The degree of path-dependence and strategic voting

Aki Lehtinen

12. February 2012

Abstract

This paper generalises Enelow (1981) and Lehtinen's (2007) model of strategic voting under amendment agendas by allowing any number of alternatives and any voting order. It is shown that strategic voting increases utilitarian efficiency also when there are more than three alternatives. A criterion for evaluating path-dependence, the degree of path-dependence, is proposed, and the generalised model is used to study how strategic voting affects path-dependence. It is shown that the existence of a Condorcet winner does not guarantee path-independence if the voters engage in strategic voting under incomplete information. When there is a Condorcet winner, strategic voting inevitably increases the degree of path-dependence, but when there is no Condorcet winner, strategic voting decreases path-dependence. Computer simulations show, however, that on average it increases the degree of path-dependence. (JEL classification numbers: D71, D81)

keywords: strategic voting, path-dependence, amendment agendas

1 Introduction[1]

Social choices are path-dependent if the outcome of voting depends on the order in which the alternatives are presented for consideration. The existence of a Condorcet winner, an alternative that has a majority against all other alternatives, guarantees that social choices are path-independent if voters act strategically under complete information (McKelvey and Niemi, 1978) or if they vote sincerely (Farquharson, 1969). For these reasons, path-dependence is commonly thought to be closely connected to the cyclicity of the preference profile. Some authors have even proven results according to which

[1]For questions, contact me at aki.lehtinen@helsinki.fi.

cyclicity is a necessary and sufficient condition for path-dependence (List, 2004). However, cyclic preferences are not necessary for path-dependence under incomplete information because the Condorcet winner is not necessarily selected under amendment agendas. There are thus cases in which social choices are path-dependent even though there is no preference cycle, and strategic voting may exacerbate the problem of path-dependence.

On the other hand, strategic voting may alleviate the problem of path-dependence when there is a preference cycle. The reason for this possibility is related to Lehtinen's (2007) results on the consequences of strategic voting: it typically increases the chances that the utilitarian winner, the alternative with the largest sum of utility, is selected. If voters' preferences cycle over all alternatives, sincere voting results in a different outcome under all different agendas but strategic voting may lead to fewer different outcomes because it increases the chances that one particular alternative is selected in many agendas. Strategic voting cannot eliminate path-dependence altogether in these cases, however, because the utilitarian winner will inevitably lose against the Condorcet winner in the last round of voting under agendas that introduce the Condorcet winner only at this last stage, insofar as the former is not the same alternative as the latter.

Given that strategic voting may both increase and decrease path-dependence, it is natural to ask which tendency is more important. This paper investigates this question by comparing the *extent* to which social choices are *path-dependent* when they engage in strategic voting and when they vote sincerely. Conducting such an investigation with a formal model requires formulating a criterion for the degree of path-dependence, and constructing a model of strategic voting that can be applied to any number of alternatives and any (amendment) agenda.

The analysis of strategic voting is based on a computer simulation framework introduced by Lehtinen (2007*b*) and a model of incomplete information via signal extraction presented in Lehtinen (2006). The welfare consequences of strategic voting are studied by comparing the utilitarian efficiency under *expected utility maximising behaviour* (EU behaviour) and *sincere voting behaviour* (SV behaviour). Under the latter behavioural assumption, all voters always vote sincerely, and under the former they vote sincerely or strategically depending on their expected utilities. The incomplete information model is based on the idea that the voters obtain a perturbed signal on other voters' utilities. The main formal contribution of this paper is to generalise this model to any number of alternatives. Such a generalisation also allows investigating whether strategic behaviour increases utilitarian efficiency when there are more than three alternatives.[2]

The structure of the paper is the following. Section 2 provides a description of a modification of Enelow's (1981) model, on which Lehtinen's model is based. Subsection 2.1 describes how the modified model can be used to study any agenda. The generalisation to any number of alternatives is based on an indexing system (ordering numbers) for pairwise contests. The details of the indexing system are of interest mainly to those who are themselves interested in constructing similar computer simulations, and its description is relegated to Appendix A.[3] Since Enelow and Lehtinen's model was limited to three alternatives, voters did not need to take other voters' strategic behaviour into account. Section 3 describes how this can be done by remodelling voters' signals. Section 4 provides a brief description of the simulation framework. Lehtinen's (2007*b*; 2007*a*; 2008) main result is that strategic voting behaviour generates higher utilitarian efficiencies than sincere behaviour. These results are most salient when some particular alternatives are commonly considered to be acceptable, i.e., when they have higher average utilities than other alternatives even though the profile is created with the anonymous impartial culture assumption. Section 4 shows how to generate such computer simulation setups with more than three alternatives: the utility of one alternative (alternative 1) is increased at the expense of others without changing voters' preference orderings. Section 5 shows an example in which the Condorcet winner is not selected, and formulates a mathematical expression for the degree of path-dependence. Section 6 shows the simulation results.

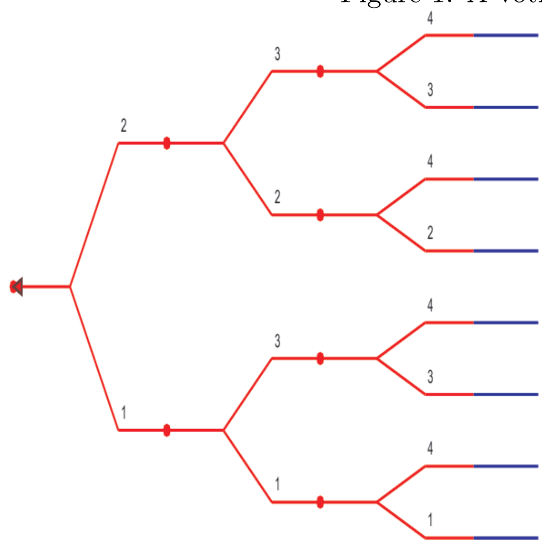
2 A model of strategic voting

Under *an amendment agenda*, two alternatives are put to a majority vote against each other in a first round of voting. The winner of this first contest is then put to vote against the third alternative in a second round, and so on. Enelow (1981) and Lehtinen (2007*b*) have presented a model of strategic voting under amendment agendas but these models were restricted to three alternatives. This model is here generalised by describing an indexing system that may be used to construct a computer model with any number of

[2]In practice, the results can be computed with seven alternatives at most. With 8 or more alternatives, the combinatorial explosion becomes unbearable for ordinary supercomputers. For example, there are n!/2 different voting orders. With 7 alternatives this is 2520 but with 8, it is 20160.

[3]I can make a reasonably well-documented computer code (in FORTRAN) available. In order to run the simulations, one needs access to a fairly new supercomputer and to IMSL libraries.

