more sense from a comparative politics perspective to talk about double ballot systems in France, rather than the double ballot system in France.
8 This distinction is different from another distinction common in the electoral system literature between mixed and non-mixed systems. An electoral system is mixed if it has, with disparate elements, e.g., a combination of single member districts elected by plurality and a ‘topping up’ or supplementary procedure that uses proportionality criteria to select from party lists.
for electing the Pope. It used to be the case that a unanimous vote was always required. Thus the threshold of exclusion was one. Furthermore, there was no rule which required any candidate to be declared ineligible for further consideration based on the vote share he received on a previous ballot, so it was always possible to advance to the next round of balloting. This created an example of what we have been calling a ‘constant rule’. Now, while it is still always possible to consider the entire pool of candidates at each ballot, the size of the majority needed for electing a Pope varies with k. Ultimately, if there has been a certain number of ballots that failed to elect a Pope, then a simple majority becomes sufficient for papal election. Thus, we now have a non-constant rule for papal elections.

Another way to have non-constancy is for the constituency in which the election occurs to shift. For example, in the U.S., if there is not a winner in the Electoral College then the next balloting takes place not just under different rules, but in a different electorate (the combined chambers of Congress, voting under a unit rule within each state’s combined delegation). Our impression is that the presidential selection rules in most countries that use runoffs are non-constant in one or both of these ways.

The U.S. Electoral College brings up yet another complication that is worth calling attention to: between rules where advancement rules are based directly on popular vote shares and those that have some more complex structure that takes into account vote shares in particular political subunits or geographical regions. The U.S. Electoral College in its first round is what is called in the game theory literature a weighted voting rule game (Grofman and Feld, 2005): each of the states (with two exceptions) uses a winner-take-all (plurality) rule for determining the winner of that state’s Electoral College vote, but the winner of the Electoral College is determined by weighting the outcomes in each state by that state’s share of the electoral vote (i.e., by each state’s total number of U.S. Senators and U.S. House members divided by 535). As it is well known, and the subject of much brouhaha every time a close presidential election seems to loom, the winner of a plurality of the popular vote need not coincide with the Electoral College winner. Most recently, George W. Bush in 2000 won a majority of the Electoral College but not a majority (or plurality) of the popular vote.

The U.S. Electoral College should be listed among the set of presidential election rules that have a distribution requirement, along with places where the distribution requirement is more obvious, such as Nigeria, Kenya, or Indonesia. In Indonesia, for example, the new presidential voting system (first used in 2004) is intended to “ensure that only broadly supported candidates are elected” (Reilly, 2003: p. 3). Presidential and vice-presidential candidates run as teams and a second round is avoided only if (a) the first round winner has over 50% of the total votes cast and (b) the candidate wins at least 20% of the vote in half of the provinces.

3. Similarities between runoff rules and voting methods requiring voters to rank-order candidates or parties

If we think about the similarities between most runoff rules and other voting rules, disregarding the issue of number of ballots required, the most obvious parallel is with simple plurality. As in plurality voting, in the usual runoff methods voters are asked to simply provide one or more x’s to indicate preferences; voters are not required to rank-order the candidates or parties.11 Indeed, when scholars classify electoral rules using degree of proportionality as a criterion, it is common to lump majority runoff procedures with plurality, due to their identical thresholds of exclusion. However, they have other important differences, such as strategic incentives for both voters and candidates/parties, so care must be taken not to treat plurality and runoff methods as basically the same. Moreover, runoffs may have features in common with other rules as well, such as approval voting – a voting rule involving only x’s which can be used to select either a single alternative or several (Brams and Fishburn, 1983) in which voters give votes to those candidates who, for that voter, pass some threshold of acceptability and the candidate with the most ‘approval’ votes wins.12

One dimension with which to compare runoff rules with other rules is in terms of district magnitude (the number of seats to be filled). Of particular importance is the simple dichotomous distinction of whether the rule is meant to select a single candidate or multiple winners. However, virtually all runoff methods in actual use operate in the context of single seat elections.

If we compare runoffs with single ballot election rules, there is one other critical point of comparison that is little

9 One consequence of the Electoral College is that the candidate whose vote is more ‘efficiently’ distributed is advanced, where efficiency is defined as the ability to translate votes into Electoral College seats (Grofman et al., 1997). Ceteris paribus, reliance on the Electoral College may make it harder for candidates whose support is entirely within particular parts of the country to win, because they are ‘wasting’ their popular votes by winning some states with far more votes than were needed to carry the state. However, there are other circumstances where a party with regionally concentrated strength may be advantaged, since a party may also find itself wasting its votes by having them too widely spread, thus winning few or no states. This latter possibility is enhanced if there are more than two ‘serious’ candidates in the contest.

