

ELECTORAL SYSTEMS AND THEIR GOALS:
A PRIMER FOR DECISION-MAKERS

By

Donald L. Horowitz
James B. Duke Professor of Law and Political Science
Duke University
Durham, North Carolina 27708-0360
U.S.A.

January 2003

© Copyright 2003 by Donald L. Horowitz

ELECTORAL SYSTEMS AND THEIR GOALS

By Donald L. Horowitz^{*}

To evaluate an electoral system or to choose a new one, it is necessary to ask first what does one want the electoral system to do. No electoral system simply reflects voter preferences or the existing pattern of cleavages in a society or the prevailing political party configuration. Every electoral system shapes and reshapes these features of the environment, and each does so in different ways. In this short paper, I want to set out several possible purposes of electoral systems that can be found in the literature in the subject and then make some observations about those purposes and the electoral systems that further them, before turning to a few comments on electoral systems in relation to Malaysian politics.

First, however, I need to underscore a point just made about a common assumption—that the best electoral system is the one that straightforwardly and most accurately reflects the preferences of voters. The nature of an electoral system is to aggregate preferences and to convert them into electoral results, and no system can do this as a passive translation of individual wishes into a collective choice. Moreover, every electoral system has biases built into its mechanisms of decision, and these then feed back into the structure of choices confronting voters, constraining and changing choices that they might have made under other systems. Consequently, not only is there imperfect reflection of voter preferences in the first instance, but voter preferences themselves are shaped by the electoral system. Preferences do not and cannot exist independently of it.

The fact that each electoral system contains a different array of biases from every other electoral system means that those who decide among such systems can choose, in effect, to prefer one set of biases over another. And to prefer one over another is to make a policy choice. Hence one can speak of the goals of the system, even though the choice of bias is not always consciously made. It follows from this that there are as many potential goals of electoral-system choice as there are combinations of biases and systems. Any enumeration of goals is, therefore, a function of objectives people living in a given society might wish to achieve and those they might wish to avoid, when

¹© Copyright 2003 by Donald L. Horowitz.

matched against the propensity of particular electoral systems to produce results in one direction or another. Many such lists of goals, in other words, could be drawn up, and all such lists have an element of arbitrariness to them.

One last preliminary caveat. When we speak of goals to be achieved, there should be no illusion that the electoral system can, by itself, achieve them. Electoral systems shape and constrain the way in which politicians and constituents behave, but they are only one small part of the forces affecting the total constellation of behaviour, even of political behaviour. Miracles do not follow from changes of electoral systems. No one should expect more than incremental changes in behavioural patterns once the configuration of electoral incentives is altered. But sometimes increments of change can be surpassingly important.

I. Six Goals

Six aims of an electoral system come readily to mind. Some of these are mutually compatible, but some others are mutually incompatible, which is why it is so important to be clear about what one is choosing. (The choice, of course, must also be geared to the pre-existing features of the political environment, since the functioning of electoral systems varies with the context.) Here are the six possible goals:

1. Proportionality of seats to votes
2. Accountability to constituents
3. Durable governments
4. Victories of the Condorcet winners
5. Interethnic and interreligious conciliation
6. Minority officeholding

I shall discuss each of these in turn.

1. *Proportionality of Seats to Votes*

Increasingly, scholars and decision makers are inclined to judge electoral systems by their ability or inability to produce proportional results. A political party that gains 20 percent of the total vote, it is argued, should win 20 percent of the total seats, rather than a few or no seats, which it may receive if elections are held on a constituency basis and its support is thinly spread rather than regionally concentrated. A party with 50 percent of the vote should, on this view, win only 50 percent of the seats rather than the 60 or 65 percent it may receive under electoral systems that often provide an inadvertent seat bonus to the largest party.

There are several ways to produce more or less proportional results, including list-system proportional representation (with its two main variants, national lists and

constituency lists) and the single transferable vote. Some such systems can be married to other purposes beside proportionality, particularly if combined with other systems in a hybrid electoral arrangement, but PR is often inimical to some of the other goals of electoral system choice, as I shall suggest later. Fairness of outcome, in the sense of proportionality of seats to votes, is, I want to stress, only one among several goals, and it may not be the most important goal. Because it is easily measurable, however, proportionality tends to preempt other goals. Discussions of proportionality should never be held in a vacuum.

