DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

EQUILIBRIA WITH UNRESTRICTED ENTRY IN MULTI MEMBER DISTRICT
PLURALITY (SNTV) ELECTIONS
Part II: Evidence from Taiwan and Japan

Olga Shvetsova

SOCIAL SCIENCE WORKING PAPER 881

March 1994

EQUILIBRIA WITH UNRESTRICTED ENTRY IN MULTI MEMBER DISTRICT PLURALITY (SNTV) ELECTIONS
Part II: Evidence from Taiwan and Japan

Olga Shvetsova

Abstract

The hypothesis that the Duverger's Law can be extended to plurality multi member district elections received some empirical support in Steven Reed's (1990) study of the Japanese elections. Here we return to electoral data of Japan and Taiwan in order to find evidence consistent with the theoretical result offered in Part I of this essay, namely, that cohesive electorates should witness either the competition among as many "serious" candidates as there are seats, or, at most, one "extra" candidate per race. We also compare the consistency of the data with two alternative sets of predictions, one – derived from our candidate-based model (strategic candidates, sincere voters), another – proposed by Gary Cox (1993) where strategic voter behavior is analyzed. The indications are that the strategic role of the candidates should be viewed as leading.

EQUILIBRIA WITH UNRESTRICTED ENTRY
IN MULTIMEMBER DISTRICT PLURALITY (SNTV) ELECTIONS
Part II: Evidence from Taiwan and Japan

Olga V. Shvetsova

California Institute of Technology

Building on the theoretical results presented in Part I of this paper, here we examine Japanese and Taiwanese election data to assess the existence of stable patterns in the number and relative electoral strength of candidates at the district level. At the same time we contrast the empirical support that our predictions receive against that received by the predictions of a "rival" conceptualization - the one offered by Cox (1993), which focuses on voters as strategic actors rather than candidates. Generally, the empirical evidence indicates the stronger impact of the strategic behavior of candidates on patterns of competition. On the other hand, that support is not so strong as to lead us to restrict ourselves to models based exclusively on sincere voting. It is clear that strategic voting must be an important consideration as well.

In the next section we discuss why studying elections in multimember districts is different from studying those in single-member districts, and how such studies advance our knowledge of electoral behavior in general. In Section 2 we state the competing hypotheses, which differ both in terms of predicted number of candidates within a district and the distribution votes among them. Sections 3 and 4 are devoted to the analysis of Japanese electoral data. Section 3 looks at the number of candidates, while Section 4 deals with the distribution of votes. Section 5 offers a parallel analysis of Taiwanese elections. Finally, in Section 6 we attempt to project theoretical findings on the role and operation of political parties.

1. STUDYING MULTIMEMBER DISTRICT ELECTIONS

Few theoretical results have been accumulated about how electoral systems work when there is more than one legislative seat to be filled per district. This lack of theory, in contrast to the well developed conceptualizations of single-member district elections, is explainable not only by the prevalence of the later case in American politics, but also by the exploding strategic complexity of multi-seat elections. However, it is the multimember district case that may play a decisive role in improving our general understanding of the true nature of the electoral game. Specifically, in deciding how to approach theorizing about elections, we must often ask one primary question: how much attention should we pay to strategic choice among voters versus among candidates.

1

Ideally, of course, we would want to consider both forms of strategic behavior. Unfortunately, in most multi-candidate election systems, doing so results in unmanageable analytic complexity. Thus, in examining the implications of a particular election law, it is useful to learn whether that law impacts electoral outcomes primarily through the strategic calculations of voters or that of candidates. If, as in the case here, we are interested in predicting the number of candidates or parties likely to compete as a function of district magnitude or allocation rule, should we focus on the possibility that voters will act strategically by voting for candidates other than those who rank highest on their preference order? Or, are we better served by supposing that voters act sincerely and by focusing instead on candidates who choose policy platforms in anticipation of the fact that competitors enter the contest whenever doing so is profitable?

At least for SNTV multi-member district elections, Reed (1990) and Cox (1993) suggest that the implications of voter and candidate rationality can be separated in terms of the number of candidates competing and the vote shares they obtain. What we require, though, are partial equilibrium models that provide the requisite empirically testable hypotheses that rigorously differentiate between the two assumptions of voter and candidate strategic action. Cox (1993) provides an appropriate analysis when voters are strategic and candidates are mere mannequins. In contrast, the model offered in Part I of this essay lets candidates be strategic (by granting them spatial mobility and by allowing them to choose whether or not to compete), and assumes that they compete for the votes of sincere Downsian voters.[1]

2. DISCRIMINATING HYPOTHESES

As Part I of this essay shows, the difference between the predictions of these two models concerns both the number of candidates that compete in equilibrium and the way in which the vote is divided among them. In Cox's model, the difference in the vote for the k winning candidates (where k is district size) should be minimal – all such candidates should receive equal electoral support. Furthermore, the electoral support of all vote-getting losers must be the same, lower than that of the last winner and should eventually drop to zero as one moves down the list. Insofar as the number of candidates is concerned, this number should be at least $k+1$. In contrast, our model predicts the equilibrium number of candidates being either k or $k+1$ in a non-polarized constituency (i.e., if the distribution of preferences is unimodal). No prediction is made about the relationship among the

---

[1] Employing a simple Nash argument in a static context allows us to draw a line between things that can and cannot be if the system is in equilibrium.

2

vote shares of the winners, but in the $k+1$ case the last winner and the first looser should enjoy approximately the same share of the vote.[2]

Thus, if we indicate the number of votes received by candidate i, $i=1,...k,...n$ as v^i (where k is the number of seats to be filled in a district, and where i is the rank of a candidate within that district), then the two hypotheses we want to test are as follows:

H1 (Cox's): $A1$: $v^i = v^{k+1}$, or $v^i = 0$ for all $i > k$, and

$B1$: $v^k/v^1 \approx 1$.

H2 (ours): $A2$: k or $k+1$ candidates compete with $v^1 \geq ... \geq v^k$, and

$B2$: In the case of $k+1$ candidate competition, $v^{k+1}/v^k \approx 1$.

Represented graphically, the two hypotheses are compared in Figure 1. Cox's hypothesis implies a uniformly distributed vote up to the last winning candidate. After k, the vote declines to some v^{k+1}, where it stays until it drops all the way to zero at some $j > k+1$ (the exact location of j will depend on the preference structure within the constituency). In contrast, our hypothesis does not preclude the decline of the vote in the set of winners - up to the kth candidate. But for those races where more than k candidates compete, our candidate-based model predicts a flat portion in the distribution of the vote between the kth and $k+1^{st}$ candidates, after which the vote should drop to zero, i.e. no other electorally motivated candidates in excess of $k+1$ should exist. Of course, since things other than electoral motivation may activate some candidates, it is sufficient to establish that a "sharp" decline in the vote between the $k+1^{st}$ and $k+2^{nd}$ candidates makes the later a "non-serious" competitor.

With this structure to our empirical analysis we can focus on the two most straightforward products of an election: the number of competing candidates and the distribution of vote among them.

---

[2] It is appropriate here to recall what assumptions were made in Part I, and what assumptions underlie Cox's model. We require that: 1) voters have single-peaked preferences over a one-dimensional policy space, and vote sincerely for the candidates nearest their ideal points; 2) each candidate maximizes his/her probability of winning a seat; 3) candidates can adjust their spatial positions, and a new candidate can enter, but only if doing so secures a positive probability, and 4) any two candidates must be spatially separated by a minimal distance $\delta > 0$. In addition to derive the results we require the voter ideal points in the district to be continuously, unimodally and symmetrically distributed. Cox when developing a voter-based model assumes that: 1) number of candidates is finite, and their policy platforms are fixed; 2) there is at least one voter for each possible preference configuration (only strict preference ranking is allowed); 3) voters vote strategically, maximizing their expected utility; 4) distribution of voter preferences is common knowledge, and all voters have identical expectations about the number of votes each candidate is going to receive.

3

However, we emphasize that it is not our intention to select any single hypothesis at the expense of another. Rather, it is to study how H1 and H2 interact and when relying on one or another better serves our purposes. Again, because both sets of results are based on stringent assumptions, we are not in the position to reject the validity of either model. Whenever evidence does not support a hypothesis, an argument can be made that it is due to the specific set of assumptions not being met by the data. With this qualification in mind, we turn first to Japanese electoral data, since Japan over the post-war period has accumulated a large body of internally consistent electoral information (consistent in the sense of district boundaries being largely unaltered and electoral rules largely unchanged).[3]

3. THE NUMBER OF CANDIDATES

Measurement problems: In developing his hypotheses about the number of "serious" candidates that compete in Japanese legislative elections, Reed counts candidates by computing their "effective" number, where this number corresponds to the inverse of a sum of the candidates' squared shares of the vote:

$$EN = \frac{1}{F} = 1/\sum_{1}^{n} \frac{(vote_i)^2}{(vote\ in\ the\ district)^2}$$

EN, then, is the inverse of a fractionalization index F, that, for the same number of candidates, increases as the candidates' vote shares become more uneven. The minimum of F for a given number of candidates is attained when their vote shares are identical. At this point EN equals the actual number of contestants. But this circumstance corresponds to EN's maximal value, so that for all other distributions of votes, the "effective" number of candidates assumes a value below the "real" number (with the bias increasing as differences in the candidates' vote shares increase).

Although convenient in some applications, "effective" counting does not help us study the behavior of candidates in the present context. A downward bias may be produced by differences in the vote for the highest ranked winners, whereas our attention is directed at competition down the line - among the lower ranked winners, losers, and potential entrants. In fact, the downward bias of the "effective" measure can be so significant, that it can result in counting fewer candidates than there are seats in the district. Consider an example where district magnitude equals four, and five candidates receive votes in a following proportion:

---

[3]   Data is taken from: Reed, Steven, *Japan Election Data: The House of Representatives, 1947-1990*, Ann Arbor, Center for Japanese Studies, 1992. Electoral data on 1993 Japanese elections are courtesy of Professor Gary Cox, University of California at San Diego.

candidate 1: 40%
candidate 2: 20%
candidate 3: 16%
candidate 4: 12% + ε

---

candidate 5: 12% - ε

Despite the fact that candidates *3, 4* and *5* are competing closely for the last seat, the "effective" measure that Reed employs returns the value $EN = 3.93$, which is less than k and which leads to the loss of the very candidate whose behavior interests us most - the $k+1^{st}$ candidate. There is, moreover, nothing extraordinary in the distribution of the vote in this example. Consider Japan's *Niigata 3rd* District in 1972:

Tanaka Kakuei	*182681*
Kobayashi Susumu	*58217*
Miyake Shoichi	*55363*
Murayama Tatsuo	*48329*
Ono Ichiro	*39867*
Furukawa Hisashi	*30747*
Magai Hideji	*18944*

The "effective" number of candidates in this race, *4.22*, is less than the size of the district. Similarly, in the *Kagawa 2nd* District, in 1980, we have for the three seats being filled,

Morita Hajime	*151546*
Kubo Hitoshi	*44027*
Kato Tsunetaro	*34535*
Kubo Fumihiko	*9829*
other candidates	*3016*

EN=2.28 is again below district size.