Figure 1: A voting tree



alternatives.[4]

Let $X = \{1, 2, 3, ...n\}$ denote the set of available alternatives, $I = \{1, 2, ..., i, ..., N\}$ a set of voters, and U^i voter i's payoff function. I will consider the case of four alternatives as an example. Figure 1 shows this agenda.

In what follows, I will call different voting orders simply *agendas*, and denote the agenda shown in Figure 1 as (1234). If, say, 3 is first put to a vote against 4, then the winner of this contest against 2, and the winner of this contest against 1, the agenda will be denoted (3421).

With four alternatives, there are $4! = 24$ different *types of voters*.

voter type	1	2	3	4	5	6	7	8	9	10	11	12
U^i_1	1	2	1	3	3	2	1	2	1	4	4	2
U^i_2	2	1	3	1	2	3	2	1	4	1	2	4
U^i_3	3	3	2	2	1	1	4	4	2	2	1	1
U^i_4	4	4	4	4	4	4	3	3	3	3	3	3

voter type	13	14	15	16	17	18	19	20	21	22	23	24
U^i_1	1	4	1	3	3	4	4	2	4	3	3	2
U^i_2	4	1	3	1	4	3	2	4	3	4	2	3
U^i_3	3	3	4	4	1	1	3	3	2	2	4	4
U^i_4	2	2	2	2	2	2	1	1	1	1	1	1

For each alternative and each individual, a utility number U^i_j is randomly

[4] See Ordeshook (1986) and Ordeshook and Schwartz (1987) for a discussion of different agendas.

drawn from the uniform distribution on (0,1). For the time being, let p^i_{jk} denote voter i's subjective probability that alternative j beats k ($j, k \in X$) in some node of the voting tree.

Maximising expected utility implies giving one's vote for that branch in the voting tree that has the greatest expected utility. In the case of three alternatives under agenda (123), voters compare two lotteries $(1, 3; p^i_{13}, 1 - p^i_{13})$ and $(2, 3; p^i_{23}, 1 - p^i_{23})$. A vote is given to 1 (more precisely, to the branch of the voting tree that contains alternative 1) if

$$p^i_{13}U^i_1 + (1 - p^i_{13})U^i_3 \geq p^i_{23}U^i_2 + (1 - p^i_{23})U^i_3 \tag{1}$$

With four alternatives the utilities for the winners in the second round need to be replaced with expected utilities in the third round. Dropping the superscripts denoting the individuals, we see that U_1 must be replaced with $p_{14}U_1 + (1 - p_{14})U_4$, U_2 with $p_{24}U_2 + (1 - p_{24})U_4$, and U_3 with $p_{34}U_3 + (1 - p_{34})U_4$. The condition for voting for the lower branch in the first round with four alternatives is thus

$$\begin{aligned} & p_{13}\left[p_{14}U_1 + (1 - p_{14})U_4\right] + (1 - p_{13})\left[p_{34}U_3 + (1 - p_{34})U_4\right] \\ \geq \; & p_{23}\left[p_{24}U_2 + (1 - p_{24})U_4\right] + (1 - p_{23})\left[p_{34}U_3 + (1 - p_{34})U_4\right]. \end{aligned} \tag{2}$$

Enumerating all formulas for expected utilities for different branches in a voting tree quickly becomes cumbersome as the number of alternatives increases. Furthermore, since the computer needs to find the winner in each pairwise comparison, and this requires computing expected utilities at each node of the voting tree, we need a general indexing method for denoting the nodes. This is done with an *ordering number*. The ordering numbers are used to denote the various probabilities and expected utilities in large voting trees. The technical details on the indexing are relegated to appendix A.

2.1 Virtual voter types

The aforementioned indexing system is designed for a single agenda. The indices themselves are not unique because they depend on the number of alternatives. However, studying path-dependence requires being able to formulate how voters act under various different agendas. Nevertheless, the indexing method can be used to study any agenda. The reason for this is based on the following observation. Suppose that we are interested in studying the behaviour of a given voter type t under some agenda A_f. Let A_1denote the agenda employed by the indexing method. Depending on the number of alternatives, it is (123), (1234), (12345), etc. There is a *virtual* *type* t' whose behaviour under agenda A_1 is *equivalent* to the behaviour of type t under agenda A_f. In order to study voters' behaviour under some agenda A_f, voters' utilities are re-ordered according to their virtual types, and their behaviour is then analysed in A_1. Thus, rather than changing the indexing method, voters types are changed and their voting is analysed as if they were voting under A_1.

In order to study voting under agenda (2341), for example, voters' utilities are converted such that they correspond to those they would have if alternative 1 were alternative 2, 2 were 3, 3 were 4, and 4 were 1. The following table shows voter type 14 ($4 \succ^i 1 \succ^i 3 \succ^i 2$) as an example.

A_1	A_f	type (14)	virtual type (7)
(1234)	(2341)		
1	2	4	1
2	3	1	2
3	4	3	4
4	1	2	3

Table 1: ordering numbers with four alternatives

Voters of type 14 under (1234) become voters of type 7 ($1 \succ^i 2 \succ^i 4 \succ^i 3$) under (2341). The ordering is constructed as follows. List A_1 and the ordering of interest A_f in two adjacent columns as in Table 1. Then take the type of interest (14) and list its ordering of alternatives in a third column. The ordering of the virtual type (7) is found as follows. Take the most preferred alternative (4), and search for this alternative in the first column. The number displayed in the second column on this row (row 4) provides the corresponding alternative (1) for the virtual type. Then take the second most preferred alternative (1), and repeat the procedure to find that the corresponding alternative is 2 for the virtual type. Go through all the alternatives in this way to find that the virtual type is 7. The expected utilities for a type 7 voter under agenda (1234) are thus the same as the expected utilities for a type 14 voter under agenda (2341) so that a type 7 voter behaves exactly the same under agenda (1234) as a type 14 voter (with identical utilities) under agenda (2341). A similar conversion is conducted for all voters. Then, once the outcomes of voting are calculated with these modified types *under agenda* (1234), the winning alternative is converted back to what it was with the original alternatives and voter types. For example, if the virtual alternative 3 wins when voter types are converted into virtual ones according to the agenda (2341), this means that alternative

2 wins under agenda (2341).

3 Signal extraction

Lehtinen (2007*b*) presented a model of signal extraction under amendment agendas. Since that model only featured three alternatives, voters did not need to take other voters' strategic behaviour into account. The only question of interest was which of two alternatives will win a pairwise contest in the last round of voting in which voters no longer have an incentive to vote strategically. With four or more alternatives, voters also need to take such behaviour into account. They are not assumed to have any knowledge on the behavioural propensities of other voters. However, since the model implies that intensively preferred alternatives are likely to obtain most strategic votes, if voters obtain perturbed information on aggregate-level differences in preference intensities, they can take other voters' strategic behaviour into account. Voters are thus not assumed to have any knowledge about individual preferences or behavioural propensities, but they can nevertheless take other voters' strategising into account if they obtain such perturbed aggregate-level information. Thus, a natural way of taking other voters' behaviour into account is to assume that voters obtain signals concerning the difference in the *sum of utility* between *each pair* of candidates. There are thus two kinds of signals: those that concern the last round, and those that concern the rounds from the second to the penultimate one. The former contain perturbed information on preference orderings whereas the latter contain perturbed information on aggregate preference intensities.