It is also true that deviations from proportionality can be limited in non-PR systems by attending to some of the sources of non-proportional outcomes that do not derive from the electoral system as such. Among those sources, malapportionment of constituencies, so that it takes two or three times as many voters to elect one representative as it does to elect another, is a very serious cause of disproportional results.

2. Accountability to Constituents

Elections to representative bodies assume some degree of representativeness – that is, accountability of legislators to those who elect them. It is generally thought that electoral systems that limit the power of central party leaders to choose candidates produce more responsive representatives. On these grounds, national list-system PR usually reposes great power in party leaders to decide which candidates will have favourable positions on the parties' lists and thus have better chances of being elected. When central party leaders have such power, the sovereignty of the voter to choose the candidates, rather than just to choose among candidates, is thought to be impaired.

On this score, constituency-based systems, such as first-past-the-post or even constituency-list PR (with small constituencies of perhaps three or four seats), are said to be preferable. But there are other ways to mitigate the domination of the process by central party leaders under list PR. One way is to allow voters to alter the order of candidates on the list, by voting for candidate 6 over candidate 3 on the list, for example. This is called open-list PR, but it can have some perverse consequences, especially in multiethnic societies.

3. Durable Governments

Obviously, an electoral system cannot represent the idiosyncratic opinions of every voter. Nevertheless, some systems make it possible for many shades of opinion to be represented, sometimes so many that the legislature ends up being fragmented, with no party having anywhere near 50 percent of the seats. In such cases, coalitions are, of course, necessary. Where the legislature is deeply fragmented, it may be difficult to put together durable coalitions. Other electoral systems may force parties to aggregate

the diverse opinions in a society for the sake of electoral success. Where this happens and diverse opinions are represented within parties rather than across parties, the reduction in the number of parties makes it more likely that durable governments can be formed. And durable governments are thought to be desirable because they promote policy consistency and responsibility and, even more importantly, may avoid the instability that can result during interregna or from the creation of fragile, unpredictable coalitions.

4. *Victory of the Condorcet Winner*

The Condorcet winner is the candidate who would receive a majority of the vote in a paired or head-to-head contest with each and every other candidate. The Condorcet winner is obviously the more popular candidate, whose victory, it is thought, ought to be preferred. But there are obstacles to this outcome. Since often there are more than two candidates, it is possible for some systems to produce results that disfavour the Condorcet winner. Sometimes first-past-the-post does this. Take a three-way contest in which candidates receive the following votes: X, 45 percent; Y, 40 percent; Z, 15 percent. Under first-past-the-post, candidate X wins. But if candidate Y faced only candidate X head-to-head, Y might be the candidate preferred by a majority of voters; and Y might also defeat Z in a paired contest. Electoral systems that can disfavour the Condorcet winner are sometimes thought to be wanting. But, of course, they may have other virtues.

There are systems that do a good job at picking the Condorcet winner. Both the alternative vote and Coombs rule (discussed later) are good at eliciting second preferences that are suppressed by first-past-the-post. But both may have other disadvantages. Again, with electoral systems, it is always a question of knowing what one wants and choosing among alternatives, all of which will have some undesirable features.

5. *Interethnic and Interreligious Conciliation*

Electoral systems that produce proportional results or accountability to constituents or durable governments may or may not foster interethnic conciliation. One way to think about electoral systems and interethnic conciliation is to ask whether a given system provides politicians with electoral inducements for moderate behaviour, for compromises with members of other ethnic groups for the sake of electoral success. Some systems can do this. The Lebanese electoral system, with its ethnically reserved seats, multi-seat constituencies, and common roll elections, gave politicians very good reasons to cooperate across group lines, for they could not be elected on the votes of their own group alone. They had to pool votes (that is, exchange support) with candidates of other groups standing in different reserved seats in the same constituency.

Similarly, systems that require candidates to achieve regional distribution of votes, in addition to a plurality, may foster conciliatory behaviour if territory is a proxy for ethnicity because groups are regionally concentrated. Nigeria pioneered this approach in its presidential elections, and now Indonesia has gone in the same direction for its presidential elections.