Not only can *EN* conceal competition, it can indicate competitiveness where it is not present. For instance, Japan's Communist Party has a policy of running a candidate in each district regardless of his or her chances of electoral success. In the majority of cases these chances are zero. *EN*, though, often counts these candidates despite their electoral hopelessness. For example, in the *Toyama 2nd* District in 1972:

Kataoka Seiichi	*76522*
Watanuki Tamisuke	*64950*
Sano Kenji	*62954*
Yoshida Minoru	*62355*
Arima Shigekazu	*9769*

Competition for the last seat clearly occurs here, but the fifth candidate is not part of it. Nevertheless, $EN=4.27$, which indicates more than $k+1$ relevant candidates.

Reed (1990) employs EN to avoid an arbitrary definition of "serious" candidates. But because "effective" counts do not carry the information we need, we must find a more satisfactory way to divide candidates into serious and non-serious categories. To that end notice that to the extent that competition among losing candidates focuses on the last seat in the district, we might expect such candidates to win comparable shares of the vote. Hence, if we order all candidates in the descending order of their vote, a sudden "significant" drop in the vote among the losers would indicate the end of the string of "serious" candidates (corresponding to the predicted drop in the distribution of the vote, as in Figure 1). The arbitrary thing to decide with this approach, though, is exactly how significant this drop should be. Since we have no theoretical basis for operationalizing "significant" one way or another, we suggest marking this drop after the candidate whose nearest competitor is 20, 33, or 50 percent behind him in terms of vote shares. Of course, that competitor should also be among the losers, as we assume that victorious candidates compete not among themselves, but with those without legislative seats.

Figure 2a shows the frequencies by the year with which a 20-percent drop in the vote occurs immediately after the kth candidate, after the $k+1^{st}$, $k+2^{nd}$, $k+3^{rd}$, or later candidate (Figures 2b and 2c consider respectively the location of the first 33-percent and 50-percent drop in the vote). A substantial portion of districts supports only as many competitive candidates as there are seats, whereas the dominant number of districts has just one competing candidate in excess of district size.

If we now take the number of candidates before the vote becomes non-competitive (i.e. drops substantially for the first time outside the set of winners) as a measure of the number of "serious" candidates, then Figure 2 allows us to make certain preliminary observations supportive of hypothesis H2. For example, we are interested in detecting the presence of a k-equilibrium in which the vote drops significantly immediately outside the set of winners. Reed's analysis does not make allowances for this possibility, whereas Cox predicts it will not happen. According to our model, however, it can occur for certain types of voter preference distributions. Hence, if we observe a persistent presence of k-candidate races, then we have one argument in favor of a candidate-based model. In Figure 2, then, we see that when the 20% criterion is applied, k-candidate races take place in 20 to 30 percent

of all electoral districts. The combined weight of k and $k+1$ candidate races by the same criterion lies - in different years - between 65% (in 1969) and 87% (in 1980, 1986, and 1990).

Naturally, strengthening the threshold of candidate "seriousness" increases the number of counted candidates in the race, so from one diagram to another the-shift should be toward the greater weight of the higher count. But the very presence of k-candidate elections even with a threshold of 50 percent means that in up to 10 percent of all races the vote drops by more than one half immediately after the last winning candidate (and in some cases the drop in the vote is as significant as 70-80 percent or more). The combined weight of k and $k+1$ races by this stronger criterion is between 33% (in 1958) and 67% (in 1980).

The pattern of one or fewer extra candidates/district: The average number of "serious" candidates in excess of district size is shown by the years of elections (1958 to 1993) in Figure 3.[4] Notice that with the relatively liberal 20% criterion, almost exactly $k+1$ candidates compete on average, whereas with the more stringent 50% criterion this number still does not go above $k+2$ in all but two elections.[5]

Our way of counting "serious" candidates is, probably, unfavorable to the hypothesis we advocate. Specifically, comparing the vote of each loser to the vote of the candidate immediately ahead of him may result in counting candidates with low vote shares, especially if the descent of the vote among the losers is smooth. Changing the way we count candidates to counting only those losers who secure vote comparable to that of the last winner allows us to better assess the competitiveness of each individual candidate. So drawing the line of "seriousness" when the fraction of the last winner's vote that the loser gathers drops below 80, 67 or 50 percent respectively (same 20, 33 and 50 percent thresholds, but applied in a different context), disregarding now closeness to the immediate frontrunner, yields an alternative count of the number of "serious" candidates. As Figure 4 shows, the

---

[4] Taking the average number of "excess" candidates allows us to avoid separating the results for 1, 2, 3, 4, and 5 member districts - something that both Reed and Cox were forced to - being at the same time consistent with our conceptualization.

[5] It is worth mentioning here that our theoretical analysis predicts the equilibrium number of candidates of either k or $k+1$ only if voter preferences are distributed unimodally, i.e. if the constituency is cohesive. This prediction does not have to hold in polarized constituencies, where $k+2$ and higher numbers of candidates can compete (see Appendix). Later we show that in the constituencies where we are not likely to expect polarization of voter preferences, the average number of candidates in the race is indeed much lower than in the districts where polarization is most likely, and basically stays at $k+1$ or below.

Notice also that the decline in the number of candidates across time is much less evident than Reed asserts on the basis of "effective" counts. If learning and adjusting to the peculiarities of SNTV occur, it most probably occurred in the period 1947-1958.

average number of excess candidates moves closer to predicted "one or below" level for stronger thresholds. However, there is no dramatic difference between the graphs generated by this and the previous, less favorable measure (compare Figures 3 and 4). This fact points at the existence of an actual gap between the candidates in the race, and indicates that steady decline in the vote outside the set of winners does not occur frequently.

Differences in the losers' competitiveness depending on their ranks: To illuminate what happens to candidates' vote shares after all seats are filled, consider Figures 5a to 5d, which show how the vote for each successive loser (the $k+1^{st}$, $k+2^{nd}$, $k+3^{rd}$, and $k+4^{th}$ candidates) is distributed as a fraction of the last winner's actual vote in that district. Most $k+1^{st}$ candidates, as Figure 5a shows, stay close to the last winner. This is but another way to say that first losers are generally competitive and pose a serious threat to the last winners. At the same time, both $k+3^{rd}$ and $k+4^{th}$ candidates (Figures 5c and 5d) concentrate at the lower end of the distribution, being thus out of the competition for seats. Of course, this assumes that the third and fourth losers are present in a district, and in many cases they are not: 592 out of 1632 observations lack the third loser, while 1106 out of 1632 lack the fourth one. We can conclude, then, that neither the third nor the fourth loser is likely to compete seriously for a seat under SNTV.

The second loser is an intermediate case, and his electoral strength is of the greatest interest to us. As Figure 5b shows, the $k+2^{nd}$ candidates distribution is generally fairly uniform with some tendency toward bimodality. The left mode of this distribution corresponds to districts with the second loser being effectively out of competition, and constitutes evidence in favor of the prediction of "one or fewer" extra candidates per race. But right the mode of this distribution signals that in a substantial number of races second losers are competitive and pose a threat to the last winner. Thus, we cannot ignore the fact that there is substantial number of $k+2$ candidate races in our data. At this point we can raise a claim that this $k+2$ pattern characterizes those constituencies where the assumptions of our model are not met. This is the issue to which we now turn.

Separating constituencies by the type of distribution of ideal points: The data summarized by Figure 5b causes us to address the question: how good are the model's predictions when we have greater confidence in the restrictive assumption about voter preferences? Our assuming unimodality can in fact explain why our hypothesis may not hold for polarized constituencies or constituencies with a broad range of preferences (with a uniform portion at the mode of preference distribution). The appendix to this paper illustrates some equilibria with more than $k+1$ "serious" contenders in non-cohesive constituencies.

A natural test, then, is to compare two samples - one formed of constituencies that are unlikely to be polarized, and another with more diverse or polarized preferences. Fortunately, basic

8

demographic character is one piece of information about Japanese electoral districts that is part of the data set. Specifically, although Japan is not ethnically or linguistically heterogeneous, we are more likely to find issue polarization in urban districts than in rural ones, if only because of the greater income, class, educational and occupational diversity found there. Hence, it is not unreasonable to suppose that our model will hold better in rural districts than in metropolitan ones.

The prediction that rural districts would follow the pattern of at most $k+1$ "serious" candidates closer than metropolitan ones is borne out by the data.[6] The number of excess candidates competing in Japanese electoral districts with different demographic characteristics is plotted in Figure 6. This is the number obtained by the first method described in Section 1, in which "serious" candidates are counted until the vote falls sharply between two adjacent candidates. The average number of excess candidates in rural constituencies is distinctly lower than in metropolitan ones. And surprisingly, the stronger the threshold used for counting "serious" candidates, the larger is the gap between rural and metropolitan districts.

And if as an alternative we determine "serious" candidates by comparing their vote shares to the share of the "weakest" winner, then not only does the separation between the number of candidates in rural and metropolitan constituencies persist, but the number of excess candidates in rural constituencies settles firmly at the $k+1$ predicted mark, regardless of the threshold we use (see Figure 7).[7]

Finally, and in accord with hypothesis $H2$, if we now analyze the relative strength of the $k+2^{nd}$ candidate (i.e. the second loser) in rural and metropolitan constituencies, we get two distinct distributions. As Figure 8 shows, in rural districts most such candidates do not pose effective competition for the kth seat (conforming to the hypothesis of at most $k+1$ true competitors in the race). In metropolitan constituencies, on the other hand, the $k+2^{nd}$ candidate is generally a competitor for a seat.

4. CANDIDATE VOTE SHARES

A direct comparison of H1 and H2: Counting the number of candidates alone is not sufficient to discriminate between H1 an H2. Although k-candidate races are allowed only under hypothesis $H2$,

---

[6] As a null hypothesis we take equal probability of $k+(>1)$ candidate races in both rural and metropolitan districts. This hypothesis is firmly rejected already at the 98% confidence interval. Probability of a $k+(>1)$ candidate race in a metropolitan district exceeds the one in a rural district by at least 20%.