A *simulated game* g consists of a set of utilities created by a random number generator, beliefs based on these utilities, voters' perturbed signals, and voting outcomes under the different behavioural assumptions. All the variables are defined for a given simulated game g, but I will omit an index denoting the game in order to avoid unnecessary clutter.

3.1 Signals for the last round of voting

Given that the beliefs for the last round are the same as in Lehtinen (2007*b*), only a very brief account is given here. Let $\succ_i$ denote voter i's preference relation. Let N denote the number of voters, and let $n(j \succ k)$ denote the number of voters who prefer alternative j to alternative k in simulated game g. Voters are assumed to obtain a randomly perturbed signal on $n(j \succ k)$. Using a standardised variable $Q(j \succ k) = \frac{n(j \succ k) - Np}{\sqrt{Np^2}}$ allows constructing signals for which reasonable values of the perturbances are independent of the number of voters because$Q\left(j,k\right) \sim N(0,1)$ for any N. Since p is the probability of success in a Bernoulli trial, $p = \frac{1}{2}$, and a *signal* of voter i concerning the preferences for j and k can be written as

$$S_i(j,k) = \frac{2n(j \succ k) - N}{\sqrt{N}} + \varepsilon \cdot r_i\left(j,k\right), \tag{3}$$

where $r_i\left(j,k\right)$ is a realization of an i.i.d. standard normal random variable, and ε is a scaling factor that reflects the *reliability* of the signals. The smaller ε is, the more reliable a voter's signals are. In this paper, voters are assumed to know the reliability of their signals. Let $R_i\left(j,k\right) = \varepsilon \cdot r_i\left(j,k\right)$. The signal can then be written as follows:

$$S_i(j,k) = Q(j \succ k) + R_i\left(j,k\right). \tag{4}$$

Lehtinen (2006) shows that voters' beliefs can be derived from such signals. They are given by equation(5).

$$p_i(j,k) = 1 - \Phi\left(\frac{-s_i(j,k)}{\varepsilon\sqrt{1+\varepsilon^2}}\right), \tag{5}$$

3.2 Utility-based signals

Let $\Delta_i(j,k) = U_i\left(j\right) - U_i\left(k\right)$, and $\Delta(j,k) = \sum_{i=1}^{N} \Delta_i(j,k) = U\left(j\right) - U\left(k\right)$. A signal consists of the difference in the sum of utility $\Delta(j,k) = U\left(j\right) - U\left(k\right)$ and a random term $\varepsilon r_i(j,k)$. A signal for a contest before the last round is given by:

$$S_i(j,k) = U\left(j\right) - U\left(k\right) + \varepsilon r_i(j,k), \tag{6}$$

The standardized variable $Q\left(j,k\right) \sim N(0,1)$ is now given by:

$$Q\left(j,k\right) = \frac{\Delta\left(j,k\right) - N * E\left[\Delta\left(j,k\right)\right]}{\sigma_\Delta^i \sqrt{N}} = \frac{\Delta\left(j,k\right)}{\sigma_\Delta^i \sqrt{N}} = \frac{U\left(j\right) - U\left(k\right)}{\sigma_\Delta^i \sqrt{N}}, \tag{7}$$

where σ_Δ^i is the standard deviation of the variable $\Delta_i\left(j,k\right)$.[5] Calculating the probability that candidate j beats candidate k ($p_i(jBk)$), given a signal $S_i(j,k)$, requires knowing the variance of Δ_i. Appendix B shows that the standard deviation of Δ_i is $\sigma_\Delta^i = \sqrt{\frac{1}{6}}$, and the utility-based signal is thus given by

$$S_i^u(j,k) = \frac{U\left(j\right) - U\left(k\right)}{\sqrt{\frac{N}{6}}} + R_i\left(j,k\right). \tag{8}$$

The corresponding probabilities are derived by applying equation 5.

[5] It is obvious that $E\left[\Delta\left(j,k\right)\right] = 0$.

4 Simulation and setups

A setup is a set of assumptions used in a set of G=1000 simulated games. The number of voters was 201. Expected utility setups differ with respect to the reliability of voters' signals (ε), and the degree of correlation between voter types and preference intensities (C). In uniform setups voters' utilities are drawn from a uniform distribution on [0,1] The simulations were thus based on the impartial anonymous culture assumption: each voter type is equally likely (see Regenwetter et al., 2006). In setups with intensity correlation the preference orderings remain the same, but the utility of alternative 1 has been increased and the utilities of all other alternatives is decreased. In order to generate such setups without affecting the interpersonal comparisons or the preference orderings, the individual utilities were derived as follows.

$U_1, U_2, ..., U_n$ were first generated from the uniform distribution on [0,1] for each voter. U_1 and U_n were then used for defining the voter's utility scale as the $[U_1, U_n]$ interval. The utility for alternative 1 was then increased, and the utilities of all other alternatives decreased. Suppose, for example, that voter i had the following utilities: U_1=.55, U_2=.80, U_3=.05, and U_4=.52 (i.e. the voter has ranking 2143). These utilities define 'scales' that express the difference in utility between the alternatives. If K_{jk} denotes the scale between j and k, we have K_{21}=.25, K_{14}=.03, and K_{43}=.47. K_{21} now expresses how much the utility of 1 can be increased without changing the preference ordering. Let the starred variables denote their values after the conversion. The utility of 1 is now increased by setting $U_1^* = U_1 + (1-C)K_{21}$. Thus, if for example, C=.5, we get $U_1^* = .55 + (1-.5).25 = .675$. If 1 is already the most preferred alternative, we use the scale between it and the second-best alternative. The scales are then redrawn such that $K_{21}^* = .125$, K_{14}^*=.155, and K_{43}^*=.47. Then the utility of alternative 2 is decreased by setting $U_2^* = U_2 - (1-C)K_{21}^* = .8 - (1-.5).125 = .7375$ Then the utility of alternative 4 is decreased by setting $U_4^* = U_4 - (1-C)K_{43}^* = .52 - (1-.5).47 = .285$. U_3 is finally decreased using the scale K_{43} into $U_3^* = U_3 - (1-.5)(.47) = -.198$. A similar conversion of utilities is conducted for all voters.

