Strategic approval voting in a large electorate

Jean-François LASLIER[*]
Laboratoire d'Économétrie, École Polytechnique
1 rue Descartes, 75005 Paris
laslier@shs.polytechnique.fr

June 16, 2006

Abstract

The paper considers approval voting for a large population of voters. It is proven that, based on statistical information about candidate scores, rational voters vote sincerly and according to a simple behavioral rule. It is also proven that if a Condorcet-winner exists, this candidate is elected.

1 Introduction

Approval Voting (AV) is the method of election according to which a voter can vote for as many candidates as she wishes, the elected candidate being the one who receives the most votes. In this paper two results are established about AV in the case of a large electorate when voters behave strategically: the sincerity of individual behavior (rational voters choose sincere ballots) and the Condorcet-consistency of the choice function defined by approval voting (whenever a Condorcet winner exists, it is the outcome of the vote).

Under AV, a ballot is a subset of the set candidates. A ballot is said to be sincere, for a voter, if it shows no "hole" with respect to the voter's preference ranking; if the voter sincerely approves of a candidate x she also approves of any candidate she prefers to x. Therefore, under AV, a voter has several sincere ballots at her disposal: she can vote for her most preferred candidate, or for her two, or three, or more most preferred candidates.[1]

It has been found by Brams and Fishburn (1983) that a voter should always vote for her most-preferred candidate and never vote for her least-preferred one. Notice that this observation implies that strategic voting is sincere in the case of three candidates. The debate about strategic voting under AV was made vivid by a paper by Niemi (1984). Niemi argued that, because there is more than one sincere approval ballot, the rule "almost begs the voter to think and behave strategically, driving the voter away from *honest* behavior" (Niemi's emphasis, p. 953). Niemi then gave some examples showing that an approval game cannot be solved in dominant strategies. Brams and Fishburn (1985), responded to this view, but the debate was limited by the very few results available about equilibria of voting games in general and strategic approval voting in particular.

For instance in one chapter of their book, Brams and Fishburn discuss the importance of pre-election polls. They give an example to prove that, under AV, adjustment caused by continual polling can have various effect and lead to cycling even when a Condorcet winner exists (example 7, p.120). But with no defined notion of rational behavior they have to postulate specific (and changing) adjustment behavior from the voters.

With the postulate that voters use sincere and undominated ("admissible") strategies but can use any of these, Brams and Sanver (2003) describe the set of possible winners of an AV election. They conclude that a plethora

---

[1] While some scholars see this feature as a drawback (Saari and Van Newenhizen, 1988), observation shows that people appreciate to have this degree of freedom: see Laslier and Van der Straeten (2006) for an experiment, and the survey Brams and Fishburn (2005) on the practice of AV.

of candidates pass this test.

This literature assumes that voters are only interested in the elected candidate and not on how many votes are obtained by the other candidates (no expressive voting) and the present paper will not depart from this standard assumption. But none of the above approaches uses an equilibrium theory of approval voting. The rationality hypothesis in a voting situation without expressive voting may have odd implications. For instance it implies that a voter is indifferent between all her strategies as soon as her votes cannot change the result of the election. As a consequence, in an election held under plurality or AV rule, any situation in which one candidate is slightly (three votes) ahead of the others is trivially a Nash equilibrium. It is fair to say that with many voters, in most voting situations, the Nash equilibrium concept is a powerless tool.

Taking a standard game-theoretical point of view, De Sinopoli, Dutta and Laslier (2006) have studied in details examples of AV games. They showed that even powerful refinement considerations such as strategic stability or regularity could not guarantee the election of a Condorcet winner when it exists nor exclude insincere behavior at equilibrium.

All these results are essentially negative ones, in the sense that this literature makes almost no prediction for AV games. A breakthrough in the rational theory of voting occurred when it was realized that considering large numbers of voters was technically possible and offered a more realistic account of political elections. This approach was pioneered by Myerson and Weber (1993). In the same paper, AV and other rules are studied on an example with three types of voters and three candidates (a Condorcet cycle). Subsequent papers by Myerson improved the techniques and tackled several problems in the theory of voting (Myerson 1998, 2000, 2002). For instance, using the so-called Poisson-Myerson model of voter participation, Myerson (2002) obtained that approval voting guarantees the choice of the majoritarian outcome in the case where there are only two types of voters. The present paper applies similar, but not identical, techniques to approval voting with no restriction on the number of candidates or voter types.

The abstract theory of the refinement of Nash equilibrium considers small perturbations in the choice of strategies or in the payoffs ("trembling hand" and other perfection criteria, strategic stability). The Poisson-Myerson model introduces uncertainty at the level of the number of players. In the present paper, uncertainty is introduced in an other way. We suppose that there is some small but strictly positive probability that each vote is wrongly recorded. This probability is supposed to be independent of the identity of the voter, of the candidates, and of the other voters' vote. It

should be interpreted as the probability of a technical mistake which is unavoidable when recording each individual vote. Referring to the controversial election of the US president in November 2000, the term "Florida tremble" was suggested to describe this hypothesis. This term rightly conveys the idea that material mistakes exist in elections. Notice however that in our model, with independence and a large number of voters, the mistakes in fact introduce no bias in the election.

Rationality implies that a voter can decide of her vote by limiting her conjecture to those events in which her vote is pivotal. In a large electorate, this is a very rare event, and it may seem unrealistic that actual voters deduce their choices from implausible premises. One can wish that a positive theory of the voter be more behavioral and less rational. This is a wise remark in general, but in the case of approval voting, the rational response turns out to be very simple. It can be described as follows. Let x_1 be the candidate who the voter thinks is the most likely to win. The voter will approve of any candidate she prefers to x_1. She will never approve of a candidate she prefers x_1 to. To decide wether she will approve of x_1 or not, she compares x_1 to the second most likely winner (the "most serious contender"). This behavior recommends a sincere ballot and its implementation does not require sophisticated computations, it only requires that the voters holds a conjecture about which two candidates are to receive the most and second-most votes. For reasons that will be transparent in the sequel, we term this behavior the "lexicographic response." Moreover, we will show that the lexicographic response is, in this case, identical to a well known "boundedly rational" but psychologically sound behavior: the choice by sequetial elimination theory of Tversky (1997a). The rational and the behavioral models are here equivalent.[2]

The paper is organized as follows. After this introduction, Section 2 describes all the essential features of the model. Section 3 informally describes the main point in the argument. Section 4 contains the formal statements and proofs of the results: the description of rational voting (Theorem 1) from which we deduce sincerity (Corollary 1), and the description of equilibrium approval scores (Theorem 2) from which we deduce Condorcet consistency (Corollary 2). Some computations are provided in an Appendix.