[7] What the test of hypotheses described above is concerned, its result sustains when alternative measure is used.

there is still a pooling effect of both models with respect to the $k+1$-candidate elections. However, some separation is achieved if we consider the relationships between the vote for certain candidates: Cox's model predicts that the vote shares of the 1st and k^{th} candidates will be approximately equal (part B of $H1$), whereas the candidate-based model sets little restriction on the relative vote of winners (part B of $H2$).[8] On the other hand, the candidate-based model requires that the kth and $k+1^{st}$ candidates in the $k+1$-candidate elections be "close".

To assess these predictions Figure 9 presents the corresponding ratios of the vote, where the average ratio of the $k+1^{st}$ vote to the kth is computed conditionally on an electoral districts having more than k "serious" candidates in them (ratios are computed for each of the "seriousness" thresholds used in the previous section). But even if we do not control for the k-candidate districts, the vote for the kth and *for the* $k+1^{st}$ candidates stay closer together on average than do the votes within the set of winning candidates.

The dynamics of these two vote ratios is meaningful: the first, v^k/v^1, says how prudent voters are in allocating their vote; the other, v^{k+1}/v^k, indicates how precisely the candidates evaluate their chances when deciding about entering the contest. We can suppose that whichever side - voters or candidates - produces the faster convergence to the predicted level is the faster learner and, potentially, the Stackelberg leader in the game. As nothing prevents us from having both candidates and voters being strategic, it then becomes important to know who learns their equilibrium strategies first, and who thereby sets the structure of the system that then persists in future elections.

Accommodating a populist vote: The counter argument to using the described vote ratios as evidence for or against the two competing theories is that the kth and 1st candidates usually are more than one step apart. In other words, unlike the $k+1^{st}$ and the kth, these candidates are not by definition immediately close to each other. Moreover, we should be concerned that the 1st-ranked candidate assumes that position because of some special, non-spatial appeal. The argument in either case would be that relying on the ratios v^k/v^1 and v^{k+1}/v^k biases our conclusions against Cox's analysis.

Thus, to reduce the potential impact of a heavy "populist" vote for the leading winner, suppose we compare the drop in the vote between v^1 and v^k versus the corresponding drop between v^2 and v^k. Depicted in Figure 10, the difference in these two ratios confirms that, indeed, frontrunners collect

---

[8] The only restriction is that their vote should not differ more than twice (for small δ - if δ goes up, so does the allowed variation in the vote).

10

a disproportionate share of the vote. But even recalculated this way, the ratio of the vote within the set of winners remains below the ratio v^{k+1}/v^{k}.[9]

5. THE EVIDENCE FROM TAIWAN

Number of candidates: Returning to our original count of "serious" candidates (when we count all candidates before the first substantial drop in the vote among the losers), we now see (Figures 11, 12) that, although less stable, a pattern similar to the one we observe in Japan arises in Taiwan as well.[10] This regularity in SNTV elections - one extra competitor over the district size - can again be taken as evidence in favor of **H2** and a candidate-based model. This is especially so if we recall that every Taiwanese election corresponds essentially to an alteration in electoral regime - for the first three elections it was the gradual change in restrictions on campaigning; for the two later cases, disruption derived from the drastic changes in district boundaries and district magnitudes as well as the rules on allowable campaigns.

This regime instability suggests that we cannot expect to infer much here about learning by candidates and voters. However, it seems that the statistics generated by a 20% and a 33% criterion quickly settle close to the predicted number of candidates, while statistics generated by the 50% criterion, although starting high, rapidly approach the other two. This circumstance points in the direction of "quantitative" rather than "qualitative" learning. Namely, the number of "serious" candidates does not change much over time, while the gap in votes between them as a group and the others in the race deepens. Slow descent of the vote in early Taiwanese elections suggests that the second measure of the number of "serious" competitors - counting those candidates whose vote is comparable to the minimal winning vote in their districts - may produce results more consistent across different thresholds. This expectation is confirmed by Figure 13. When we only count losers whose vote is comparable to the minimal winning vote in their district, the number of excess candidates stays mostly at the predicted "one or below" for the 20 percent threshold, and all three thresholds generate comparable numbers.

---

[9] Another place to use the per seat decline in the vote is to see if there is a discontinuity in the safety of the $k-1^{st}$ and the kth candidates that reflects the fact that one is "safe" while the other is under the pressure of the competition. In fact, we find the vote ratio of the kth to *the* $k-1^{st}$ vote slightly lower than the ratio of the first loser's to the last winner's vote when k-candidate races are excluded from the sample. The null hypothesis that the two ratios are the same can be rejected in favor of the hypothesis that the ratio of the vote for two weakest winners is lower than the one of the vote for the first looser to the vote for the last winner. Taking into account the higher denominator of the first ratio, this translates into substantially greater decline of the vote in absolute terms between the two last winners, than on the edge of the set of winning candidates.

[10] Taiwanese electoral data are courtesy of Professor Emerson Niou, Duke University.

The difference between the two ways of counting candidates, then, is much greater in Taiwan than in Japan. However, an interesting feature of at least the first three Taiwanese elections (1980, 1983, and 1986), is that no institutionalized (party) alternative to the Kuomintang existed, which implied minimal coordination within the opposition with respect to entry and campaigns. Hence, a large number of unaffiliated candidates could run, with each of them likely to either miscalculate his chances or to act in pursuit of non-electoral goals. Indeed, unlike Japanese elections between 1958 and 1993, candidate participation is high in Taiwan. Nevertheless, if one ignores crowds of non-serious candidates, the same pattern emerges as in Japan. The fact that electoral districts were so few at that time (eight, returning to ten candidates each), allows us to look at the by-district dynamics. The 20% drop occurs fairly early, and districts display some stability in terms of the number of "serious" excess candidates that is unrelated to the overall number of "actual" candidates in those districts (see Addendum to Part 1).

Distribution of votes: As with the number of "serious" candidates, Taiwanese elections generate a pattern similar to the one we observe in Japan in terms of candidates' relative strengths. Figure 14 shows the relation between the vote for differently ranked candidates. Recall that to correspond to our analytical predictions, the ratios of the vote should be close to one on the boundary of the set of winners (i.e. the first loser should have nearly as many votes, as the last winner), and somewhere below one within the set of winners (the first and last winners can have electoral support of substantially different size). In contrast, the implications of Cox's strategic voter model are met if the ratio of the vote within the set of winners is approximately one, and somewhere (although not necessarily "drastically") below one on the boundary of the set of winners. In fact, the values of various vote ratios stay close to the values we report for Japan, thereby reinforcing our conclusion about the superior performance of a candidate-based model.[11]

6. IMPLICATIONS FOR PARTY SYSTEMS UNDER SNTV

It follows from our theoretical analysis that candidates within a district have considerable incentive to coordinate strategies. Neither a mechanism of enforcement nor *ad hoc* cooperative assumptions are needed to explain the potentially coherent behavior among the winning candidates.

---

[11] Cox's vote ratio within the set of winners equals about .6, which is below the number for Japan. But Taiwanese electoral districts are on average larger, and, therefore, the ranks of the candidates being compared by this ratio are further apart. Elimination of the most popular candidate in each district (to neutralize any extreme populist vote) does not change the results as sharply as it did in Japan, but one explanation may be the lower number of nationally renown politicians competing in districts. And as in Japanese case, the ratio of the vote for the first loser to the vote for the last winner stays fairly high, averaging to 0.91 across elections.

The mutual goal of safe winners is to deter entry and, thus, to preserve the status-quo. And, at least in theory, all but one winning candidate are safe winners in equilibrium, while the candidates who are to become the last winner and the first loser cannot improve their situation as they already occupy the best available locations.

The existence of strong national parties then, even if they participate only by nominating or endorsing candidates, obscures the picture of within-district electoral competition. Indeed, the reason why candidates would play cooperatively in a non-cooperative game is the satiation they achieve once they have secured seats for themselves, as no single candidate can aspire to more than one seat. The common interest of candidates in a district pushes them to coordinate their actions across party lines, rather than to seek a partisan expansion. This is not true for parties. Parties may have candidate-like incentives only if they are unlikely to place more that one candidate per district into the parliament. Once parties become sufficiently strong to attempt an increase of the number of seats they control in the district, implicit cooperation among the set of winning candidates is threatened. National parties can introduce challengers in the districts that not only jeopardize the safety of their incumbents, but through the national campaign can disrupt equilibria in "unchallenged" districts as well.

For this reason factionalism in the Japanese party system may have been stabilizing, since factions would normally have not more than one candidate per district, and would resist nominations from their own party that could diminish the security of their candidates. Factions can also distance themselves from a party campaign if that campaign is not to their advantage.

Another feature of Japanese elections is the existence of independent candidates who are endorsed by one or more than one party. Their presence can also be viewed as a manifestation of the gap between district-level and national-level politics. Namely, a structure of voter preferences at the district level at odds with the national picture may require substantial adjustments in party political platforms, so that an independent candidate endorsed by the party would do better. Alternatively, the distribution of voter ideal points may be unusually "thin" where party platforms are spatially concentrated. As a consequence this residual vote has to be taken care of in cooperation with other parties through joint nomination of an independent candidate, so as not to invite organized entry.

Another important relationship - between parties and voters - is modified as well by the electoral imperatives generated by SNTV. SNTV provides strong incentives for parties to coordinate voter strategies within a district. Indeed, as we have seen in our theoretical development (Part I of this essay), successful candidates (and their respective local party organizations), once they anticipate enough support to win a seat, should be primarily concerned with preserving a "properly structured" competition within the district. Theoretically at least, electorally-minded candidates may wish to support other competitors in their district with some of their own vote to preserve the overall balance.

Alternatively, candidates may find it useful to suppress their own turnout so as to make entry less attractive. Either way, we are talking about candidates manipulating the effective distribution of voter preferences, which can be done if parties can influence voters strategies.

This later possibility, then - that parties might be the instruments of influencing the observable distribution of voter preferences - suggests yet another role for parties. Suppose, for a given district magnitude and preference distribution, that no equilibrium campaign configuration exists. We know, however, that for some derivative preference distribution - after adjustments in turnout, for instance - there may be an equilibrium. Parties then may be a tool for "correcting imperfections" in the distribution of preferences so as to generate a favorable equilibrium - one that maintains the viability of almost all incumbent politicians.

APPENDIX
DEPARTING FROM SOME ASSUMPTIONS: POLARIZED PREFERENCES

Part I of this essay does not treat cases of complex electoral preferences within districts, such as non-spatial issues or bimodally or multimodally distributed spatial preferences. Instead we assume that voter ideal points are unimodally distributed over the single-dimensional policy. But this assumption may not be met in many electoral districts included in our data. Voter preferences in some districts may correspond to this assumption closer than in the others. As was mentioned earlier, the structure of competition appears significantly different in rural and metropolitan constituencies. We believe that rural constituencies correspond better to our assumptions, and in fact the predictions of our model are met closer in those districts.