5 Path-dependence

To the best of our knowledge, path-independence (Plott, 1973) has only been studied in complete information settings. This is why the existence of a Condorcet winner is often considered sufficient for path-independent social choices. Hammond (1977) shows that a social welfare functional satisfies a condition which is related to path-independence (metastatic consistency) if it satisfies Arrow's Independence of Irrelevant Alternatives. The result means that there is a close relation between ordinal choice and path-independence. The existence of a Condorcet winner does not guarantee path-independence. It is thus natural to investigate how important the existence of a Condorcet winner is for path-independence under incomplete information, but we will also study how the distribution of intensities affects path-dependence.

5.1 Two examples

Let us now use the model presented in the previous section to study an example of path-dependence. Assume that the preferences of three voters A, B, and C, for alternatives 1, 2 and 3 can be described with the following table:

A	B	C
2 (1)	2 (1)	1 (1)
1 (0.9)	1 (0.9)	3 (0.9)
3 (0)	3 (0)	1 (0)

Table 2: An example of path-dependence

The numbers in parentheses denote voters' utilities. 1 is the utilitarian winner and 2 the Condorcet winner. Assume first that the agenda is (123). If all voters engage in sincere voting behaviour, the Condorcet winner 2 will beat 1 in the first round and 3 in the second round, and emerges as the final outcome. Suppose now that voters maximised expected utility with incomplete information. Assume that all three voters have identical beliefs such that $p_{23} = 0.7$, and $p_{13} = 0.9$. The voters thus believe that it is likely that 2 beats 3, but even more likely that 1 beats 3 in the last round. These beliefs are fairly 'reasonable' because both 1 and 2 beat 3 if they survive the first round of voting. Furthermore, 1 beats 3 by three votes to zero, and 2 beats 3 by two votes to one. These beliefs are of course essentially plugged out of nowhere, but voters might well have such beliefs if they derived from the model of incomplete information described in the previous section.

Voters A and B vote sincerely for 2 in the first round if $U(1) < \frac{p_{23}}{p_{13}}$ (i.e. if $0.9 < \frac{0.7}{0.9} = 0.7778$). Since this is untrue, A and B will vote strategically for 1 in the first round of voting. Voter C has a weakly dominant strategy to vote for 1 in the first round of voting. 1 is thus the outcome because it beats 2 in the first round and 3 in the second. The utilitarian winner 1 is chosen if the voters maximise expected utility, but the Condorcet winner

2 is chosen if all voters vote sincerely. We may conclude that a Condorcet winner is not necessarily chosen under amendment agendas. Note, however, that the Condorcet winner is always selected in some agenda, because if it enters the voting in the last round, it beats any other alternative.

Our example also shows that social choice may be path-dependent even though a Condorcet winner exists: If the agenda is (231), voters A and B have a dominant strategy to vote sincerely for 2 in the first round. Whatever voter C does, 2 wins in the first round. In the second round 2 beats 3 and emerges as the winner under agenda (231).

Cyclic preferences are thus not a necessary condition for path-dependent social choices, but they are sufficient. Consider the payoffs displayed in example 2. Assume that $p_{12} = 0.55$, $p_{13} = 0.3$ and $p_{23} = 0.6$ for all voters.

A	B	C
1 (1)	2 (1)	3 (1)
2 (0.9)	3 (0.5)	1 (0.1)
3 (0)	1 (0)	2 (0)

Table 3: An example of path-dependence

You may verify that even though alternative 2 would be the outcome under agendas (123) and (132), it could not win under agenda (231), because it faces 1 in the second round, and 1 has a majority against it. Similar reasoning applies to any other agenda (and alternative). If the preferences are cyclic, there is always an agenda in which an alternative is not selected.

Example 2 also shows how strategic voting may lead to a lower degree of path-dependence than sincere voting. Note that EU behaviour resulted in two different outcomes, whereas sincere voting yielded a different outcome in all three agendas.

5.2 The degree of path-dependence

I will now present a criterion for the *extent* to which social choices depend on the order of voting. It will be called the *degree of path-dependence*. It is defined by applying a functional form that has been put to similar tasks in various different fields.[6]

[6] Corrado Gini was one of the first to apply it for studying the inequality of the distribution of income. Carnap (1952, pp. 65-68) discussed it under the name 'degree of order'. Carnap also used the terms 'homogeneity' and the 'uniformity of the world'). Patil and Taillie (1982) show various applications (e.g., biodiversity, industrial concentration) for the functional form used here and related forms. Political scientists are perhaps best

When the number of alternatives is n, there are $\frac{n!}{2}$ different binary agendas and thus $\frac{n!}{2}$ different voting orders. Each simulated election g is here taken to have a fixed preference profile, but the outcomes from this election may be different under different agendas. Let a_k^g denote the number of agendas in which alternative k is the outcome in the simulated election g. Let f_k^g denote the relative frequency of agendas in which alternative k is the outcome in a given simulated game g, i.e. the ratio between the number of agendas in which k is the outcome and the number of all possible agendas:

$$f_k^g = \frac{a_k^g}{\left(\frac{n!}{2}\right)}. \tag{9}$$

The *degree of path-dependence* DPD^g in a simulated election g is one minus the sum of squared relative frequencies for each of the alternatives

$$1 - \sum_{k=1}^{n} \left(f_k^g\right)^2$$

With n=3, for example, this is 1- $\left[(f_1^g)^2 + (f_2^g)^2 + (f_3^g)^2\right]$. The degree of path-dependence obtains its theoretical maximum when each alternative wins in an equal number of agendas: $1 - \frac{1}{n} = \frac{n-1}{n}$, and the theoretical minimum when one alternative wins under all agendas, i.e., when voting is path-independent. The theoretical minimum is always zero but the theoretical maximum depends on the number of alternatives. We will be more interested in the *average degree of path-dependence*:

$$DPD = 1 - \frac{1}{G} \sum_{g=1}^{G} \sum_{k=1}^{n} \left(f_k^g\right)^2. \tag{10}$$

Let us now show that our criterion yields reasonable results. Assume that in simulated elections g_1 and g_1 the relative frequencies of five alternatives are distributed in the following manner.

	election 1	election 2
number of agendas under which alternative 1 wins	57	20
number of agendas under which alternative 2 wins	1	20
number of agendas under which alternative 3 wins	1	20
number of agendas under which alternative 4 wins	1	0
number of agendas under which alternative 5 wins	0	0

acquainted with the so called 'effective number of parties' Laakso and Taagepera (1979).