---

[2] Coming back to Florida, the final certified margin of Bush against Gore was 537 votes and it was often heard after the election that the 97,421 voters who voted for Ralph Nader acted irrationally precisely because they failed to reason on the basis of the pivotal event "Gore almost tie with Bush".

2   The model

Candidates and voters

Let X denote the finite set of candidates. We follow Myerson and Weber (1993) in considering that there exists a finite number of voter types $\tau \in T$. A voter of type τ evaluates the utility of the election of candidate $x \in X$ according to a von Neumann and Morgenstern utility index $u_\tau(x)$. A preference on X is a transitive and complete binary relation. In this paper, for simplicity, all preferences are supposed to be strict: no voter is indifferent between two candidates. Preferences are denoted in the usual way: $x\, P_\tau\, y$ means that τ-voters prefer candidate x to candidate y, thus:

$$x\, P_\tau\, y \iff u_\tau(x) > u_\tau(y).$$

Notice that one has to assume utility functions besides preference relations, because voters take decisions under uncertainty. But it turns out that the obtained results can all be phrased in terms of preferences. That is: It will be proven that rational behavior in the considered situation only depends on preferences.

We wish to consider a large electorate. To do so, we consider a fixed finite number n of voters and replicate this set n times in the following way. Let p_τ denotes the fraction of type-τ voters, with:

$$\sum_{\tau \in T} p_\tau = 1.$$

In the n-fold replicate economy the number of type-τ voters is nvp_τ and the total number of voters is nv.

Voting

An approval voting ballot is just a subset of the set of candidates. The relative approval **score** $s(x)$ of candidate x is the fraction of the electorate that approves of x. If all type-τ voters choose ballot $B_\tau \subseteq X$, the score is the sum of p_τ over types τ such that $x \in B_\tau$:

$$s(x) = \sum_{\tau:\, x \in B_\tau} p_\tau.$$

Here the number of votes in favor of x is $nvs(x)$. Of course $0 \leq s(x) \leq 1$, but $\sum_{x \in X} s(x)$ is usually larger than 1.

Trembling Ballots

We consider the following perturbations. For each voter and each candidate there is a probability $\varepsilon > 0$ that this vote is not recorded. This probability is supposed to be small. Precisely, it will be sufficient to suppose that $\varepsilon < 1/v$ (independently of n). We also suppose that these mistakes occurs independently of the voter, of the candidate, and of the voter approving or not other candidates[3]. For candidate x and voter i, let $\eta_{i,x}$ be equal to 1 with probability ε and to 0 with the complement probability, when the (intended) number of votes for candidate x is $nvs(x)$, the realized number of votes is a random variable $\mathcal{S}^n(x)$. Denote by $AV(i)$ the set of the $nvs(x)$ voters who approve of x, then:

$$\mathcal{S}^n(x) = \sum_{i \in AV(x)} (1 - \eta_{i,x}).$$

The random variable $\mathcal{S}^n(x)$ is binomial, with expected value and variance:

$$
\begin{aligned}
\mathrm{E}[\mathcal{S}^n(x)] &= (1-\varepsilon)nvs(x) \\
\mathrm{V}[\mathcal{S}^n(x)] &= \varepsilon(1-\varepsilon)nvs(x).
\end{aligned}
$$

To get intuition about this model, denote by $\mathbf{b}(x)$ the realized score:

$$\mathbf{b}(x) = \frac{\mathcal{S}^n(x)}{nv}.$$

For n large, the central limit theorem implies that the random variable $\mathbf{b}(x)$ is approximately normal. One can write:

$$
\begin{aligned}
\mathbf{b}(x) &\rightsquigarrow {}_{n\to\infty}\mathcal{N}\left(a(x),\ \mathrm{var}\right), \\
a(x) &= (1-\varepsilon)s(x), \\
\mathrm{var} &= \frac{\varepsilon(1-\varepsilon)s(x)}{nv}.
\end{aligned}
$$

The expectation of $\mathbf{b}(x)$, denoted $a(x)$, is just a linear transformation of the score $s(x)$, and a and s rank candidates the same way. The variance of $\mathbf{b}(x)$ is decreasing as $1/n$. It is therefore tempting to directly use this continuous model and to define as "approval scores" independent Gaussian variables $\mathcal{N}\left(a(x),\ \frac{\varepsilon(1-\varepsilon)s(x)}{nv}\right)$. This continuous model delivers some correct qualitative intuition, but using it without care may also lead to mistaken conclusion, as is shown in Appendix B.

---

[3]The model here differs from Myerson-Poisson models, in which the population of voters is uncertain. If uncertainty about candidate scores were to arise from uncertainty about the population of voters, then errors would be correlated.

3   The argument

In this section the main argument for the proof is informally stated.
More or less serious races

The elected candidate is one with highest score. Ties are resolved by a
fair lottery. Given a strategy profile and its score vector s, the most probable
event is that the candidate with highest score wins, but it may be the case
that mistakes are such that another candidate does, and it may be the case
that two (or more) candidates are so close that one vote can be decisive. In
what follows, it will be needed to evaluate the probabilities of some of these
events when the number of voters is large.

One ballot may have consequences on the result of the election only if the
two (or more) first ranked candidates have scores that are within one vote
of each other. The probability of such a pivotal event is small if n is large,
but some of these events are even much less probable than others. It will
be proved that three (or more)-way ties are negligible in front of two-way
ties, and that different two-way ties are negligible one in front of the other
according to a simple lexicographic pattern: The most probable one is a tie
between candidates x_1 and x_2: that is the most "serious race", the second
most serious race is x_1 against x_3, which is less serious than x_1 against x_4,
etc. Lemma 2 states this point precisely.