Even though we are not in the position to give an exhaustive description of what might happen in the non-unimodal cases, we can offer examples of equilibria for certain special distributions, that may prove generalizable - notably, some bimodal (polarized) distributions and quasi-concave configurations that begin to approximate a uniform distribution.

In Part I we establish the existence of $k+1$ equilibria (k odd) for all symmetric unimodal distributions and for distributions that approximate them "sufficiently closely". And although we do not know whether equilibria exist for other unimodal distributions, we know that if they exist, they cannot be but of the same form - either k, or $k+1$ with two tied candidates at the mode.

Now there are certain things that we can say about the number of candidates in polarized (bimodal) constituencies. First, building on the above mentioned result, if the distribution of voter preferences is bimodal and symmetric, as well as symmetric around each mode, then at least for $k=4a+1$ (where a is a positive integer) a $k+1$ equilibrium exists (see Figure A1). Without symmetry around each mode, a $k+1$ equilibrium may exist that looks like the configuration in Figure A2. Figures A3 and A4 illustrate $k+2$ equilibria. Note, that the configuration in Figure A4 is simply a combination of two $k+1$ equilibria for unimodal distributions. Note also that the same configuration is an equilibrium if district magnitude is five, six or seven.

The intermediate case between a strict unimodal distribution and a polarized (bi-modal) constituency occurs when the preference distributions has a flat region for a mode. In this case the number of candidates in excess of k that compete in equilibrium depends on the size of the district and the specific form of the distribution, because as we know from Part I, a uniform preference distribution can support a great variety of equilibria configurations. The existence of finite slopes in $f(x)$ and how candidates respond to them restrict this variety. Figures A5 and A6 illustrate the $k+1$ and $k+3$ equilibria.

REFERENCES:

Baerwald, Hans, *Party Politics in Japan*, 1986, London, Allen & Unwin.

Balinski, M.L., and H.P.Young, *Fair Representation Meeting thr Ideal of One Man One Vote*, New Haven, Yale University Press, 1982.

Cox, Gary W., "Electoral Equilibrium in Double-Member Districts", *Public Choice*, 1984, 44:443-451.

Cox, Gary W., "Electoral Equilibrium under Alternative Voting Institutions", *American Journal of Political Science*, 1987, 31: 82-108.

Cox, Gary W., "Centripetal and Centrifugal Incentives in Electoral Systems", *American Journal of Political Science*, 1990, 34: 903-935.

Cox, Gary W., and Emerson Niou, Seat Bonuses Under th e Single Non-Transferable Vote System: Evidence from Japan and Taiwan", *Comparative Politics*, 1993 .

Cox, Gary W., "Strategic Voting Equilibria under the Single Non-Transferable Vote", 1993.

Curtis, Gerald L., *Campaigning Japanese Style*, 1972 New York, Columbia University Press.

Denzau, A., A.Kats, and S.Slutsky, Multiagent Equilibria with Market Share and Ranking Objectives, *Social Choice and Welfare*, 1985, 2:95-117.

Feddersen, Timothy J., Itai Sened, and Stephen J. Wright, "Rational Voting and Candidate Entry under Plurality Rule", *American Journal of Political Science*, 1990, 34 1005-1016.

Greenberg, Joseph, and Shlomo Weber, "Multiparty Equilibria under Proportional Representation", *American Political Science Review, 1985, 79: 693-703.*

Filippov, Mikhail G., and Olga Shvetsova, "Ethnic Composition of Federal Units and the Survival of Democratic Federal States", presented at the APSA Annual Meeting, Washington D.C., September 1993.

Holler, Manfred J., ed., *The Logic of Multiparty Systems*, International Studies in Economics and Econometrics Volume 17, Boston, Kluwer Academic Publishers, 1987.

Horowitz, Donald L., *Ethnic Groups in Conflict*, Berkeley, University of California Press, 1985.

Hrebenar, Ronald J., ed., *The Japanese Party System*, 1986, Boulder, Colo., Westview Press.

Lijphart, Arend, *Democracies in Plural Societies: A Comparative exploration*, New Haven, Yale University Press, 1977.

Lijphart, Arend, *Democracies*, New Haven, Yale University Press, 1984.

Linz, Juan J., "The Perils of Presidentialism", *The Journal of Democracy*, 1: 51-69.

Mainwaring, Scott, "Presidentialism, Multipartism, and Democracy - the Difficult Combination", *Comparative Political Studies*, 1993, 26: 198-228.

Myerson, Roger B., and Robert J. Weber, "A Theory of Voting Equilibria", The Center for Mathematical Studies in Economics and Management Science, Northwesern University, *Discussion Paper No. 782*, 1991.

Ordeshook, Peter C., and Olga Shvetsova, "Ethnic Heterogeneity, District Magnitude, and the Number of Parties", *American Journal of Political Science*, vol. 38, 1994.

Palfrey, Thomas R., "Spatial Equilibrium with Entry", *Review of Economic Studies*, 1984, 51:139-156.

Palfrey, Thomas R., "A Mathematical Proof of Duverger's Law", in P.C.Ordeshook, ed., *Models of Strategic Choice in Politics*, Ann Arbor: University of Michigan Press, 1989.

de Palma, A., Gap-Seon Hong, and J.-F. Thisse, "Equilibria in Multi-Party Competition Under Uncertainty", *Social Choice and Welfare*, 1990, 7:247-259.

Reed, Steven R., "Structure and Behavior: Extending Duverger's Law to the Japanese Case", *British Journal of Political Science*, 1990, 20:335-356.

Reed, Steven R., *Japan Election Data: The House of Representatives 1947-1990*, Ann Arbor: Center for Japanese Studies, 1992.

Riker, William H., *Liberalism Against Populism: A Confrontation Between the Theory of Democrac and the Theory of Social Choice*, Illinois, Waveland Press, 1982.

Shugart, Matthew S., and John M. Carey, *Presidents and Assemblies: Constitutional Design and Electoral Dynamics*, Cambridge University Press, 1992.

Winkler, Edwin A., "Institutionalization and Participation in Taiwan: from Hard to Soft Authoritarianism", *China Quarterly*, 1984, 89:481-499.

Figure 1. Alternative Hypotheses Generated by Voter-Based (H1 – Cox 1993) and Candidate-Based Models
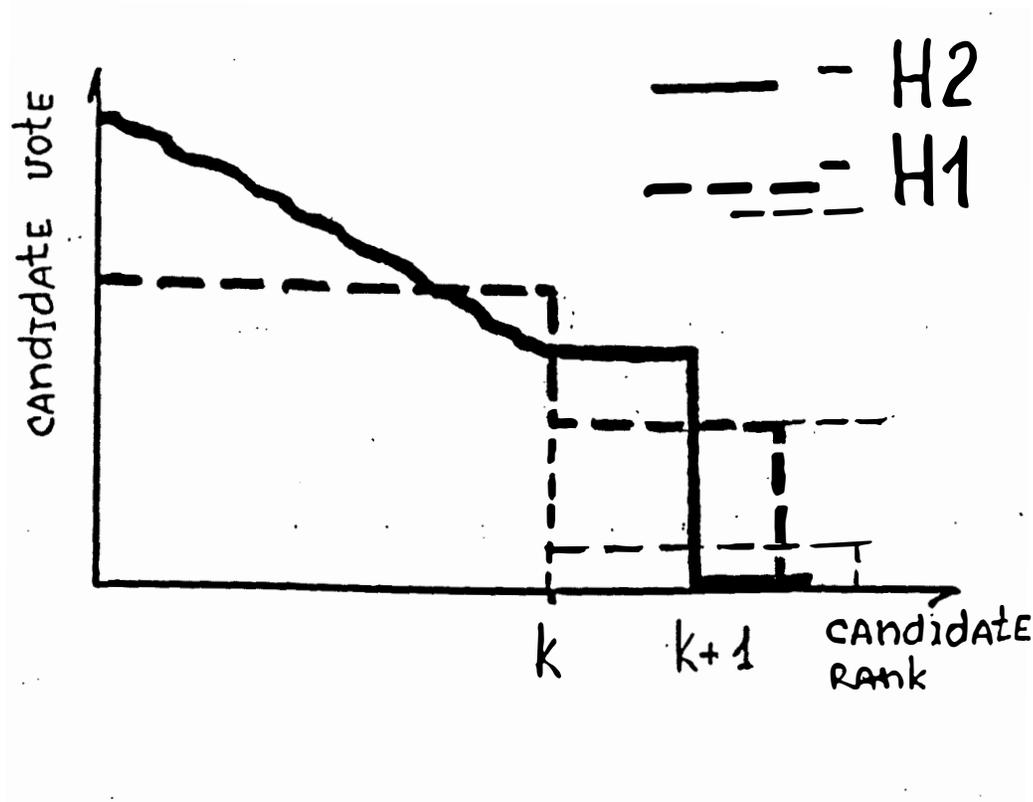
Figure 2. Location of the First Substantial Decline in Vote Outside the Set of Winners