Thus $a_1^{g1} = 57$, $a_2^{g1} = 1$, $a_3^{g1} = 1$, $a_4^{g1} = 1$, and $a_5^{g1} = 0$, and $a_1^{g2} = 20$, $a_2^{g2} = 20$, $a_3^{g2} = 20$, $a_4^{g2} = 0$, and $a_5^{g2} = 0$. In election 1 we have $\sum_{k=1}^{n} \left(f_k^{g1}\right)^2 = \left(\frac{57}{60}\right)^2 + \left(\frac{1}{60}\right)^2 + \left(\frac{1}{60}\right)^2 + \left(\frac{1}{60}\right)^2 = 0.9033$ so that the degree of path-dependence is $1 - 0.9033 = 0.0967$. In game 2 we have $\sum_{k=1}^{n} \left(f_k^{g2}\right)^2 = \left(\frac{20}{60}\right)^2 + \left(\frac{20}{60}\right)^2 + \left(\frac{20}{60}\right)^2 = \frac{1}{3}$ so that the degree of path-dependence in this game is $1 - \frac{1}{3} = 0.6667$. Clearly, our criterion of the degree of path-dependence yields intuitively reasonable results.

6 Simulation results

6.1 Utilitarian efficiencies

Figures 2, 3 and 4 display utilitarian efficiencies with four, five and six alternatives. It seems clear that strategic voting continues to increase utilitarian efficiencies also when there are more than three alternatives. Furthermore, under EU behaviour the utilitarian efficiencies only have a slight tendency to decrease as the number of alternatives increases. Indeed, when n=6, an efficiency of 50 per cent is quite remarkable under the impartial anonymous culture. Note, finally, that the utilitarian efficiencies were here calculated as averages over all agendas. In other words, these results also mean that Lehtinen's (2007b) result that strategic voting increases utilitarian efficiency under amendment agendas does not depend on using a particular (123) agenda.

6.2 Degree of path-dependence

Figures 5, 6 and 7 show degrees of path-dependence with 4, 5 and 6 alternatives. These figures show that the degree of path-dependence is higher under EU than under SV behaviour. The main reason for this result is that there is often a Condorcet winner even when the number alternatives is relatively large. With 6 alternatives, for example, there is a Condorcet winner among the alternatives in 69.9 per cent of the simulated elections. Under those elections, DPD_{SV} is always zero but DPD_{EU} is usually far from zero. However, when there is no Condorcet winner, sincere voting yields a very high degree of path-dependence, but strategic voting decreases the degree of path-dependence because the utilitarian winner obtains many more strategic votes than other alternatives. This can be seen from figures 8 and 9. Thet display degrees of path-dependence under those elections in which there is no Condorcet winner.