This observation turns out to be sufficient to infer rational behavior. A
rational voter will obey a simple heuristic and consider in a sequential way
the different occurrences of her being pivotal, according to the magnitude
of these events.

Rational behavior and the Law of Lexicographic Maximization:

We will show that the rational behavior can be deduced from a simple
heuristic, which can be described as follows in very general terms. Let D
be a finite set of possible decisions and $\Omega = \{\omega_1, ..., \omega_N\}$ a finite partition
of events, with π the probability measure on Ω and u a von Neumann-
Morgenstern utility. The maximization problem of a rational decision-maker
can be written

$$\max_{d \in D} \sum_{k=1}^{N} \pi(\omega_k) \, \mathrm{E}u(d, \omega_k),$$

where $\mathrm{E}u(d, \omega_k)$ denotes the expected utility of decision d conditional on
event ω_k. Suppose that π is such that, for $1 \leq k < k' \leq N$, the probability
$\pi(\omega_k)$ is large compared to the probability $\pi(\omega_{k'})$:

$$\pi(\omega_1) \gg \pi(\omega_2) \gg ... \gg \pi(\omega_N).$$

Then the above maximization problem is solved recursively by the following algorithm:

- $D_0 = D$.

- For $k = 1$ to N, $D_k = \arg\max_{d \in D_{k-1}} \mathrm{E}u(d, \omega_k)$.

For instance, if the utility $\mathrm{E}u(d, \omega_1)$ in the most probable event ω_1 is maximized at a unique decision, that is to say if D_1 is a singleton set, then this decision is the optimal one. If D_1 contains several elements, then searching for the best decision can proceed by leaving aside all decisions which are not in D_1 and going to the next most probable event in order to distinguish between the elements of D_1 and finding D_2. The algorithm proceeds until a single decision is reached, or until all events have been considered and thus the remaining decisions give the same utility in any event.

With the above "trembling ballot" model and a large enough number of voters, it will be proved that the law of lexicographic maximization applies, so that a voter's rational behavior is the "lexicographic response" described in the introduction: candidates are compared to the announced winner, and the announced winner himself is compared to his most serious contender. This is the content of Theorem 1. The most probable event (ω_1) is that there is no tie and that my vote makes no difference; this event leaves availlable all possible decisions (ballots) and $D_1 = D_0$ is the set of all possible ballots. The second most probable event (ω_2) is a tie between the two first-ranked candidates x_1 and x_2; in this case, I know which are the good decisions to take: depending on my preference, I approve of x_1 or x_2 but not both. This defines D_2, which is a strict subset of D_1 but not yet a singleton because I have not decided yet wether I approve of x_3 or not. It will be seen that for all $k \geq 2$, x_k appears in the algorithm to be compared to x_1, thereby defining a unique best response. Note that this is not an equilibrium effect: this strategy actually describes the best response to any strategy profile which generates no ties in the scores.

The above algorithm is sound from the behavioral point of view. Behavioral theories of individual choice often incorporates the idea that actual choices are obtained after a qualitative and discrete process of simplification and elimination, rather than after the weighting of probabilities and utilities which caracterises rational behavior. We refer the interested reader to the second volume of the classic book Suppes et al. (1989), pages 436-457, and to Tversky (1972a, 1972b) or Tversky and Sattah (1979). The main argu-

ment of the present paper is that this "boundedly rational" rule of behavior is in fact rational if the number of voters is large.

4   Results

We start by a simple lemma which is just the formal description of the voter's best responses when there is no uncertainty.

Lemma 1 *Suppose a given voter knows how the other voters vote. Let s^* be the highest score (computed from these other voters votes), let Y_1 be the set of candidates with score s^*, and let Y_2 be the set of candidates with score $s^* - 1$ (Y_2 can be empty). The best responses for this voter only depend on Y_1 and Y_2. Denoting by $\mathcal{B} = 2^X$ the set of ballots $B \subseteq X$, let $\phi(Y_1, Y_2) \subseteq \mathcal{B}$ be set of best responses.*

- *If $Y_1 = \{x_i\}$ and $Y_2 = \emptyset$: $\phi(Y_1, Y_2) = 2^{\mathcal{B}}$.*

- *If $Y_1 = \{x_i, x_j\}$, $Y_2 = \emptyset$ and the voters prefers x_j to x_i:*
  $\phi(Y_1, Y_2) = \{B \in \mathcal{B} : x_i \notin B, \ x_j \in B\}$.

- *If $Y_1 = \{x_i\}$, $Y_2 = \{x_j\}$ and the voters prefers x_j to x_i:*
  $\phi(Y_1, Y_2) = \{B \in \mathcal{B} : x_i \notin B, \ x_j \in B\}$.

- *If $Y_1 = \{x_i\}$, $Y_2 = \{x_j\}$ and the voters prefers x_i to x_j:*
  $\phi(Y_1, Y_2) = B \setminus \{B \in \mathcal{B} : x_i \notin B, \ x_j \in B\}$.

Proof. Clearly the voter's behavior only depends on the scores of the candidates that can get elected, thus the voter can condition her decision on the different possibilities for Y_1 and Y_2.

If Y_1 contains a single candidate and Y_2 is empty, then the voter can have no effect on who is elected so that all ballots are identical for him. One can write $\phi(Y_1, Y_2) = 2^{\mathcal{B}}$. Consider now the cases with two candidates.

If $Y_1 = \{x_i, x_j\}$ and $Y_2 = \emptyset$, then the voter should vote for the candidate she prefers among x_i and x_j, say x_j, and not for the other one. It does not matter for her wether she votes or not for any other candidate since they will not be elected: $\phi(Y_1, Y_2) = \{B \in \mathcal{B} : x_i \notin B, \ x_j \in B\}$.