2a: Considered decline is 20% or more
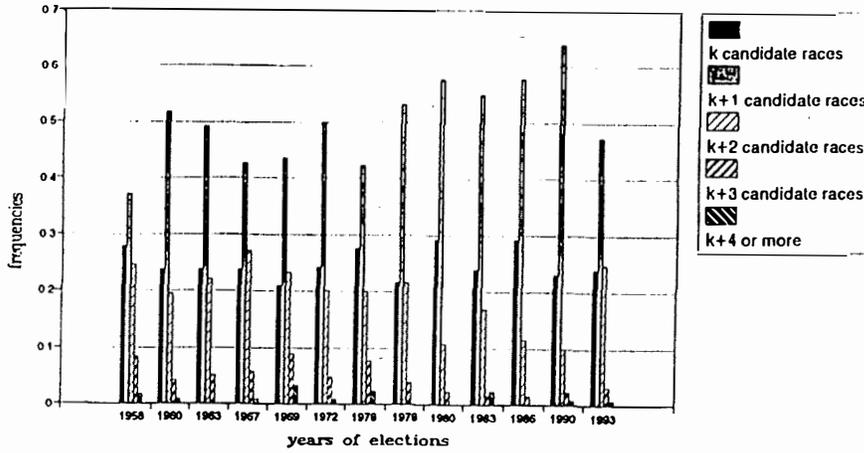
2b: Considered decline is 33% or more
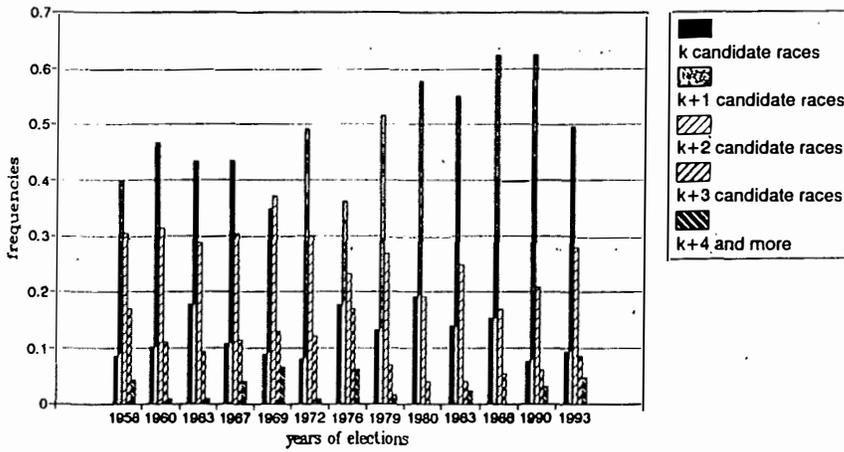
2c: Considered decline is 50% or more
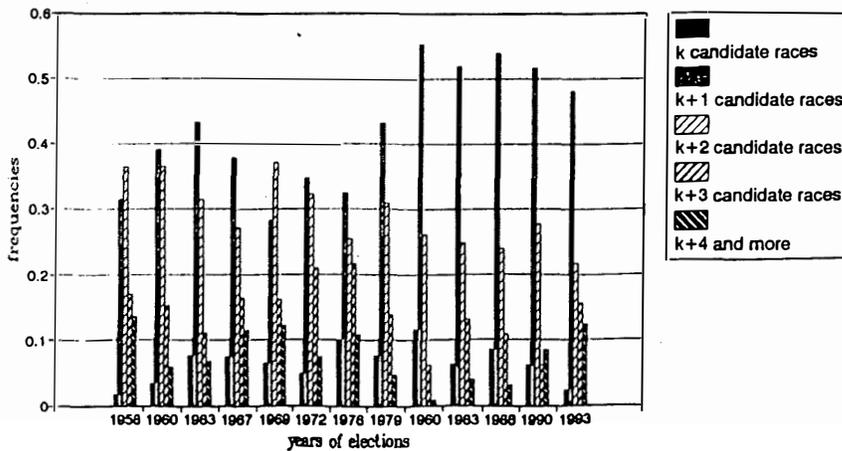
Figure 3. Average Number of "Serious" Candidates per Race in Excess of the District Size
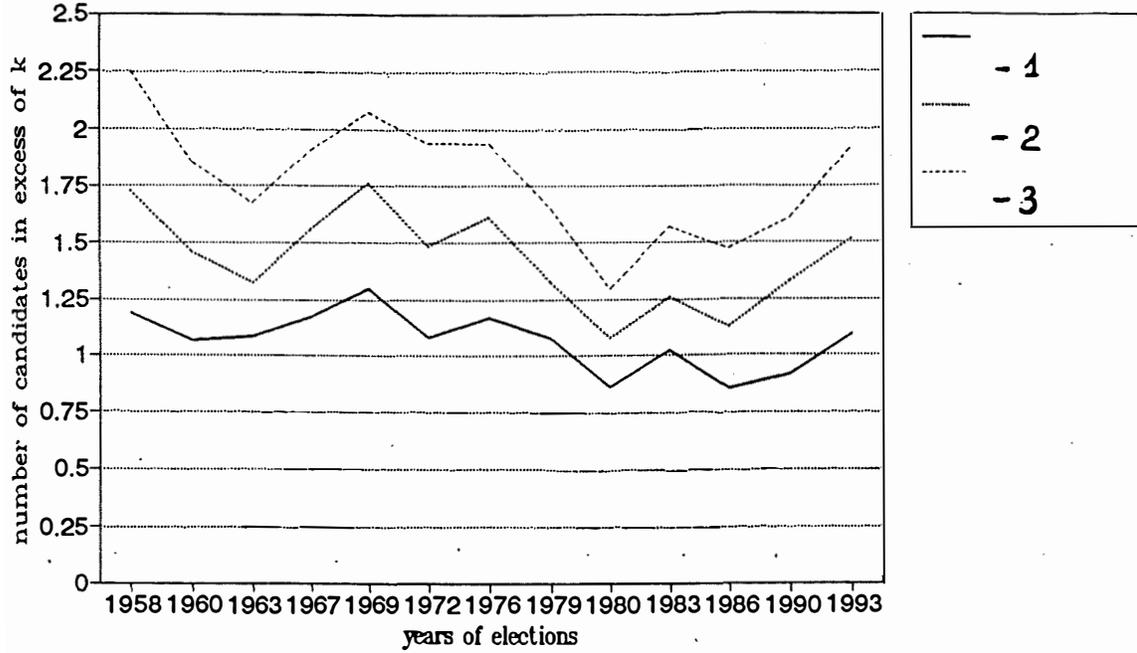
1 -   before the first 20% drop in vote;
2 -   before the first 33% drop in vote;
3 -   before the first 50% drop in vote.

Figure 4. Average Number of "Competitive" Candidates per Race
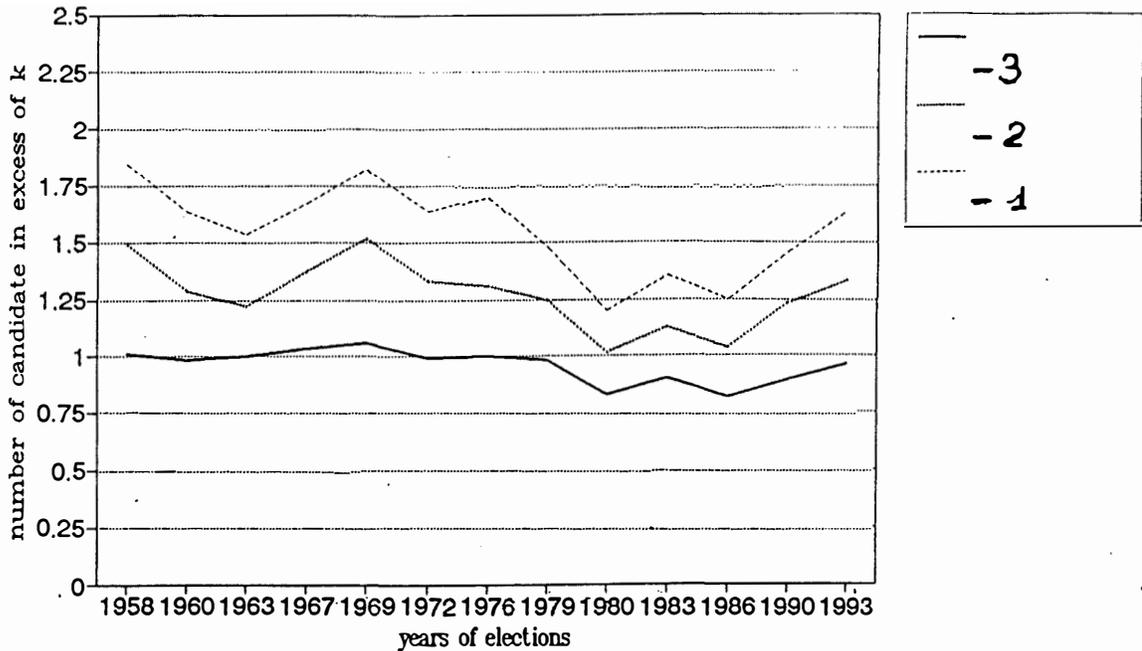
1 – considering losers with as little as 50% of the minimal winning vote in
their districts;
2 – with 67 or more percent;
3 – with 80% or more.

Figure 5. Performance of Differently Ranked Losers Relatively to Their Districts' Last Winner Vote - Japan

FIRST LOSER'S VOTE AS A PERCENTAGE OF LAST WINNER'S VOTE - JAPAN
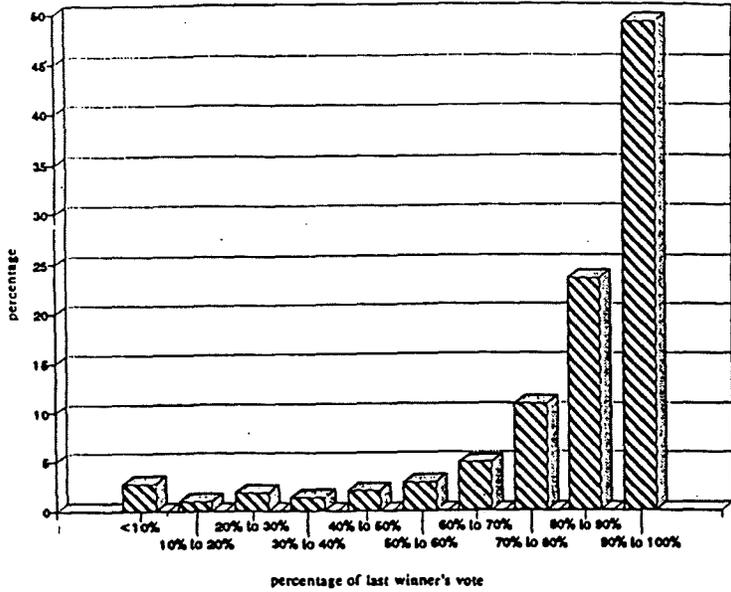
percentage of last winner's vote