Figure 2: Utilitarian efficiencies with n=4

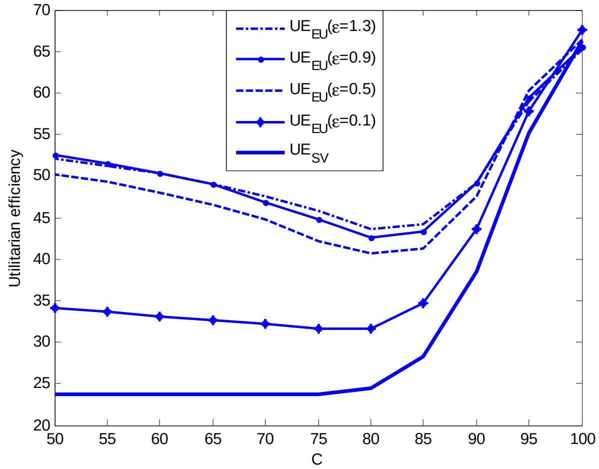


Figure 3: Utilitarian efficiencies with n=5

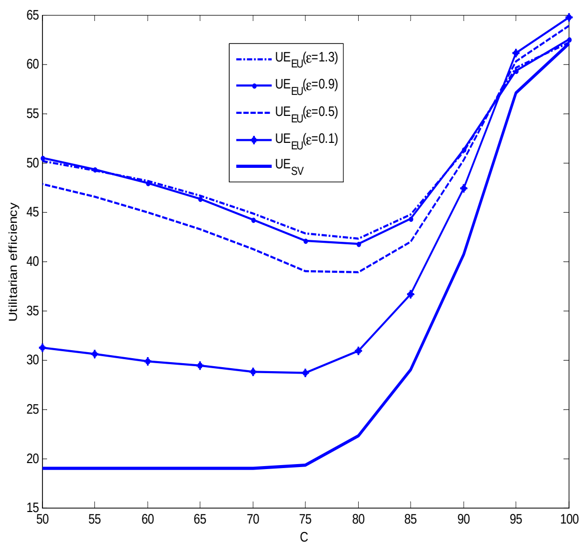


Figure 4: Utilitarian efficiencies with n=6

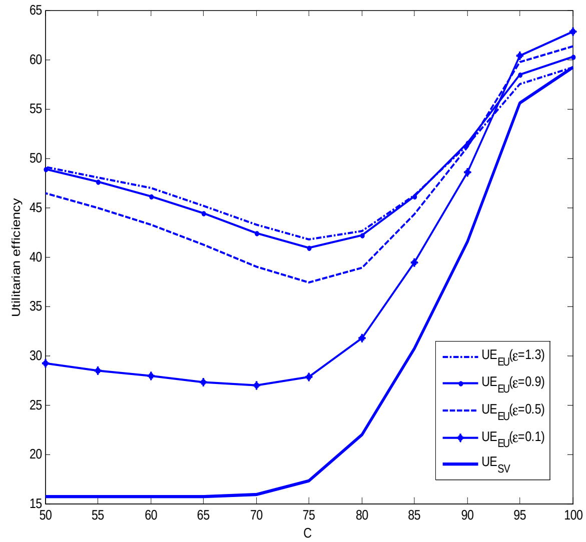


Figure 5: Degree of path-dependence with n=4

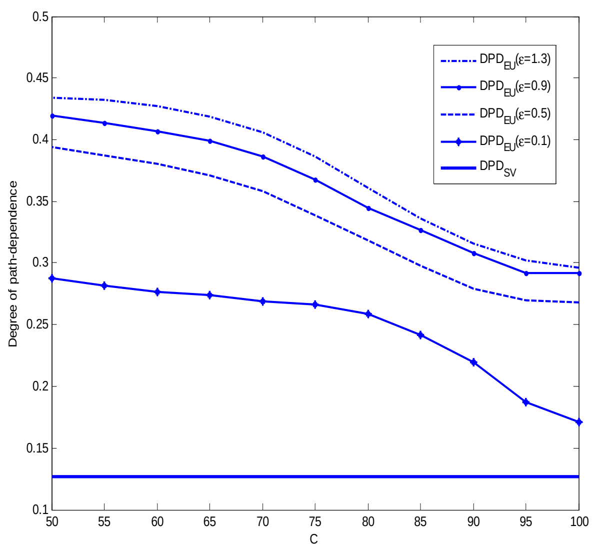


Figure 6: Degree of path-dependence with n=5

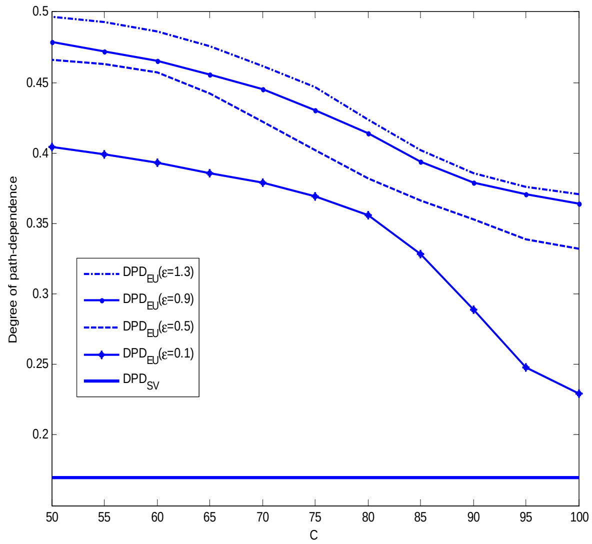


Figure 7: Degree of path-dependence with n=6

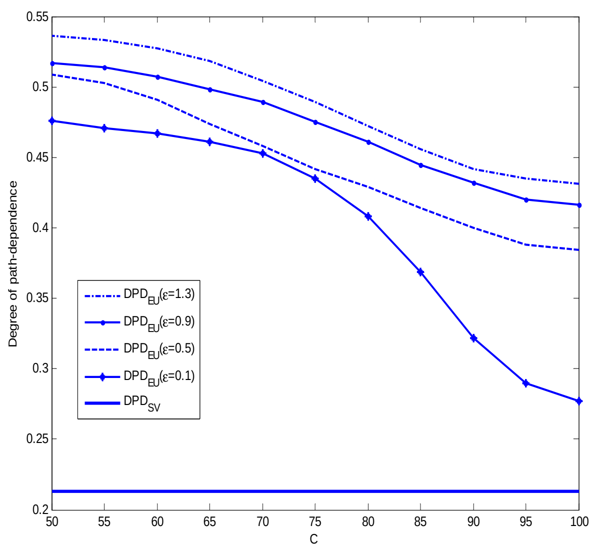


Figure 8: DPD when there is no Condorcet winner

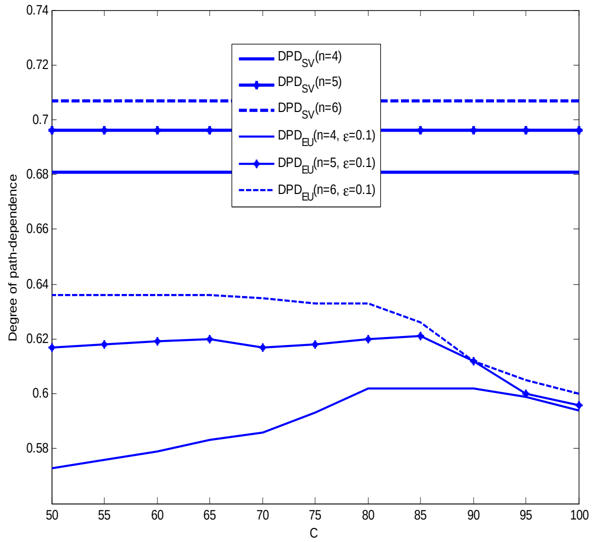
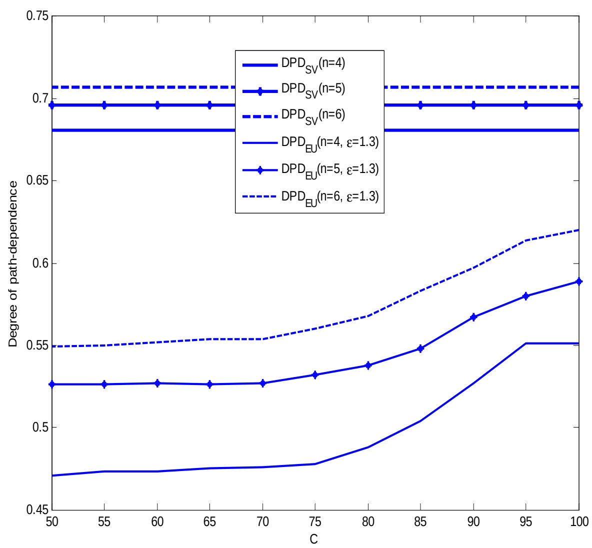


Figure 9: DPD when there is no Condorcet winner



Conclusions

Strategic voting increases the degree of path-dependence mainly because it may generate path-dependent choices even when there is a Condorcet winner. The simulation setups were constructed in such a way that the Condorcet winner and the utilitarian winner are often different. Given that strategic voting increases the number of votes for the latter, it increases utilitarian efficiency. However, under amendment agendas, the Condorcet winner always wins under at least those agendas under which it is introduced on the last round of voting. This creates a situation in which utilitarian winners win under a large number of agendas, but never all of them if they are not the same alternative as the Condorcet winners. This, in turn, leads to relatively high degrees of path-dependence. When there is no Condorcet winner strategic voting decreases the degree of path-dependence because it concentrates votes on the utilitarian winner.

A Appendix: Indexing the voting trees

A.1 Ordering numbers

Consider all the pairs of alternatives for four alternatives as shown in Table 4:

ordering number	first alternative	second alternative
1	1	2
2	1	3
3	1	4
4	2	3
5	2	4
6	3	4

Table 4: ordering numbers with four alternatives

The first pair of alternatives is always $\{1, 2\}$. Adding a third alternative yields the pairs $\{1, 3\}$ and $\{2, 3\}$ because the n:th alternative is put against each alternative in the last round. By the same principle, adding a fourth alternative yields $\{1, 4\}$, $\{3, 4\}$, $\{2, 4\}$, $\{3, 4\}$, and so on. These sequences of pairs of alternatives provide us with a corresponding sequence of ordering numbers.