If $Y_1 = \{x_i\}$ and $Y_2 = \{x_j\}$, then the voter can have either x_i elected or produce a tie between x_i and x_j. (1) If she prefers x_j to x_i, she also prefers a tie between them to x_i winning, so that a best response for her is any ballot that contains x_j and not x_i: $\phi(Y_1, Y_2) = \{B \in \mathcal{B} : x_i \notin B, \ x_j \in B\}$. (2) If she prefers x_i to x_j, she should either vote for both x_i and x_j, or

9

not vote for x_j and it does not matter wether or not she votes for any other candidate. This means that she must avoid ballots that contain x_j and not x_i. Her set of best responses is the complement of the previous set: $\phi(Y_1, Y_2) = B \setminus \{B \in \mathcal{B} : x_i \notin B, \ x_j \in B\}$.

If $Y_1 \cup Y_2$ contains three or more candidates then $\phi(Y_1, Y_2)$ is some subset of $\mathcal{B}$ that we do not need to specify. Notice that, in all cases, the set of best responses $\phi(Y_1, Y_2)$ does not depend on n. ∎

The next lemma describes how small the probability of a pivotal event is. To do so, some notation is helpful.

Definition 1 *For each non-empty subset Y of candidates, denote by $\mathsf{pivot}(n, Y)$ the event:*

$$\forall \, y \ \in \ Y, \ \mathcal{S}(y) \geq \max_{x \in X} \mathcal{S}(x) - 1$$
$$\forall \, y' \ \notin \ Y, \ \mathcal{S}(y') < \max_{x \in X} \mathcal{S}(x) - 1.$$

*Given two subsets Y, Y' of X, Y is a **more serious race** than Y' if*

$$\lim_{n \to +\infty} \frac{\Pr\left[\mathsf{pivot}(n, Y')\right]}{\Pr\left[\mathsf{pivot}(n, Y)\right]} = 0.$$

This is denoted
$$Y' \ll Y.$$

Lemma 2 *Suppose that there are no ties in the score vector s, then the two-candidate races involving a given candidate x_i are ordered:*

$$\{x_i, x_K\} \ll \{x_i, x_{K-1}\} \ll ... \ll \{x_i, x_{i+1}\} \ll \{x_i, x_{i-1}\} \ll ... \ll \{x_i, x_1\}.$$

Moreover, if Y contains three or more candidates then

$$Y \ll \{x_i, x_j\}$$

for x_i, x_j two of them.

(This lemma is proved in the appendix.) Remark that the lemma does not say, for instance, which one of the two races $\{x_1, x_4\}$ and $\{x_2, x_3\}$ is the most serious. The lemma says that $\{x_1, x_4\}$ is the most serious among all the races that involve candidate x_4, and this is all what will be needed about candidate x_4, and similarly for the other candidates. We can now state and prove the key result in this paper.

Theorem 1 *Let s be a score vector with two candidates at the two first places and no tie: x_1 and x_2 such that $s(x_1) > s(x_2) > s(y)$ for $y \in X$, $y \neq x_1, x_2$. There exists n_0 such that, for all $n > n_0$ and for all type τ, any type-τ voter has a unique best-response ballot B_τ^*. This ballot is described by the following rule (the "lexicographic response" to s):*

- *for τ such that $u_\tau(x_1) > u_\tau(x_2)$, $B_\tau^* = \{x \in X : u_\tau(x) \geq u_\tau(x_1)\}$,*

- *for τ such that $u_\tau(x_1) < u_\tau(x_2)$, $B_\tau^* = \{x \in X : u_\tau(x) > u_\tau(x_1)\}$.*

Proof. Suppose first that $u_\tau(x_1) > u_\tau(x_2)$. To prove that B_τ^* is the unique best response, we prove that any other ballot B is not. More precisely, given $B \neq B_\tau^*$, we will find a strictly better response B' which differs from B only on one candidate: B' will be of the form either $B' = B \cup \{x\}$ or $B' = B \setminus \{x\}$. Hence the following observation will be very useful.

First point. Let $\mathrm{E}\left[u_\tau(B)\right]$ and $\mathrm{E}\left[u_\tau(B')\right]$ denote the expected utility of strategy B and B' for a voter i of type τ, given the other voters' strategies. Suppose that $B' = B \cup \{x\}$ for some $x \notin B$. Then if the vote of i in favor of x is not recorded, which was denoted as $\eta_{i,x} = 0$ and happens with probability ε, voting B' is the same as voting B, thus:

$$\mathrm{E}\left[u_\tau(B')\right] = \varepsilon \mathrm{E}\left[u_\tau(B)\right] + (1 - \varepsilon)\mathrm{E}\left[u_\tau(B') \mid \eta_{i,x} = 1\right]$$

where $\mathrm{E}\left[u_\tau(B') \mid \eta_{i,x} = 1\right]$ denotes the expected utility of casting ballot B' knowing that there is no mistake in the recording of the approval of candidate x by voter i. It follows that

$$\mathrm{E}\left[u_\tau(B') - u_\tau(B)\right] = (1 - \varepsilon)\mathrm{E}\left[\left(u_\tau(B') - u_\tau(B)\right) \mid \eta_{i,x} = 1\right],$$

which means that the comparison between B and $B' = B \cup \{x\}$ for a voter can be made under the assumption that a vote in favor of x would be recorded for sure. Of course, the same is true when comparing B and $B' = B \setminus \{x\}$ for $x \in B$. To lighten notation, in the sequel we will forget the condition and write $\mathrm{E}\left[\cdot\right]$ for $\mathrm{E}\left[\cdot \mid \eta_{i,x} = 1\right]$.

This first point being made, we shall distinguish three cases, depending where x_1 stands in the voter's preference and ballot.

(i) Suppose that there exits $x \in X$ such that $u_\tau(x) > u_\tau(x_1)$ and $x \notin B$. Let then $B' = B \cup \{x\}$.