On the other hand, electoral systems that allow politicians to be elected without behaving moderately may make post-electoral conciliation more difficult. Coalitions that are created after elections merely to form a government of 50 percent plus one of the seats in parliament may prove to be fragile when divisive ethnic issues arise. So, for interethnic conciliation, the question is how the electoral system affects the preelectoral calculations of parties and politicians.

6. *Minority Officeholding*

Some writers, policy-makers, and ethnic-group activists think that group proportionality ought to be a goal of electoral systems. The (debatable) assumption is that, if group A comprises 10 percent of the population, it ought to comprise 10 percent of members of the legislature. Many electoral systems produce results that underrepresent members of minority groups in legislatures, if by *represent* we mean produce a share of electoral victors that is proportionate to the minority share of the population. Under first-past-the-post elections, for example, if minorities are geographically well distributed, winning candidates may be drawn largely from the majority population. The goal of minority officeholding can thus be seen as an instance of the same phenomenon that gives rise to attempts to achieve proportionality between votes and seats, except that proportionality in that respect is *party* proportionality, rather than *group* proportionality.

In some countries, notably the United States, serious efforts have been made to increase the minority share of legislators. In the context of plurality elections, the Voting Rights Act has been interpreted to require the redrawing of constituency boundaries in the direction of greater homogeneity, so as to facilitate the election of minority representatives where minorities constitute more than half of an electorate. There have been few suggestions that the electoral system itself be changed for this purpose. Some minority advocates, however, have urged consideration of the cumulative vote, an electoral system that allows voters in a multimember constituency to cast some or all of their votes for a single candidate, thereby maximizing the chance that a minority candidate could, on the basis of such cumulation, achieve victory over other candidates whose support was more widely but less intensely distributed. By providing incentives for candidates to take extreme positions so as to heighten their ability to attract all of the votes of a subset of voters, however, the cumulative vote could open the way to polarized politics.

As this possibility shows, there may be a tradeoff among the goals of group officeholding and interethnic conciliation. Similarly, proportionate minority officeholding does not guarantee that minority interests will receive attention in the legislative process. Indeed, minority officeholding may come at the expense of minority representation in this larger sense, for the creation of more homogeneous constituencies means not only more minority-dominated constituencies but also, correlatively, more constituencies in which majority-group voters dominate and in which majority-group candidates do not need to worry about minority support or minority interests.²

²For one among many assessments, see Charles Cameron *et al.*, “Do Majority-Minority Districts Maximize Substantive Black Representation in Congress?” *American Political Science Review*, Vol. 90, no. 4 (December 1996), pp. 794-812.

II. Choosing Among Goals

Many decision makers try to design the electoral system to maximize more than one goal. The German system, for example, is a constituency-based system, but with a proportional overlay, so that legislators have reasons to respond to their constituents but parties receive an overall number of seats that is more or less proportional to the votes they have won nationally. There is an increasing trend toward hybrid systems to achieve multiple goals, as New Zealand, Italy, and Japan have modified their former systems.

There are also, however, strong cultural continuities in electoral arrangements. The U.K., the U.S., Canada, India, many Anglophone African countries, and Malaysia all use the first-past-the-post system, which is regarded as a system common in the English-speaking world. By contrast, continental Europe tends to use list PR, and so do Francophone countries in Africa. Some very poor decisions can be made on the basis of cultural affinity. Benin, a former French colony, opted, like France, for a presidential system, with a run-off election if no candidate receives a majority on the first ballot. By choosing the run-off system, Benin turned a relatively benign tripolar ethnic conflict into a much more serious bipolar conflict. Deliberate choice, not cultural affinity, ought to be the basis of decisions about electoral systems.

As post-colonial countries rethink their electoral arrangements, they often depart from their inherited electoral systems. There is some evidence of adoption of systems from outside zones of cultural or colonial affinity, but generally their choices stay close to those of the former metropole. When English-speaking countries, such as Ireland or Australia, choose proportional representation, they tend not to choose list-system PR. Instead, they tend to opt for the single transferable vote, sometimes called the Anglo-Saxon version of PR. And in other cases when they depart from first-past-the-post, they stay with single-member constituency systems, such as the alternative vote, which has been used in Australia, Papua-New Guinea, and Fiji.