An ordering number is used to identify the nodes and the corresponding probabilities in the voting trees.[7] The ordering number determines the probability p_{jk} used in a given branch, and the corresponding label for the expected utility expression EU_{jk}[8]. A node always corresponds to a comparison of a pair of alternatives. With n alternatives there are $\binom{n}{2}$ different pairwise comparisons. Given any two alternatives j and k such that $k > j$, and a number of alternatives n, an ordering number $o(n, j, k)$ is given by the following formula:

$$o(n, j, k) = (n - \frac{j}{2})(j - 1) + k - j \tag{11}$$

The formula can be derived as follows. The pairwise comparisons are ordered such that all pairs in which alternative 1 is involved obtain the first ordering numbers in an ascending order, then those in which alternative 2 is involved but not alternative 1, then those in which 3 is involved but not alternatives 1 and 2, and so on. Alternative 1 is involved in $n - 1$ pairs in which it is the first alternative, alternative 2 in $n - 2$ such pairs and so on. There are 0 pairs in which alternative 1 is involved before alternative 1, $(n-1)$ pairs before alternative 2 is the first alternative, $(n-1)+(n-2) = 2n-3$ pairs before alternative 3 is the first alternative, $(n-1)+(n-2)+(n-3) = 3n-6$ comparisons before alternative 4 is the first and so on. The number of pairs before alternative j can thus be written as a sum of $(j-1)n$, and the sum of an arithmetic series $0, -1, -2, ..., -(j-1) = -\sum_{i=0}^{j-1} i$. The sum of this series is

[7]We do not use the so called *indexed traversal order* here. This is the ordering imposed on the nodes of a game tree by a lexicographic ordering of the nodes when each node is identified by the sequence of branch numbers necessary to reach it. If the number of alternatives is large, there are a large number of branches in which the alternatives n and n-1, or n and n-2 etc. are put against each other. There are thus usually several different paths to a given pair of alternatives. The indexed traversal order is not used because expected utilities would have to be calculated several times for identical sub-branches in a voting tree. The benefit in CPU-time and memory from using an ordering number instead of the indexed traversal order is considerable in computer simulations. Using an ordering number implies the assumption that irrespective of the path with which a branch is obtained, any branch with identical alternatives is to be evaluated in the same way. It implies the assumption that there is no learning because if the agents could learn in a significant manner from previous voting rounds, it would not be justifiable to assume that the agents always evaluate the expected utility of a branch only on the basis of the alternatives in that branch.

[8]Each sub-branch in a voting tree does not have a unique ordering number, because there are several identical subtrees in any binary agenda with at least four alternatives. The number of repetitions increases with the number of alternatives so that with six alternatives, for example, the ordering numbers for the last round are 5, 15, 14, 15, 12, 15, 14, 15, 9, 15, 14, 15, 12, 15, 14, 15, and 4, 13, 11, 13, 8, 13, 11, 13, for the penultimate round etc.

$\frac{(j)[0-(j-1)]}{2} = \frac{-j(j-1)}{2}$ so that the number of pairs before j is $(j-1)n - \frac{j(j-1)}{2} = (n - \frac{j}{2})(j-1)$. $k-j$ expresses the number of pairs between alternative j and alternatives from $j+1$ to k. For example, the comparison $\{1,4\}$ corresponds to the ordering number $o_{n14} = (1-1)*4 - \sum_{i=0}^{1-1} i + 4 - 1 = 3$, and $\{3,4\}$ corresponds to $o_{n34} = 6$. The sequence $\{1,4\}$, $\{3,4\}$, $\{2,4\}$, $\{3,4\}$ corresponds thus to (3,6,5,6). Such sequences and pairs of alternatives allow us two write the expected utility expressions for any number of alternatives.

As in Lehtinen (2007), voters are assumed to obtain perturbed signals $S_i(j,k)$ concerning the number of voters who prefer one alternative to another in a pairwise comparison. The signals and the probabilities $p_i(j,k)$ in a vote between any two branches, as well as the corresponding expected utilities $EU_i(j,k)$, are labeled with the ordering numbers. p_1 thus corresponds to p_{12}, p_2 corresponds to p_{13},..., and $p_{[\binom{n}{2}-1]\binom{n}{2}}$ corresponds to $p_{\binom{n}{2}}$. Signals and expected utilities are similarly labeled. Each player i thus obtains a set of perturbed signals $\{s_1, s_2, ..., s_{\binom{n}{2}}\}$.[9]

A.2 Virtual ordering numbers

In order to use the ordering numbers for studying any agenda, it is necessary to find the virtual ordering numbers. The condition for voting for the lower branch under agenda 1234 was.

$$p_{13}\left[p_{14}U_1 + (1-p_{14})U_4\right] + (1-p_{13})\left[p_{34}U_3 + (1-p_{34})U_4\right] \quad (12)$$
$$\geq \quad p_{23}\left[p_{24}U_2 + (1-p_{24})U_4\right] + (1-p_{23})\left[p_{34}U_3 + (1-p_{34})U_4\right]. \quad (13)$$

If this is expressed in terms of ordering numbers, we get

$$p_2\left[p_3U_1 + (1-p_3)U_4\right] + (1-p_2)\left[p_6U_3 + (1-p_6)U_4\right]$$
$$\geq \quad p_4\left[p_5U_2 + (1-p_5)U_4\right] + (1-p_4)\left[p_6U_3 + (1-p_6)U_4\right]. \quad (14)$$

[9]Another possibility would be to assume that each player obtains one observation for each individual and each possible pair of alternatives. With n alternatives there are $\binom{n}{2}$ different pairwise comparisons. Each player would then obtain $\binom{n}{2}*(N-1)$ observations in each simulated game g. If a setup has M elections, there would then be $M*N*\binom{n}{2}*(N-1)$ observations in one setup. With $M = 1000$, $N = 201$, and $n = 7$, this is 844200000 observations in each setup. We have not used this latter approach because computing the beliefs on the basis of them requires a considerable amount of memory and CPU-time. The most important reason for using ordering numbers is that they save computing time. Most of the time is spent in calculating the probabilities from the normal distribution. This is why it is done only once.

Consider now agenda 2341 and eq. (12). Replacing alternative 2 with 1, 3 with 2, 4 with 3 and 1 with 4, we get

$$\begin{aligned} & p_{24}\left[p_{21}U_2+(1-p_{21})U_1\right]+(1-p_{24})\left[p_{41}U_4+(1-p_{41})U_1\right] \\ \geq\ & p_{34}\left[p_{31}U_3+(1-p_{31})U_1\right]+(1-p_{34})\left[p_{41}U_4+(1-p_{41})U_1\right]. \end{aligned} \quad (15)$$

if utilities are expressed in terms of the virtual ordering numbers. The computer is able to transform the labels for probabilities into virtual ordering numbers by using the following technique. Consider Table 5. The six lowest rows express ordering numbers for alternative pairs when the ordering of the items in the pair is interchanged. The virtual ordering numbers are found by listing the virtual alternatives in the pairwise comparisons as in columns '1. virtual' and '2. virtual', and by finding which ordering number (from the first column) correspoinds to them. Thus, for example, if 1. virtual is 4, and 2. virtual is 1, the corresponding virtual ordering number is 9. Replacing the labels of all pairs of alternatives in (14) with virtual ordering numbers then yields

$$\begin{aligned} & p_5\left[p_7U_2+(1-p_7)U_1\right]+(1-p_5)\left[p_9U_4+(1-p_9)U_1\right] \\ \geq\ & p_6\left[p_8U_3+(1-p_8)U_1\right]+(1-p_6)\left[p_9U_4+(1-p_9)U_1\right]. \end{aligned} \quad (16)$$

We can also write this in terms of expected utilities:

$$\begin{aligned} & p_5EU_7+(1-p_5)EU_9 \\ \geq\ & p_6EU_8+(1-p_6)EU_9. \end{aligned} \quad (17)$$

If there are five alternatives, the expression for voting for the lower branch can be obtained by replacing the utilities U_j with expected utilities from contests between j and 5. Thus, it is helpful to calculate the expected utilities for all last-round contests. These expected utilities are then put in place of utilities in equation (12), using appropriate ordering numbers.