SECOND LOSER'S VOTE AS A PERCENTAGE OF LAST WINNER'S VOTE - JAPAN

percentage of last winner's vote

THIRD LOSER'S VOTE AS A PERCENTAGE OF LAST WINNER'S VOTE - JAPAN
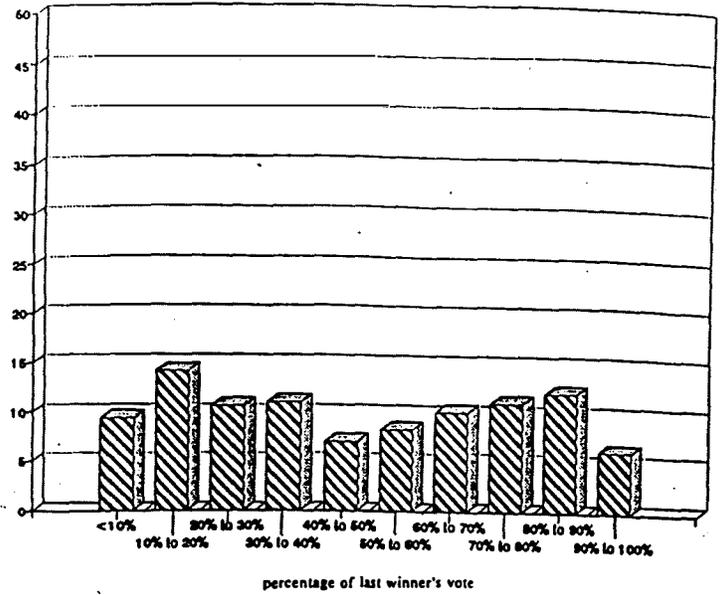
percentage of last winner's vote
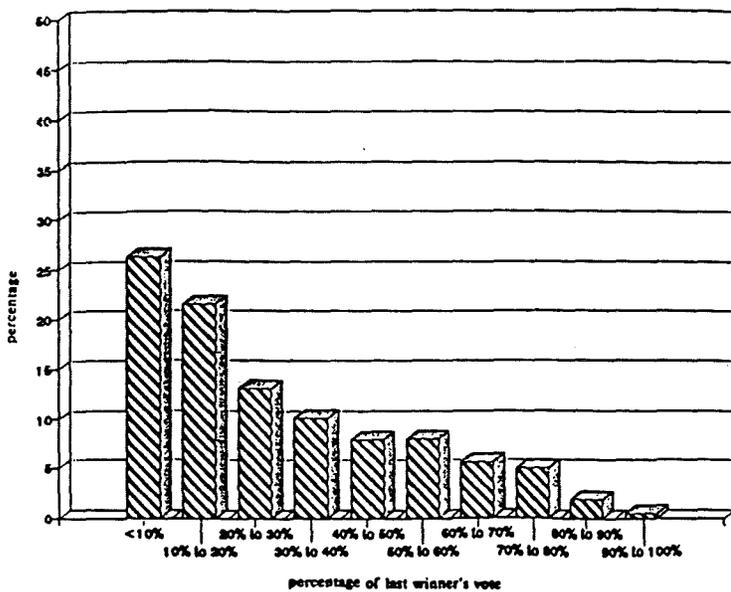
FOURTH LOSER'S VOTE AS A PERCENTAGE OF LAST WINNER'S VOTE - JAPAN

percentage of last winner's vote

22
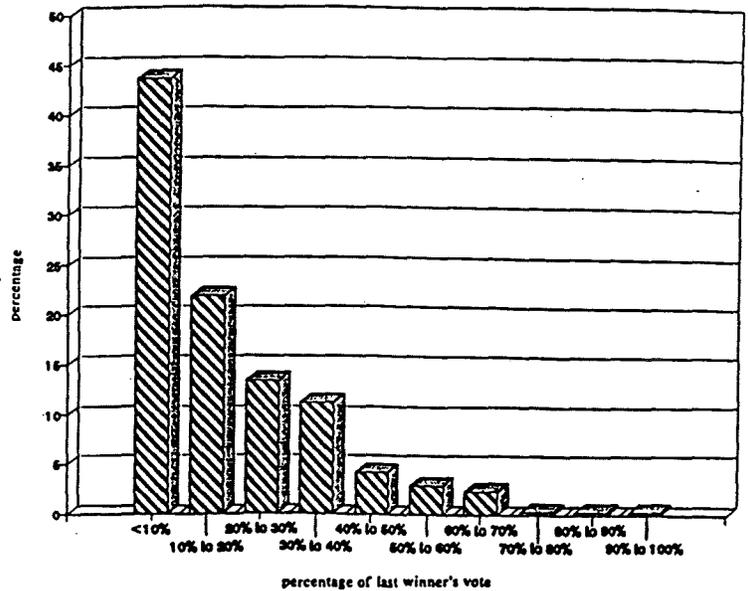
Figure 6. Average Number of "Serious" Excess Candidates before the First Substantial Drop in the Vote among Losers in Constituencies of Different Demographic Types - Japan

6a: Considering a drop of 20% or more
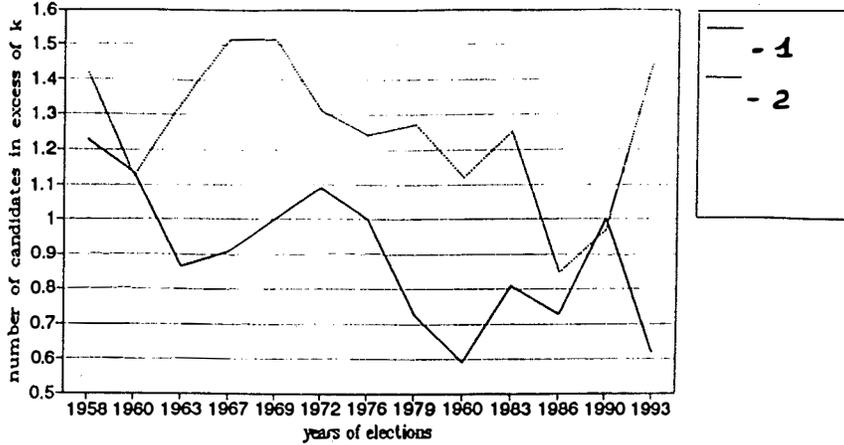
6b: Considering a drop of 33% or more
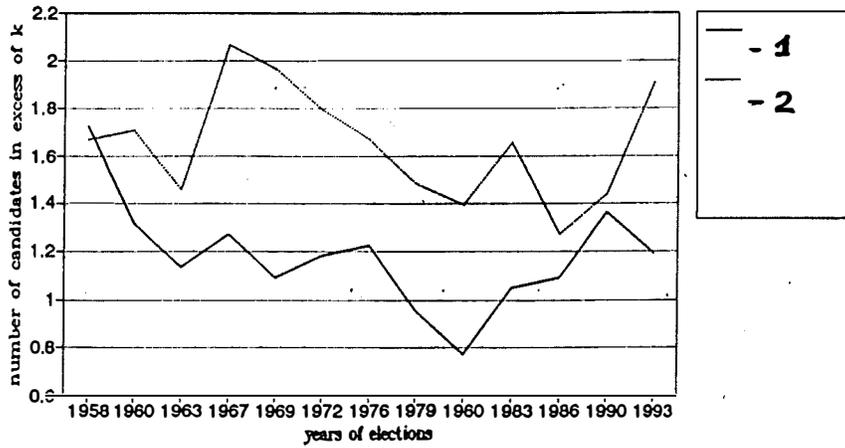
6c: Considering a drop of 50% or more
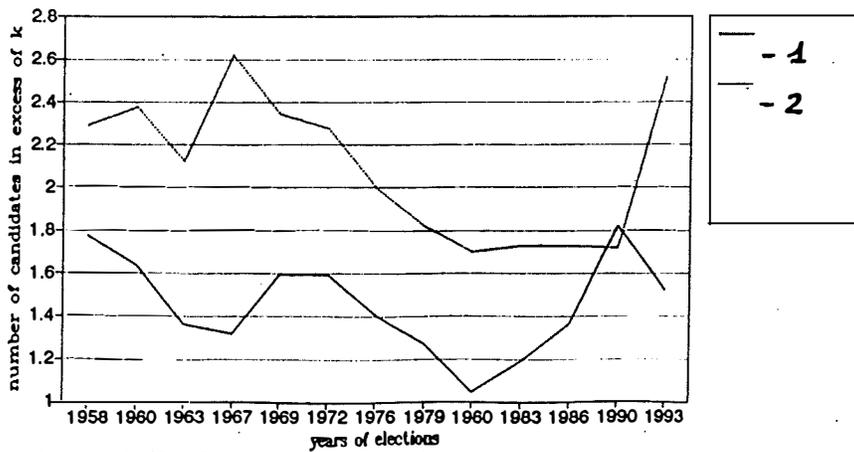
1 - in rural districts;   2 - in metropolitan districts

Figure 7.   Average Number of "Competitive" Losers per Race in Constituencies of
Different Demographic Types – Japan

7a: Considering candidates with 80% of their districts' minimal winning vote

7b: Considering candidates with at least 50% of their districts' minimal winning vote

1 – in rural districts;   2 – in metropolitan districts

24

Figure 8.   Second Loser's Vote as a Percentage of the Last Winner's Vote – Japan