New electoral systems have effects on party formation, party behaviour, and party systems. It is a serious mistake to take the preexisting party configuration and project it unchanged into the future if a new electoral system is adopted. When mixed-member proportional representation was adopted in New Zealand, it became much harder to form a government after the first election, and a single right-wing party gained power in the governing coalition that it could never have had under the former first-past-the-post system. When the alternative vote was adopted in Fiji, two multiethnic coalitions were formed, embracing almost all parties, and one was able to form the government. In both cases, the new system altered the strength of preexisting parties and changed their alignments. Any new system could be expected to have comparably strong effects elsewhere.

Conventionally, it is thought that first-past-the-post elections promote a party system with relatively few parties (sometimes only two). Under first-past-the-post, a party with as little as 48 percent of the vote is virtually guaranteed to receive more than 50 percent of the seats. (Even 40 percent of the vote gives a party a fair chance at a majority of seats.) Because of this seat bonus, first-past-the-post typically makes the formation of governments easier than it might be under some other systems and also makes governments more durable. Those who value stability and value consistency in policy-making often prefer first-past-the-post. But, of course, first-past-the-post does not have these effects everywhere. In India, Malaysia, and Canada, first-past-the-post has been compatible with multiparty systems, because the structure of social cleavages makes it impossible to compress all the main tendencies into two or three parties. Still, plurality systems do provide inducements for the aggregation or amalgamation of divergent interests into a few parties.

List-system proportional representation, on the other hand, is said to facilitate the representation of social cleavages. Minority opinions that are unrepresented in plurality-winner systems may find expression where parties with five or ten percent of the vote nationwide are accorded five or ten percent of the seats. Where, however, there are multiple social cleavages, adoption of a highly proportional list system creates incentives for fragmentation rather than amalgamation of political tendencies. If many social groups are organized into separate parties, each of which can gain a small fraction of the total seats, the likelihood is that political differences will be magnified rather than compressed. Governments may be formed only with difficulty, their composition may be unpredictable, and their durability may be doubtful. This situation has been called by Giovanni Sartori “polarized pluralism,” a situation fostered by PR and conducive to immobilism. Parties cultivate only their own supporters, and compromise is hard to come by.³

Not everyone agrees with such a diagnosis. Theorists of consociational democracy argue that ethnic differences can be composed by coalitions once all major groups are represented in parliament through proportional representation (and through adoption of a variety of norms of conflict management, in addition to PR). But, as I indicated, the electoral system does not merely represent preexisting tendencies; it also shapes them. Factions that would, under other systems, remain within one political party, for fear of being unable to win seats on their own, may, under PR, be tempted to go it alone. In a severely divided society, with many group and subgroup affiliations, and many shades of opinion, list PR is likely to produce a great deal of party fragmentation.

³Sartori first made these points in an essay in Joseph LaPalombara and Myron Weiner, eds., *Political Parties and Political Development* (Princeton: Princeton University Press, 1966). The theme was pursued in Sartori’s volume on political parties and more recently in his book on constitutional engineering.

By representing many shades of opinion—and by proliferating those shades of opinion—list PR is indifferent to the goal of choosing the Condorcet winner. And vice versa: systems that are good at choosing the Condorcet winner leave the interests associated with many losing candidates unrepresented, as PR does not. Note, however, that the Condorcet winner is likely to be a generally moderate candidate. PR winners, on the other hand, may be moderate or not. The whole thrust of PR is to represent all opinions, regardless of their position on the political spectrum. Sartori's descriptive phrase, *polarized pluralism*, is apt where the spectrum is broad or there is more than one spectrum, as there is when class cleavages coexist with but do not overlap with ethnic, religious, and regional cleavages.

A system that chooses the Condorcet winner may thus also foster interethnic conciliation, simply by favouring moderates or compromise candidates over extremists. The Lebanese system appears to do this by favouring candidates who can gain some support from voters outside their own group. The same goes for systems that require regionally well distributed support, in addition to a plurality, for victory : they make it probable that a candidate who manages to achieve both requirements has broad appeal and so is unlikely to be an extremist, popular only with his or her own segment of the population.