We will also need to know how the ordering number for the *next* round of voting is identified for calculating how the voters will vote after a result from a previous round has been revealed. This can be done using table 6 above. In order to construct this table, take Table 5, pick the ordering numbers from the rows in which the alternative that is introduced in the last round of voting (i.e. $n = 4$) is found on the 'second' column, and arrange the ordering numbers in a single column to obtain a column $\begin{bmatrix} 3 \\ 5 \\ 6 \end{bmatrix}$. Then pick the rows

ordering number	first	second	1. virtual	2. virtual	virtual number
1	1	2	2	3	4
2	1	3	2	4	5
3	1	4	2	1	7
4	2	3	3	4	6
5	2	4	3	1	8
6	3	4	4	1	9
7	2	1			
8	3	1			
9	4	1			
10	3	2			
11	4	2			
12	4	3			

Table 5: virtual ordering numbers for agenda 2341 with four alternatives

1. round	**2. round**	**3. round**
		6
	4	5
1	2	3

Table 6: ordering numbers

with $n-1=3$ from this column to get $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$. Continuing like this until alternative 2, putting the results in consecutive columns, and flipping them up and down yields Table 6.

The columns were flipped up and down because the ordering numbers are now positioned in the same way as the branches in the voting tree. The leftmost column in this table displays the ordering number in the first round of voting, the second column the ordering numbers in the second, and the $(n-1)^{th}$ column from the left the $(n-1)^{th}$ (i.e. the last) round of voting. If the lower branch obtains a majority of votes, the ordering number for the next round is found on the next column directly to the right of the present one, and if the upper branch obtains more votes, the ordering number is found on the next column to the right on the south-west - north-east diagonal.

For example, with four alternatives, if 1 beats 2 in the first round, we know that we will need to calculate the expected utilities for the two branches emanating from the branch corresponding to ordering number 2, because 2 lies directly to the right of 1, the ordering number for the vote between alternatives 1 and 2. The ordering number for one branch that emanates from branch 2 must be directly to the right from 2 (3), and the other must be on the diagonal (6).

Using these indexing methods, we may thus express the expected utilities for all branches in a voting tree, and the corresponding probabilities for any number of alternatives in any voting round under an amendment agenda.

B Appendix: The standard deviation of Δ_i

The sum of utilities for candidate j can be viewed as the sum of N random variables U_i, one for each voter: $U_1+U_2+,...,U_i,...,+U_N = \sum_{i=1}^{N} U_i = U(j)$. Let $\Delta_i(j,k) = U_i(j) - U_i(k)$, the variance of Δ_i is

$$Var(\Delta_i) = E\left[U_i(j) - U_i(k) - E\left[U_i(j) - U_i(k)\right]\right]^2.$$

Each U_i is a uniformly distributed random variable on (0,1) with expected value $\frac{1}{2}$, and variance $\int_0^1 \left(v_i - \frac{1}{2}\right)^2 dV_i = \frac{1}{12}$. It is obvious that $E\left[U_i(j) - U_i(k)\right] = 0$. The variance of Δ_i is thus given by

$$Var(\Delta_i) = E\left[U_i(j)^2\right] - 2E\left[U_i(j)\,U_i(k)\right] + E\left[U_i(k)^2\right].$$

Since by definition

$$E(U_i(j)^2) = Var(U_i(j)) + \left[E(U_i(j))\right]^2 = \frac{1}{12} + \left(\frac{1}{2}\right)^2 = \frac{1}{3},$$

$U_i(j)$ and $U_i(k)$ are independent random variables so that $E\left[U_i(j)\,U_i(k)\right] = E\left[U_i(j)\right]E[U_i(k)] = \frac{1}{2}\frac{1}{2} = \frac{1}{4}$. The variance of Δ_i is thus $\frac{1}{3} - 2\left(\frac{1}{4}\right) + \frac{1}{3} = \frac{1}{6}$. Since Δ is the sum of N independent random variables $\Delta = \Sigma_{i=1}^{N}\Delta_i$, the standard deviation of Δ, σ_Δ, is$\sqrt{\frac{N}{6}}$.

References

Carnap, Rudolf. 1952. *The continuum of inductive methods*. Chicago: University of Chicago Press.

Enelow, James. 1981. "Saving amendments, killer amendments, and an expected utility theory of sophisticated voting." *The Journal of Politics* 43(4):1062–1089.

Farquharson, Robin. 1969. *Theory of voting.* UK: Hodge.

Hammond, Peter J. 1977. "Dynamic restrictions on metastatic choice." *Economica* 44(176):337–350.

Laakso, Markku and Rein Taagepera. 1979. "Effective number of parties: A measure with application to West Europe." *Comparative Political Studies* 12:3–27.

Lehtinen, Aki. 2006. "Signal extraction for simulated games with a large number of players." *Computational Statistics and Data Analysis* 50:2495–2507.

Lehtinen, Aki. 2007*a*. "The Borda rule is also intended for dishonest men." *Public Choice* 133(1-2):73–90.

Lehtinen, Aki. 2007*b*. "The welfare consequences of strategic voting in two commonly used parliamentary agendas." *Theory and Decision* 63(1):1–40.

Lehtinen, Aki. 2008. "The welfare consequences of strategic behaviour under approval and plurality voting." *European Journal of Political Economy* 24(3):688–704.

List, Christian. 2004. "A Model of Path-Dependence in Decisions over Multiple Propositions." *American Political Science Review* 98(3):495–514.

McKelvey, Richard and Richard Niemi. 1978. "A multistage game representation of sophisticated voting for binary procedures." *Journal of Economic Theory* 18:1–22.

Ordeshook, Peter. 1986. *Game theory and political theory: An introduction.* Cambridge: Cambridge University Press.

Ordeshook, Peter and Thomas Schwartz. 1987. "Agendas and the control of political outcomes." *American Political Science Review* 81:179–200.

Patil, G.P. and C. Taillie. 1982. "Diversity as a concept and its measurement." *Journal of the American Statistical Association* 77(379):548–561.

Plott, Charles. 1973. "Path independence, rationality, and social choice." *Econometrica* 41(6):1075–1091.

Regenwetter, Michel, Bernard Grofman, A. A. Marley and Ilia Tsetlin. 2006. *Behavioral social choice: probabilistic models, statistical inference, and applications.* Cambridge: Cambridge University Press.