10 In the event that both requirements are not met, the second round will be a straight two-person contest among the two top vote getters, with no distributional requirement required for victory (Reilly, 2003: p. 3). This rule was adapted from the presidential voting rule used in Nigeria, another large and ethnically plural country (Reilly, 2003: p. 3). A similar rule is in place in Kenya, with at least a 25% vote share in five or more of the eight provinces (Matthew Shugart, personal communication, January, 2008). The two most important voting systems where voters are required to provide rankings of alternatives are the single transferable vote (and its single seat special case, the alternative vote) and the Bordas count (See Bowler and Grofman, 2000 for a discussion of the former; and Black, 1958, and Saari, 1994, 1995 for a discussion of the latter). The Coombs rule, while little known, is a ranked procedure of considerable theoretical interest.

12 On the other hand, the single-nontransferable vote (SNTV), which is the single vote version of limited voting, while only requiring a single x vote, involves multi-seat elections, and it can be viewed as a form of proportional representation. Under SNTV the m candidates with the most votes win, but each voter has only a single vote to cast. This means that the threshold of exclusion, the largest vote share with which a party can still be denied a seat, is the same for SNTV as it is for the d’Hondt version of list PR. Pure list PR systems also involve multi-seat elections, but again (except in open list systems) voters have only a single x to cast – here normally for a party, rather than for a candidate.
commented on in the electoral systems literature: for any runoff procedure using only X’s, we can always create a single ballot procedure using rank-ordered ballots which can be thought of as an analogue to that runoff rule. For example, the ‘alternative vote’ can be regarded as a direct analogue of the MRSE (majority rule sequential elimination procedure) where we drop the candidate with lowest vote share. Indeed, the alternative vote has been pushed by a leading electoral reform group in the U.S. (Center for Voting and Democracy), under the name instant runoff, as a replacement for simple plurality elections. However, despite the similarities between MRSE and the alternative vote, they are not completely logically equivalent rules for the simple reason that the information about (inferred) voter preferences available to the voters on the kth ballot in MRSE (k > 1) is more complete than the information available to the voter who must decide how to list preferences on his/her alternative vote ballot. Only if we disregard the potential for strategic voting, and posit that each voter simply lists his most preferred candidates in rank-order on the alternative vote ballot, do MRSE and the alternative vote become equivalent.

In like manner, as noted by van der Kolk (2006), there is a direct analogy between the two ballot majority runoff procedure and the supplementary vote procedure (SV) that was recently proposed by the Plant committee in Britain and advocated by some prominent British political scientists. However, as van der Kolk also notes, the peculiar variant of the SV procedure advocated by the Plant committee limited voters in how many choices they could rank (to two choices), which makes the analogy an imperfect one. Instead, as van der Kolk points out, the closest analogue between the two ballot

majority runoff method and a single ballot ranked scheme is with the Contingent Vote (CV) system used in Queensland, Australia between 1892 and 1942 (Reilly, 1997). However, while all runoff procedures can be given ranked method ‘near twins’, not all ranked voting procedures have runoff variants. Consider the Coombs rule, an interesting variant on the alternative vote. The Coombs rule (Straffin, 1980; Grofman and Feld, 2004) requires voters to submit a rank-ordering of alternatives, but now we drop the candidate with most last place preferences rather than the candidate with fewest first place votes, as we would do in AV. Because we cannot determine last place preferences from first place X’s on ballots, there is no runoff method involving only first place rankings which is directly analogous to the Coombs rule.

4. Emphasizing differences among the variety of voting rules selecting a single alternative from a larger set

Some recent literature has argued that paying primary attention to the distinction between majority/plurality systems on the one hand, and PR systems, on the other, misses much of importance about the incentives structuring political competition. Other distinctions, such as between candidate centered and party centered ballots (see, e.g., Carey and Shugart, 1995; Grofman and Bowlar, 1997; Grofman, 2005), or between rank-ordered and x ballots, may lead us to view the similarities/differences among electoral rules in a quite different light. For example, Kurrild-Klitgard (2001) points out that both the usual PR methods and single seat plurality or runoff methods are alike in that they each commonly only elicit information about first preferences.