The alternative vote also favours moderates. AV is an electoral system that, unlike first-past-the-post, requires 50 percent of the vote plus one for victory. Where no candidate receives 50 percent of voters' first preferences, AV requires that the candidate with the fewest first-preference votes be eliminated and that that candidate's second-preference votes be redistributed as if they were first preferences. The process is repeated until a candidate receives 50 percent. In this way, victory goes to candidates who have some support outside the core of supporters who accord them first preferences. A variation on this is the Coombs rule, under which, if no candidate has 50 percent of first preferences, the candidate with the largest number of last preferences, rather than the smallest number of first preferences, is eliminated first, and the process is carried out until a 50-percent-plus-one winner is found. Coombs is asserted to be better than the alternative vote at choosing the Condorcet winner.⁴

Note, however, that neither the alternative vote nor Coombs will produce proportional results, if by that phrase we mean proportionality of first preference votes to seats. But, of course, under such preferential systems, it makes no sense to judge proportionality by first-preference votes alone. The whole point of the system is to count second and subsequent preferences of voters, rather than to discard them, as they

⁴See Bernard Grofman and Scott Feld, "If You Like the Alternative Vote (a.k.a. the Instant Runoff), Then You'll Really Like the Coombs Rule," unpublished paper, March 1, 2002.

are discarded by first-past-the-post and list PR alike. Proportionality is generally indifferent to moderation, and moderation is indifferent to proportionality. These are two quite different goals.

If preferential systems such as the alternative vote or Coombs are intended to reflect the full array of voter preferences, they may also shape those preferences, and they may shape the behaviour of the parties competing under them. The key to this is the 50-percent threshold for victory in each constituency. As parties recognize that they may not be able by themselves to secure 50 percent of the vote in a given constituency or across a run of constituencies, they are likely to form coalitions before the election in order to exchange second and subsequent preferences. If they do not do this, their opponents will. Something of an analogous sort (but usually with less intensity) occurs under first-past-the-post, for a party wishing to be the plurality winner may try to become a broadly aggregative organization that encompasses a range of views, tendencies, and social groups.

Nevertheless, the plurality and majority thresholds may create altogether different incentives for parties in a fragmented party system. In a four-way contest under the plurality rule, a party can win with as little as 26 or 30 or 35 percent of the vote; and, in recurrent three-way and four-way races, it may secure 40 or even 50 percent of the seats on a much smaller percentage of the total vote. Such a party need not broaden out its support unduly, and it need not compromise with other social groups, in order to win the election. But where the threshold for victory in each constituency is 50 percent and the party system is fragmented, there are powerful incentives for parties to make agreements with each other for the exchange of second and subsequent preferences. The success of such exchanges is likely to determine the winner.

Such preferential systems thus encourage the formation of preelectoral coalitions, and those coalitions in turn depend upon the ability of parties to compromise their differences. Hence the conciliatory thrust of systems of this sort under the conditions specified.

Another preferential system, the single transferable vote, provides fewer such incentives. That is because the threshold for success is governed by a different formula. STV operates in multimember constituencies. To be elected, a candidate must achieve a quota, as follows:

$$Quota = \frac{1}{\text{Number of Seats} + 1}$$
 In other words, a candidate is deemed elected if, in a four-member constituency, he or she receives one vote more than one-fifth of the votes. Such a low threshold provides few incentives to interparty agreements to transfer votes, even though perhaps the last seat in a constituency may be decided on transferred votes, as votes above the quota for victory

are transferred in accordance with the second and subsequent preferences of voters providing those surplus votes. Overall, then, STV provides weaker incentives to compromise than preferential systems with majority thresholds do.

If parties do not seek votes (in this case in the form of second and subsequent preferences) from voters whose primary allegiance is to another party, an important consequence follows. In socially polarized settings, parties may not need to moderate their positions, and their supporters need not be cognizant of the claims of other social groups. After the election, of course, interparty arrangements may have to be made in order to govern. Under list PR or STV or FPP, in conditions of party fragmentation, postelectoral coalitions will probably be necessary, but they will not be based on the sort of understanding that is required of parties that make preelectoral coalitions in order to maximize their share of the vote, rather than merely to put together enough seats to form a government.