8a: Rural constituencies

frequencies

vote gathered / the last winner's vote

8b: Metropolitan Constituencies

frequencies

vote gathered / the last winner's vote
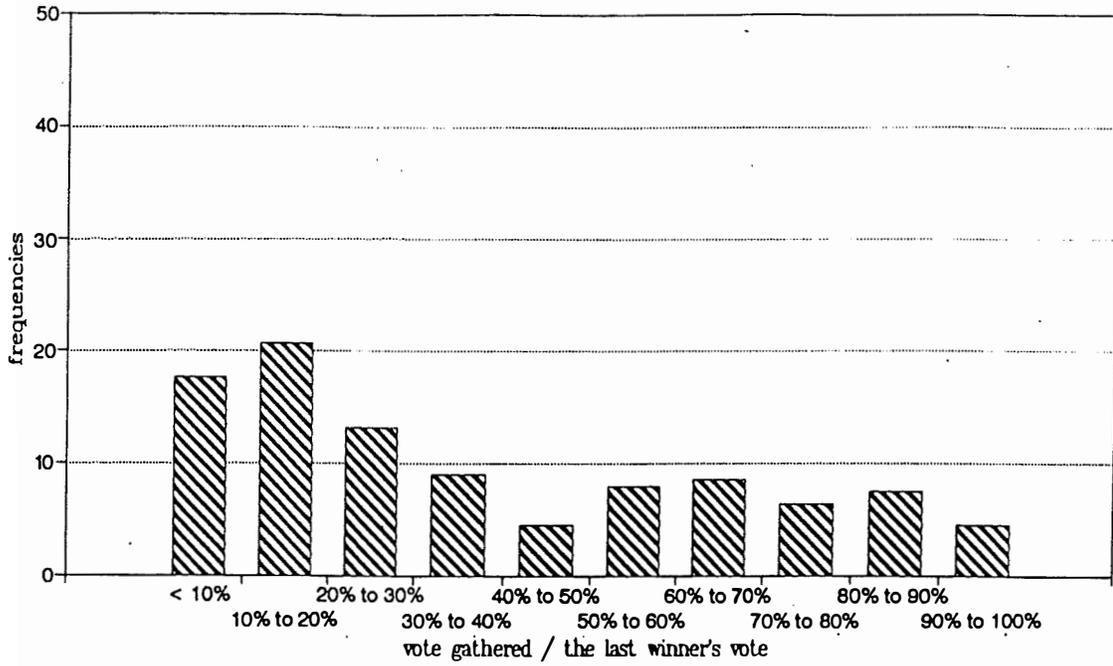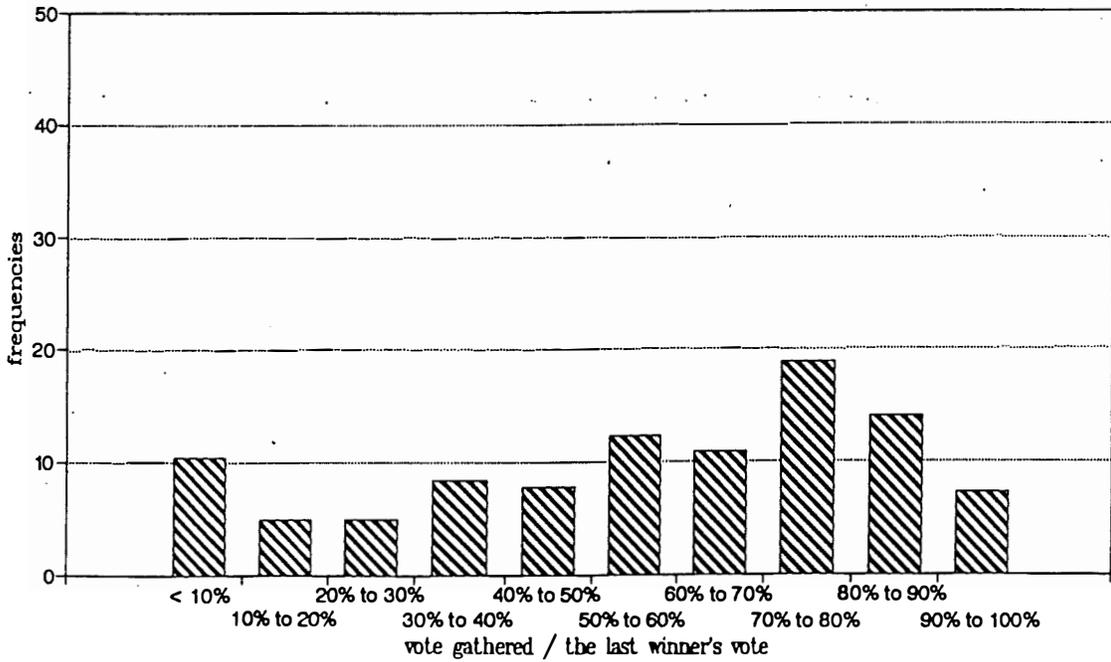
Figure 9.   **Comparison of H1 and H2 in Terms of Vote Dynamics**
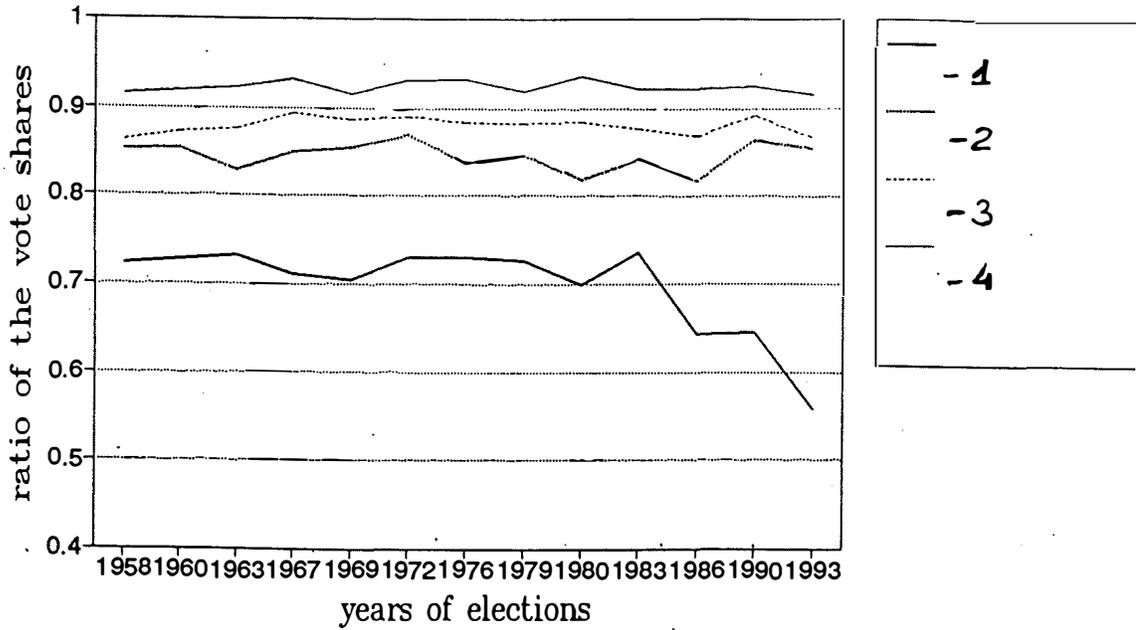
1 -   ratio of the vote within the set of winners - Cox 1993 (H1: $k^{th}/first=1$)

                   ratio of the vote for the first loser to the vote for the last winner:

2 -   in all races;

                 in competitive races

3 -   when the first loser is not more than 20% behind the last winner;

4 -   when the first loser is not more than 50% behind the last winner.

Figure 10.   Comparison of H1 and H2 Controlled for the Populist vote – Japan
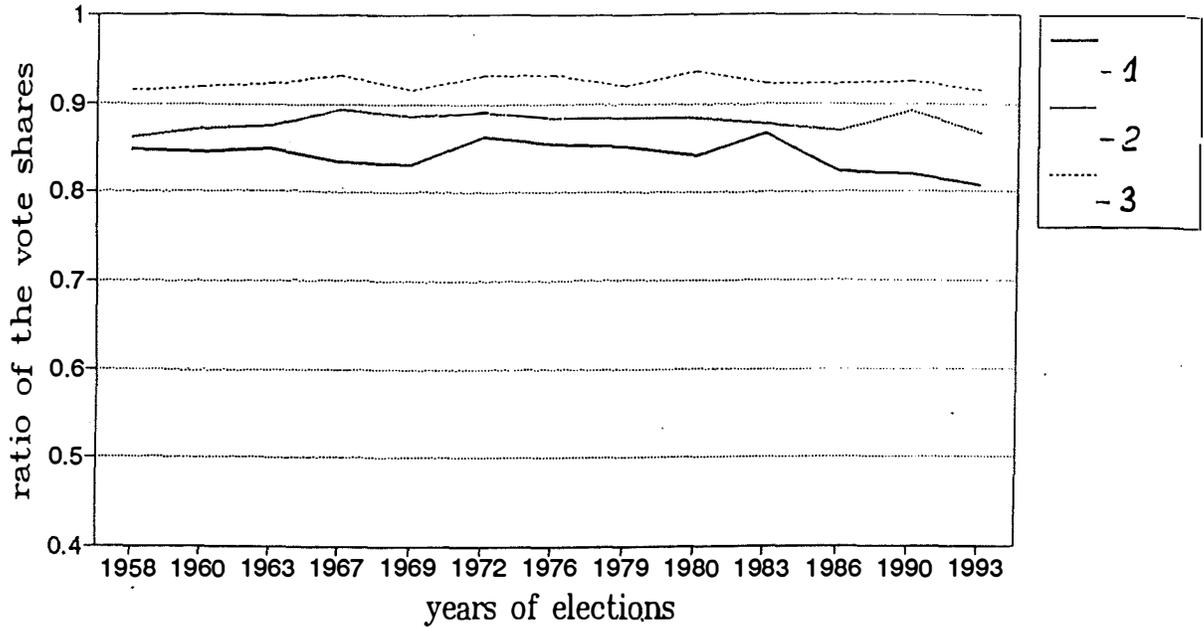
1 -   ratio of the vote for the last winner to the vote for the second winner;
          ratio of the vote for the first loser to the vote for the last winner
          in competitive races:
2 -   when the first loser is not more than 50% behind the last winner;
3 -   when the first loser is not more than 20% behind the last winner.

Figure 11. Location of the First 33% Decline in the Vote Outside the Set of Winners - Taiwan
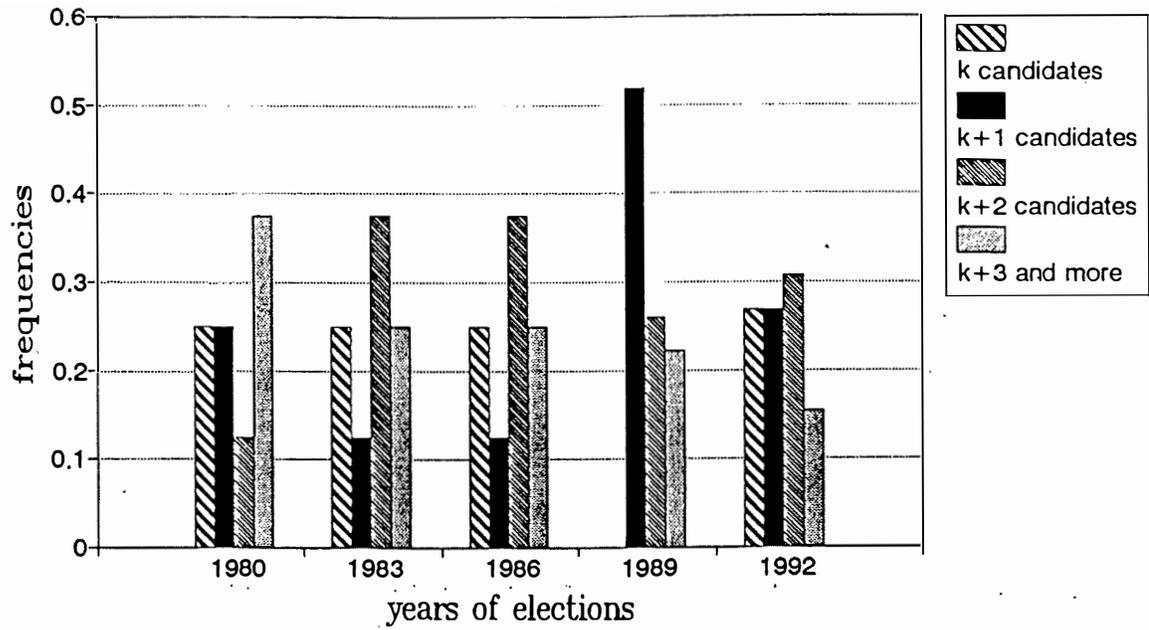
Figure 12. Average Number of "Serious" Candidates per Race in Excess of the District Size - Taiwan
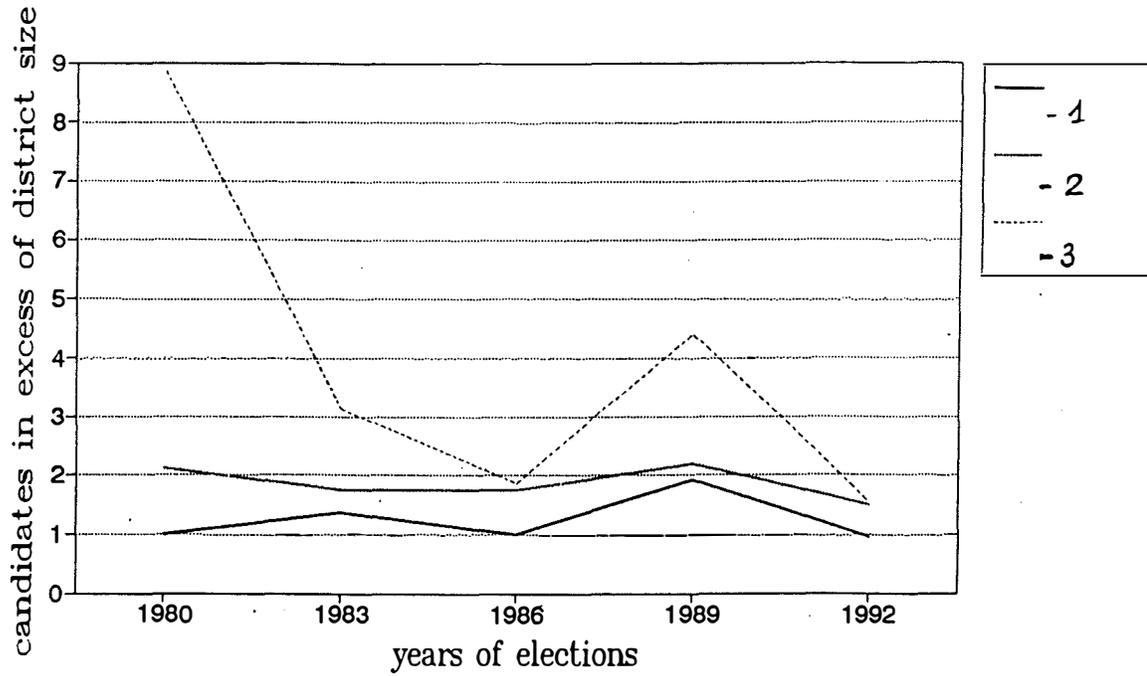
1 -   before the first 20% drop in the vote outside the set of winners;
2 -   before the first 33% drop in the vote outside the set of winners;
3 -   before the first 50% drop in the vote outside the set of winners.