To compare ballots B and B', the voter computes the difference

$$\Delta = \sum_Y \Pr[\mathsf{pivot}(n, Y)] \, \mathrm{E}\left[u_\tau(B', Y) - u_\tau(B, Y)\right],$$

where $Eu_\tau(B, Y)$ denotes the expected utility for her of choosing ballot B conditional on the event $\mathsf{pivot}(n, Y)$. The events $\mathsf{pivot}(n, Y)$ with $x \notin Y$ have no consequences therefore the sum in Δ can run over the subsets Y such that $x \in Y$. Among these is $\{x, x_1\}$. From Lemma 1, it follows that in all three cases ($Y_1 = \{x, x_1\}$ and $Y_2 = \emptyset$, $Y_1 = \{x\}$ and $Y_2 = \{x_1\}$, $Y_1 = \{x_1\}$ and $Y_2 = \{x\}$), the voter's utility is strictly larger voting for x than not. Thus $E\left(u_\tau(B', \{x, x_1\}) - u_\tau(B, \{x, x_1\})\right) > 0$.

As proved in Lemma 2, the probability of $\mathsf{pivot}(n, Y)$ is decreasing in n if $\#Y \geq 2$ in such a way that for all $Y \subseteq X$ with $x \in Y$, $Y \neq \{x, x_1\}$ and $\#Y \geq 2$,

$$\lim_{n \to \infty} \frac{\Pr[\mathsf{pivot}(n, Y)]}{\Pr[\mathsf{pivot}(n, \{x, x_1\})]} = 0.$$

One can thus factor out $\Pr[\mathsf{pivot}(n, \{x, x_1\})]$ in Δ and write

$$\frac{\Delta}{\Pr[\mathsf{pivot}(n, \{x, x_1\})]} = a + \sum_{\substack{Y \neq \{x, x_1\} \\ \#Y \geq 2}} o_Y(n)$$

for

$$a = E\left[u_\tau(B', \{x, x_1\}) - u_\tau(B, \{x, x_1\})\right]$$

and

$$o_Y(n) = \frac{\Pr[\mathsf{pivot}(n, Y)]}{\Pr[\mathsf{pivot}(n, \{x, x_1\})]} E\left[u_\tau(B', Y) - u_\tau(B, Y)\right],$$

with a being strictly positive and with $o_Y(n)$ tending to 0 when n tends to infinity. It follows that, for n large enough, $\Delta > 0$. This establishes that B is not a best response, more exactly there exists a number $n_{B, \tau}$ depending on B and τ such that, for all $n > n_{B, \tau}$, B is not a best response. Because there is a finite number of ballots and types, $n_{B, \tau}$ can be chosen independent of B and τ.

(ii) Suppose that there exists $x \in X$ such that $u_\tau(x) < u_\tau(x_1)$ and $x \in B$. Let then $B' = B \setminus \{x\}$. The reasoning is the same as in the previous case: The relevant events are again $\mathsf{pivot}(n, Y)$ for $x \in Y$, the voter's utility is strictly larger not voting for x, and the relevant race is again $\{x, x_1\}$.

(iii) Suppose that $x_1 \notin B$. The conclusion follows considering $B' = B \cup \{x_1\}$, the race $\{x_1, x_2\}$ is here relevant and is the most serious one.

From items (i), (ii) and (iii) it follows that the voter's best response must satisfy (i) $u_\tau(x) \geq u_\tau(x_1)$ if $x \in B$, (ii) $u_\tau(x) \leq u_\tau(x_1)$ if $x \notin B$, and (iii) $x_1 \in B$. Therefore $B_\tau^* = \{x \in X : u_\tau(x) \geq u_\tau(x_1)\}$ as stated. The argument is identical in the case $u_\tau(x_1) < u_\tau(x_2)$. ∎

Notice that the previous result implies that, for n large enough, all voters of a given type use the same strategy when responding to a score vector that satisfy the mentioned properties. The next definition is standard in the study of approval voting.

Definition 2 *A ballot B is **sincere** for a type-τ voter if $u_\tau(x) > u_\tau(y)$ for all $x \in B$ and $y \notin B$.*

As a direct consequence of the previous theorem, one gets the following corollary.

Corollary 1 *For a large electorate, in the absence of tie the best response is sincere.*

It should be emphasized that the previous theorem and corollary are true even out of equilibrium; they just describe the voter's response to a conjecture she holds about the candidate scores. Precisely, the voter takes into account that she estimates the candidate scores with a statistical disturbance of order $1/n$. We now turn to equilibrium considerations. The piece of notation $p[x, y]$ will be used to denote the fraction of voters who prefer x to y:

$$p[x, y] = \sum_{\tau:\, x\, P_\tau\, y} p_\tau,$$

so that for all $x \neq y$, $p[x, y] + p[y, x] = 1$. The number of voters who prefer x to y is $nvp[x, y]$.

Theorem 2 *Let s be a vector with two candidates at the two first places and no tie: x_1 and x_2 such that $s(x_1) > s(x_2) > s(y)$ for $y \in X$, $y \neq x_1, x_2$. There exists n_0 such that, for all $n > n_0$, if s is the score vector of an equilibrium of the game with n voters, then*

- *the score of the first-ranked candidate is his majoritarian score against the second-ranked candidate:*

$$s(x_1) = p\,[x_1, x_2]\,,$$

- *the score of any other candidate is his majoritarian score against the first-ranked candidate:*

$$x \neq x_1 \Rightarrow s(x) = p\,[x, x_1]\,.$$

Proof. This is a direct consequence of Theorem 1. Each voter approves of x_1 if and only of she prefers x_1 to x_2. For $x \neq x_1$, she approves of x if and only if she prefers x to x_1. ∎

Definition 3 *A **Condorcet winner** at profile p is a candidate x_1 such that $p[x_1, x] > 1/2$ for all $x \neq x_1$. A **best contender** to a candidate x_1 is a candidate x_2 such that $p[x_2, x_1] \geq p[x, x_1]$ for all $x \neq x_1$.*

As it is well known, the existence of a Condorcet winner is a substantial assumption. On the contrary the assumption that there is a unique best contender to a candidate x is an innocuous one, inasmuch as one can assume away exact ties in the preference profile.[4]

Corollary 2 *For a large electorate, if there is an equilibrium with no tie, the winner of the election is a Condorcet winner. If the preference profile admits a Condorcet winner and the Condorcet winner has a unique best contender then the game has a unique equilibrium, in this equilibrium the Condorcet winner is elected.*