Relatedly, we have recently argued that the distinctions among the many different types of rules that select a single candidate from a larger set may loom almost as large as the standard majoritarian/pluralitarian versus proportional/minoritarian distinction (Grofman, in press). In particular, (1) voting rules can be characterized in part by which voting rule they reduce to when applied within a single seat constituency; and (2) voting methods can be divided into three groupings based on how, for single seat constituencies, they treat voter preferences. Some rules can best be regarded as majoritarian in thrust; i.e., they are perhaps best judged by how likely they are to satisfy the Condorcet criterion (Black, 1958) of always picking an alternative that defeats each and every other alternative in paired contest if one exists. Both plurality and runoff methods, especially the French double ballot system and variants thereof, are often regarded as falling into this category. For example, Colomer (2007) reports

13 We have defined runoff methods as those that potentially require more than one round of balloting, and we have confined ourselves to runoff methods which ask voters to indicate preferences only with X’s. It is, however, perhaps a matter of definitional choice as to whether methods that require the voter to provide a rank-ordering of candidates, with lower level preferences potentially coming into play as candidates are eliminated, should count as ‘runoff methods’. However, the converse is not true, since rank-ordered ballots provide more information that can be used to specify voting rules than ballots which are X’s only.

14 These reforms have been successful in getting the alternative vote adopted to replace plurality elections in several jurisdictions, perhaps most notably for city council elections in San Francisco. We would note that using the term ‘instant runoff’ to refer only to AV is highly misleading since there are other ranked procedures (e.g., SV and Coombs) that can also serve as ‘instant runoffs’.

15 There are many different types of strategic voting in addition to not voting for the candidate who is most preferred (e.g., strategic rankings of candidates in an STV election, or where the electoral rules permit such crossovers, voting in a non-preferred party’s primary).

16 Also, of course, new information may be received by a voter between the first and second round of an election.

17 The Plant committee recommendations did not influence electoral rule changes in the UK at the level at which they were directed, national parliamentary elections; nor, despite innovations in electoral system choice that occurred as part of the process of devolution (see discussion in Dunleavy and Margetts, 2001, 2005), did their recommendations influence electoral system choices at the regional parliamentary level. However, van der Kolk points out that some dozen municipalities in England have adopted SV (with a ‘rank only two candidates’ restriction on voters) as part of a shift to directly elected mayors (for further discussion see van der Kolk et al., 2004, 2006).

18 Van der Kolk (2006) reminds us that a version of SV with a constraint allowing voters to rank only two alternatives was used in Alabama in Democratic Party primaries between 1915 and 1931; and one SV version with a constraint allowing voters to only rank three alternatives has been used since 1982 in Sri Lankan presidential elections (Reilly, 1997).

19 This is a many to one mapping; i.e., in general, there are many rules for selecting m candidates that will reduce to the same voting rule when m = 1.
the extent to which presidential (and parliamentary) elections in Latin America choose Condorcet winners (or Condorcet losers; i.e., candidates who might be expected to lose in pairwise choice against each and every other alternative). He finds that runoff methods are generally superior to simple plurality (cf. Grofman and Feld, 2004).

Other rules can be thought of as consensual in character, such as the Balinski and Laraki (2006) méthode majoritaire, which is a ranked method that looks to median preferences; or approval voting, which asks voters to indicate with x’s what alternatives they would be ‘satisfied’ with (Brams and Fishburn, 1983). The Borda rule, a ranked procedure, also can be thought of as a consensual procedure. While we certainly can evaluate these methods in terms of whether or not they are likely to satisfy the Condorcet criterion, the inventors of the méthode majoritaire explicitly assert that it offers an alternative to both the classic Condorcetian view and to the approach of Jean Charles Borda who has recently been espoused by Donald Saari (1994, 1995). Similarly, Saari has defended the Borda rule against both advocates of approval voting and those who support the Condorcet criterion approach.

Still other procedures can be thought of as going beyond consensual, by being structured to avoid outcomes that a minority is deeply unhappy about; e.g., the anti-plurality rule where voters cast their x’s to indicate the candidate they least wish to see elected, or the Coombs rule, a ranked method mentioned previously that sequentially eliminates the candidates with the last place votes (Grofman and Feld, 2004).

By identifying variants of runoff methods in this article, we have seen just how complex rules can be even when we are only selecting a single candidate from among several. Similarly, by noting some of the mathematical links between runoff methods and rank-ordered methods, we hope to have contributed to bridging the gap between work on electoral rules within political science, which primarily focuses on rules actually in use for parliamentary and presidential elections, and the more abstract measuring characteristic of the social choice literature in mathematical economics.

References


Grofman, B., Feld, Scott L., 2004. If you like the alternative vote (a.k.a. the instant runoff) then you ought to know about the Coombs rule. Electoral Studies 23, 641–659.


The anti-plurality rule, like the plurality rule is an example of what is known as a scoring rule. There are an infinite number of such rules, including the Borda rule, which we may regard as exactly intermediate between plurality and anti-plurality methods.

22 Inventoring runoff election rules in use around the world in terms of the five-fold classification proposed in this article remains to be done (see, however, Birch, 2003).

23 Also see Regenwetter et al. (2006).