Figure 13.   Average Number of "Competitive" Losers per Race – Taiwan
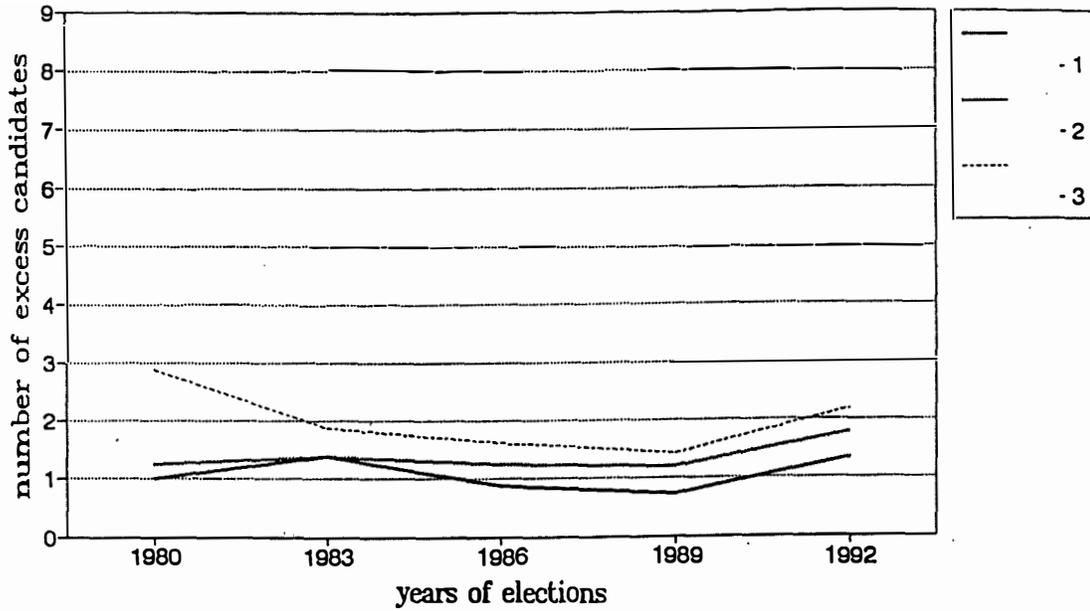
1 –   considering candidates with at least 80% of their districts' minimal winning vote;
2 –   considering candidates with at least 67% of their districts' minimal winning vote;
3 –   considering candidates with at least 50% of their districts' minimal winning vote;

**Figure 14.** Comparison of H1 and H2 (Vote Ratios) - Taiwan Data
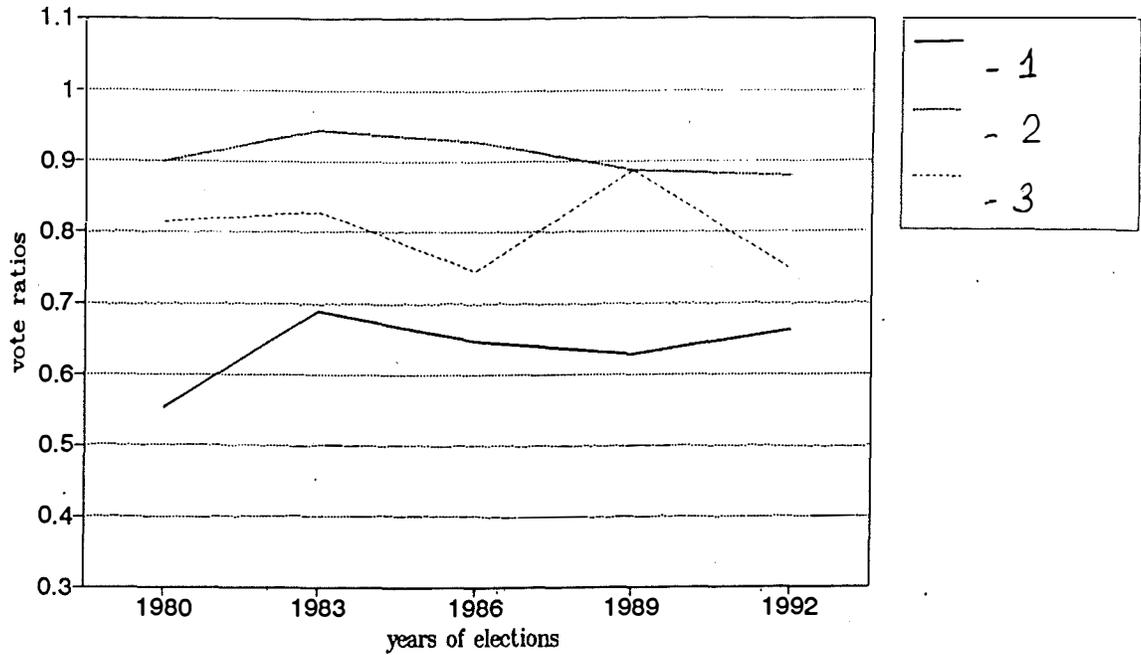
1 - ratio of the vote within the set of winners - Cox 1993 (H1: $k^{th}/first=1$)

2 - ratio of the vote for the first loser to the vote for the last winner: in all races;

3 - in competitive races when the first loser is not more than 20% behind the last winner.

Examples of Equilibria Configurations for Some Non-Cohesive Preference Distributions
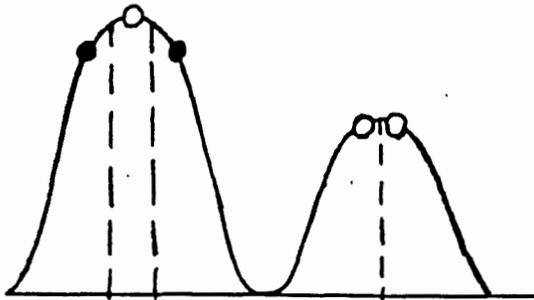
A1: Example of a *k+1* equilibrium existence for the case of district magnitude *k=5*.
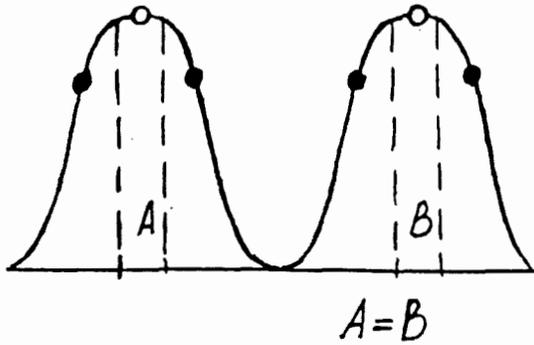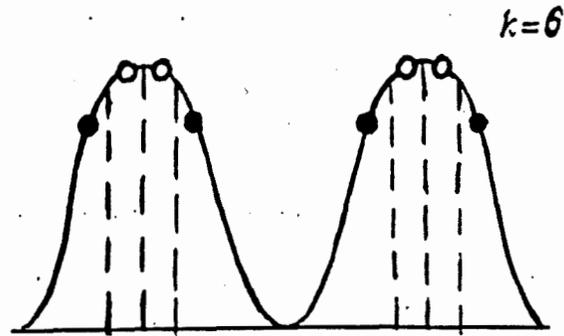
A2: Example of a *k+1* equilibrium existence for the case of district magnitude *k=3*.
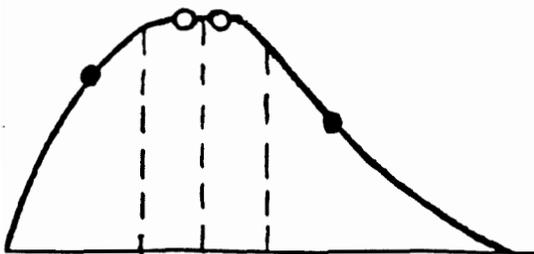

$$A = B$$
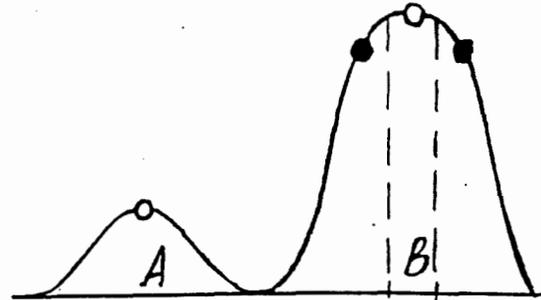
A3: Example of a *k+2* equilibrium existence.

A4: *k+2* Equilibrium as a combination of two *k+1* equilibria.

$$k=6$$

A5: Possibility of a *k+1* equilibrium when distribution of preferences has a flat portion on the top.

A6: Possibility of a *k+3* equilibrium when distribution of preferences has a flat portion on the top.

$$k=4$$