Proof. Notice that, as a consequence of Theorem 1, no voter votes simultaneously for x_1 and x_2, and each of them votes for either x_1 or x_2. Here,

$$s(x_2) = 1 - s(x_1) < s(x_1)$$

implies $s(x_2) < 1/2$, thus, for $x \neq x_1$

$$p[x, x_1] = s(x) \leq s(x_2) < \frac{1}{2}.$$

∎

These results have a non-trivial implication for the case of a preference profile with no Condorcet winner, although they do not allow for a complete characterization of equilibrium in that case. In that case, at equilibrium, the score vector must exhibit a tie between the two top candidates or between several second-ranked candidates. Approval voting is not a solution to the so-called Condorcet paradox. Indeed, approval voting retains the basic disequilibrium property of majority rule: if there is no Condorcet winner then any announced winner will be defeated, according to approval voting, by another candidate, preferred to the former by more than half of the population.

---

[4] Things would be different in a model of party competition with endogeneous candidate positions.

The key to this result is that, with a large population, voters strategic thinking has to put special emphasis on pairwise comparisons of candidates, even if the voting rule itself is not defined in terms of pairwise comparisons. Therefore a natural avenue of research is to check wether the arguments here provided in the case of approval voting can be extended to other voting rules.

A Appendix: Probabilities of pivotal events

A.1 Proof of Lemma 2

When a voter i, in the n-replicate population, wonders whom to approve of, the pivotal events to be taken into account are the almost ties arising from the other voters strategies. These events are described either by random variables of the form $\mathcal{S}_k^n \rightsquigarrow \mathcal{B}(nvs(k), 1 - \varepsilon)$, obtained from $nvs(k)$ votes, or by random variables $\widetilde{\mathcal{S}}_k^n \rightsquigarrow \mathcal{B}(nvs(k) - 1, 1 - \varepsilon)$, obtained from $nvs(k) - 1$ votes; the second case arises if the voters of the same type as i approve of k at the considered profile. In the first part of the proof ("Computation in the binomial model") we will consider that all variables are of the form $\mathcal{S}_k^n$. In the second part of the proof ("Environmental invariance") we will check that the proof is valid in the general case because the ties arising from variables $\mathcal{S}_k^n$ or $\widetilde{\mathcal{S}}_k^n$ have the same order of magnitude.

A.1.1 Computation in the binominal model

For an integer n, let S_k^n, for $k = 1, ..., K$, be independent binomial random variables of parameters nvs_k and $1 - \varepsilon$:

$$\mathcal{S}_k^n \rightsquigarrow \mathcal{B}(nvs(k), 1 - \varepsilon)$$

The probability that $\mathcal{S}_k^n = i$ is:

$$\Pr\left[\mathcal{S}_k^n = i\right] = C_{nvs(k)}^i \left(1 - \varepsilon\right)^i \varepsilon^{nvs(k)}$$

for $i = 0, ..., nvs(k)$. To lighten notation, write

$$n_i \equiv nvs(k_i).$$

We suppose that the parameters $s(k)$ are strictly ordered:

$$s(1) > s(2) > ... > s(K).$$

First step. We must consider the occurence of ties up to one vote. In a first step, the event of an *exact* tie at the top between two candidates. Let k_1 and k_2 with $k_1 < k_2$ such that:

$$\forall k \neq k_1, k_2 , \ \mathcal{S}_{k_1}^n = \mathcal{S}_{k_2}^n > \mathcal{S}_k^n .$$

The probability of this event is:

$$
\mathsf{P}(n, k_1, k_2) \ = \ \sum_{i=0}^{n_2} C_{n_1}^i C_{n_2}^i \left(1 - \varepsilon\right)^{2i} \tag{1}
$$
$$
\times \varepsilon^{n_1 + n_2 - 2i} \Pr\left[\forall k \neq k_1, k_2, \mathcal{S}_k^n < i\right]
$$

Now consider three candidates:

$$k_1 < k_2 < k_3.$$

We will prove that the probability of the race between k_1 and k_3 is very small compared to the probability of the race between k_1 and k_2. Write

$$
\begin{aligned}
\mathsf{P}(n, k_1, k_3) &= \sum_{i=0}^{n_3} \sum_{j=0}^{i} C_{n_1}^i C_{n_3}^i C_{n_2}^j (1-\varepsilon)^{2i+j} \\
&\quad \times \varepsilon^{n_1+n_2+n_3-2i-j} \Pr\left[\forall k \neq k_1, k_2, k_3, \mathcal{S}_k^n < i\right].
\end{aligned}
$$

To decipher this formula: the index i is the value of $\mathcal{S}_{k_1}^n = \mathcal{S}_{k_3}^n$ and j the value of $\mathcal{S}_{k_2}^n$.

I claim that, for $j < i, C_{n_3}^i C_{n_2}^j < C_{n_3}^j C_{n_2}^i$. To see this, write

$$
\begin{aligned}
\frac{C_{n_3}^i C_{n_2}^j}{C_{n_3}^j C_{n_2}^i} &= \frac{n_3(n_3-1)\dots(n_3-i+1)}{n_3(n_3-1)\dots(n_3-j+1)} \\
&\quad \times \frac{n_2(n_2-1)\dots(n_2-j+1)}{n_2(n_2-1)\dots(n_2-i+1)} \\
&= \frac{(n_3-j)\dots(n_3-i+1)}{(n_2-j)\dots(n_2-i+1)}
\end{aligned}
$$

and notice that for each of the $i-j$ remaining term, $n_3 - u < n_2 - u$, hence the ratio is smaller than 1.