That is not to say that STV does not serve any useful functions. In general, it produces relatively proportional results, if proportionality is measured by first preferences, although that proportionality depends heavily on how many seats each constituency has (the more seats, the more proportional the results). And since STV is constituency based, it provides some of the same constituency responsiveness that first-part-the-post or AV does, although, of course, multimember seats and low thresholds may mitigate the need of candidates to cater to the interests of those who belong to groups with which they themselves are not identified. Constituency-list PR is also a system that provides for accountability to constituents and a kind of local focus that national-list PR does not provide.

So far I have discussed electoral systems in terms of their goals, as if the major differences in outcome derived entirely from differences among features inherent in those systems. But a major caveat must be introduced here. Electoral outcomes are not just produced by systems, but by the preexisting pattern of social cleavages, single or multiple, bipolar or multipolar. Is there only one axis of difference in the society that will be reflected in the party system, or is there more than one? To what extent do the various axes overlap? Are there two or three or many positions at which voters locate themselves along each axis? And then there are the specific features of electoral arrangements in a polity. These can truly skew the results in one direction or another. For constituency-based systems, constituency delimitation and apportionment are inordinately important. Are constituency boundaries drawn so as to maximize homogeneity or heterogeneity along the relevant axis? The exigencies of securing election will vary accordingly. Are boundaries drawn so that some groups waste votes to elect favoured candidates by large majorities, while other groups are able to use their votes more efficiently? Are constituencies more or less equal or unequal in population? Malapportionment—that is, unequal population of constituencies—is a major cause of disproportionality between votes and seats. And, for list PR, an analogous issue is

whether there is a low (one or two percent) or high (five percent or more) threshold for a party to secure representation. A system in which any party with one percent of the vote can win a seat or seats in parliament provides great inducements for parties to split into their component factions, and it may enable very small parties that can make or break governments to have disproportionate impact in determining policy and receiving patronage. Ironically, the more perfect the proportionality built into such a system, the more disproportionate the ultimate policy results may be.

This should suffice to show that the details matter. Electoral systems need to be tailored closely to what those who design them want them to do. Of course, most of the time, those who design such systems want them to produce results that favour the interests they prefer. My point is that there are other ways to evaluate electoral systems and the specific arrangements made pursuant to them. And very likely, no matter what the intentions of the designers, there will be some unanticipated consequences.

In every analysis, the way to proceed is to ask what goals should be fostered, which goals preferred over others, and what the likely effects of various alternatives really are. No one can answer these questions without an extended analysis that links electoral reasoning to the party system and the pattern of social cleavages.

III. Attending to Details

There are ways to subvert even the most carefully chosen electoral system. One very common way concerns the boundaries and size of constituencies. Constituencies with vastly different numbers of voters are a prime source of disproportional outcomes. Constituencies whose boundaries have been delimited to advantage one ethnic group over another can undo the effects of electoral systems that have ethnically conciliatory features. Boundaries drawn to favour incumbents can make necessary political change impossible to achieve. It is, therefore, crucially important, not merely to choose an apt electoral system, but to make certain that boundaries are delimited by governmental bodies that are as neutral, professional, and independent as possible.

Constituency delimitation has some special requirements where systems are adopted with interethnic or interreligious conciliation as the primary goal. Whether the alternative vote or the Coombs rule or the Lebanese system is chosen, constituencies must be demarcated to assure that they are heterogeneous in composition.

The same attention to detail is required to make the Nigerian-Indonesian presidential electoral system work. The 1978 Nigerian system required the winning candidate to gain a plurality plus at least 25 percent of the vote in no fewer than two thirds of the states. While this system creates incentives for broadly distributed support, it also opens the possibility that no candidate will secure the requisite regional distribution to

be elected. It is, therefore, necessary to provide a clear and decisive fallback provision to choose a president if this contingency occurs. The Indonesian constitution makers recently opted for a 50-percent-plus-one threshold plus a provincial distribution requirement, thereby making it especially difficult for any candidate to win on the first round if there are several candidates, which Indonesia's fragmented party system makes likely. It is important not to set thresholds for victory so high as to undo the benefits of the system being adopted. Electoral choice, in short, cannot stop at the level of the electoral system alone but must extend to all the details of implementing them so as to gain the benefits of the system that is preferred.