It follows that if we take again the above formula for and exchange i and j in $C_{n_2}^i$ and $C_{n_3}^j$ and if we denote

$$
\begin{aligned}
B_n &= \sum_{i=0}^{n_3} \sum_{j=0}^{i} C_{n_1}^i C_{n_2}^i C_{n_3}^j (1-\varepsilon)^{2i+j} \\
&\quad \times \varepsilon^{n_1+n_2+n_3-2i-j} \Pr\left[\forall k \neq k_1, k_2, k_3, \mathcal{S}_k^n < i\right],
\end{aligned}
$$

we obtain:

$$\mathsf{P}(n, k_1, k_3) < B_n.$$

Notice that B_n is a part of the summation that defines $\mathsf{P}(n, k_1, k_2)$:

$$
\begin{aligned}
\mathsf{P}(n, k_1, k_2) &= \sum_{i=0}^{n_2} \sum_{j=0}^{i} C_{n_1}^i C_{n_2}^i C_{n_3}^j (1-\varepsilon)^{2i+j} \\
&\quad \times \varepsilon^{n_1+n_2+n_3-2i-j} \Pr\left[\forall k \neq k_1, k_2, k_3, \mathcal{S}_k^n < i\right] \\
&= B_n + C_n,
\end{aligned}
$$

17

where C_n is the same sum for i going from $n_3 + 1$ to n_2. One can write:

$$
\begin{aligned}
B_n &= \sum_{i=0}^{n_3} \Pr\left[\mathcal{S}_{k_1}^n = \mathcal{S}_{k_2}^n = i\right] \Pr\left[\forall k \neq k_1, k_2, \mathcal{S}_k^n < i\right], \\
C_n &= \sum_{i=n_3+1}^{n_2} \Pr\left[\mathcal{S}_{k_1}^n = \mathcal{S}_{k_2}^n = i\right] \Pr\left[\forall k \neq k_1, k_2, \mathcal{S}_k^n < i\right].
\end{aligned}
$$

Using that $\Pr\left[\forall k \neq k_1, k_2, \mathcal{S}_k^n < i\right]$ is obviously increasing with i,

$$
\begin{aligned}
B_n &< \Pr\left[\forall k \neq k_1, k_2, \mathcal{S}_k^n < n_3\right] \sum_{i=0}^{n_3} \Pr\left[\mathcal{S}_{k_1}^n = \mathcal{S}_{k_2}^n = i\right], \\
C_n &> \Pr\left[\forall k \neq k_1, k_2, \mathcal{S}_k^n < n_3\right] \sum_{i=n_3+1}^{n_2} \Pr\left[\mathcal{S}_{k_1}^n = \mathcal{S}_{k_2}^n = i\right].
\end{aligned}
$$

For ε such that $n_2 < (1 - \varepsilon)n_1$, $\Pr\left[\mathcal{S}_{k_1}^n = i\right]$ is increasing with i for $i \in \{0, ..., n_2\}$: To check this point, note that for $i \leq (1 - \varepsilon)n_1 - 1$,

$$
\begin{aligned}
\Pr\left[\mathcal{S}_{k_1}^n = i + 1\right] &= C_{n_1}^{i+1} (1 - \varepsilon)^{i+1} \varepsilon^{n_1 - i - 1} \\
\frac{\Pr\left[\mathcal{S}_{k_1}^n = i + 1\right]}{\Pr\left[\mathcal{S}_{k_1}^n = i\right]} &= \frac{(n_1 - i)(1 - \varepsilon)}{(i + 1)\varepsilon} \geq \frac{n_1 - i}{\varepsilon n_1} \geq \frac{\varepsilon n_1 + 1}{\varepsilon n_1} > 1.
\end{aligned}
$$

The inequality $n_2 < (1 - \varepsilon)n_1$ can be written: $\varepsilon < (vs(k_1) - vs(k_2))/vs(k_1)$, and $vs(k_1)$ and $vs(k_2)$ are two different integers smaller than the initial number of voters v. Therefore this inequality is true for all n for all $\varepsilon < 1/v$. Under this condition, it follows:

$$
\begin{aligned}
B_n &< \Pr\left[\forall k \neq k_1, k_2, \mathcal{S}_k^n < n_3\right] \Pr\left[\mathcal{S}_{k_1}^n = n_3\right] \sum_{i=0}^{n_3} \Pr\left[\mathcal{S}_{k_2}^n = i\right], \\
C_n &> \Pr\left[\forall k \neq k_1, k_2, \mathcal{S}_k^n < n_3\right] \Pr\left[\mathcal{S}_{k_1}^n = n_3\right] \sum_{i=n_3+1}^{n_2} \Pr\left[\mathcal{S}_{k_2}^n = i\right]
\end{aligned}
$$

and thus

$$
\frac{B_n}{C_n} < \frac{\Pr\left[\mathcal{S}_{k_2}^n \leq n_3\right]}{\Pr\left[\mathcal{S}_{k_2}^n > n_3\right]}.
$$

When n tends to infinity, the variable $\mathcal{S}_{k_2}^n/n$ is approximately normal, with mean $\mu = (1 - \varepsilon)vs(k_2)$ and variance $\varepsilon(1 - \varepsilon)vs(k_2)/n$. Like previously,

$\varepsilon < 1/v$ implies that $vs(k_3) < \mu$. The weak law of large number implies that the probability that $\mathcal{S}_{k_2}^n/n$ is less than $vs(k_3)$ tends to 0 when n tends to infinity, which means here that $\Pr\left[\mathcal{S}_{k_2}^n \leq n_3\right]$ tends to 0 and thus $\Pr\left[\mathcal{S}_{k_2}^n > n_3\right]$ tends to 1. Therefore $\frac{B_n}{C_n}$ tends to 0. Recall that $\mathsf{P}(n, k_1, k_3) < B_n$ and $\mathsf{P}(n, k_1, k_2) = B_n + C_n$, it follows that

$$\lim_{n \to \infty} \frac{\mathsf{P}(n, k_1, k_3)}{\mathsf{P}(n, k_1, k_2)} = 0.$$

Second step. Consider next the event $\mathsf{pivot}(n, \{k_1, k_2\})$ of an almost exact tie at the top between the two candidates k_1 and k_2 with $k_1 < k_2$. This is the union of the three disjoint events:

$$\forall k \neq k_1, k_2 \ , \ \mathcal{S}_{k_1}^n = \mathcal{S}_{k_2}^n > \mathcal{S}_k^n$$
$$\text{or} \ \ \forall k \neq k_1, k_2 \ , \ \mathcal{S}_{k_1}^n + 1 = \mathcal{S}_{k_2}^n > \mathcal{S}_k^n$$
$$\text{or} \ \ \forall k \neq k_1, k_2 \ , \ \mathcal{S}_{k_2}^n + 1 = \mathcal{S}_{k_1}^n > \mathcal{S}_k^n$$

Denote by $\mathsf{P}(n, k_1, k_2)$, $\mathsf{P}'(n, k_1, k_2)$, and $\mathsf{P}''(n, k_1, k_2)$ the probabilities of these three events. For k_3 with $k_1 < k_2 < k_3$, the same reasoning which led in the first step to $\lim_{n \to \infty} \frac{\mathsf{P}(n,k_1,k_3)}{\mathsf{P}(n,k_1,k_2)} = 0$ can be made and leads to the two same other conclusions $\lim_{n \to \infty} \frac{\mathsf{P}'(n,k_1,k_3)}{\mathsf{P}'(n,k_1,k_2)} = 0$ and $\lim_{n \to \infty} \frac{\mathsf{P}''(n,k_1,k_3)}{\mathsf{P}''(n,k_1,k_2)} = 0$. Therefore $\lim_{n \to \infty} \frac{\mathsf{P}(n,k_1,k_3)+\mathsf{P}'(n,k_1,k_3)+\mathsf{P}''(n,k_1,k_3)}{\mathsf{P}(n,k_1,k_2)+\mathsf{P}'(n,k_1,k_2)+\mathsf{P}''(n,k_1,k_2)} = 0$ and we obtain that

$$\lim_{n \to \infty} \frac{\Pr[\mathsf{pivot}(n, \{k_1, k_3\})]}{\Pr[\mathsf{pivot}(n, \{k_1, k_2\})]} = 0. \tag{2}$$

Third step. It remains to show that three-way (or more) ties are negligible compared to two-way ties. This easy point is left to the reader.

A.1.2 Environmental invariance

Let

$$\widetilde{\mathcal{S}}_k^n \rightsquigarrow \mathcal{B}(nvs(k) - 1, 1 - \varepsilon).$$

Consider a strategy profile in which type-τ voters approve of candidate k, for a type τ voter, the number of votes in favor of k is $\widetilde{\mathcal{S}}_k^n$ if this particular voter decides not to approve of k,.or $\mathcal{S}_k^n$ if he decides to approve of k. We now check that this feature has no consequence on the voter's decision because the ties arising from variables $\mathcal{S}_k^n$ or $\widetilde{\mathcal{S}}_k^n$ have the same order of magnitude. To see this, consider for instance the event of an exact tie at the top between k_1 and k_2 when the score of k_1 is

$$\widetilde{\mathcal{S}}_{k_1}^n \rightsquigarrow \mathcal{B}(nvs(k_1) - 1, 1 - \varepsilon)$$

and the score of k_2 is

$$\mathcal{S}^n_{k_2} \rightsquigarrow \mathcal{B}(nvs(k_2), 1 - \varepsilon).$$

The event

$$\forall k \neq k_1, k_2 , \ \widetilde{\mathcal{S}}^n_{k_1} = \mathcal{S}^n_{k_2} > \mathcal{S}^n_k$$

has the following probability:

$$\begin{aligned}
\mathsf{Q}(n, k_1, k_2) &= \sum_{i=0}^{n_2} C^i_{n_1-1} C^i_{n_2} (1 - \varepsilon)^{2i} \\
&\times \varepsilon^{n_1 - 1 + n_2 - 2i} \Pr\left[\forall k \neq k_1, k_2, \mathcal{S}^n_k < i\right]
\end{aligned}$$

(we still denote $n_1 = ns(k_1)$ and $n_2 = ns(k_2)$), to be compared with $\mathsf{P}(n, k_1, k_2)$ in formula 1. Because $C^i_{n_1-1} < C^i_{n_1}$, it is easy to see that

$$\mathsf{Q}(n, k_1, k_2) < (1/\varepsilon)\mathsf{P}(n, k_1, k_2).$$

Using the fact that

$$C^i_{n_1} = \frac{n_1}{n_1 - i} C^i_{n_1-1} \leq \frac{n_1}{n_1 - n_2} C^i_{n_1-1},$$

one can also see that

$$\mathsf{P}(n, k_1, k_2) < \varepsilon \frac{n_1}{n_1 - n_2} \mathsf{Q}(n, k_1, k_2).$$

We thus obtain that, for all n,

$$\varepsilon < \frac{\mathsf{P}(n, k_1, k_2)}{\mathsf{Q}(n, k_1, k_2)} < \varepsilon \frac{s(k_1)}{s(k_1) - s(k_2)}. \tag{3}$$

We leave to the reader the easy verification of similar formulas for all events of the form

$$\forall k \neq k_1, k_2 , \ \widehat{\mathcal{S}}^n_{k_1} = \widehat{\mathcal{S}}^n_{k_2} > \widehat{\mathcal{S}}^n_k ,$$

where the variables $\widehat{\mathcal{S}}^n_k$ are equal either to $\mathcal{S}^n_k$ or to $\widetilde{\mathcal{S}}^n_k$, as well as for the almost tie events of the form

$$\forall k \neq k_1, k_2 , \ \widehat{\mathcal{S}}^n_{k_1} + 1 = \widehat{\mathcal{S}}^n_{k_2} > \widehat{\mathcal{S}}^n_k$$

and

$$\forall k \neq k_1, k_2 , \ \widehat{\mathcal{S}}^n_{k_2} + 1 = \widehat{\mathcal{S}}^n_{k_1} > \widehat{\mathcal{S}}^n_k .$$

From the observation 3 one can deduce that the conclusion 2 that the race $\{k_1, k_3\}$ is negligeable compared to the race $\{k_1, k_3\}$ holds for all the voters, independently of who this voter is voting for. This point plays the role of the "environmental invariance" property that Myerson obtains with Poisson random variables. With binomial variables, different voters do not have exactly the same information, but the difference is so small that it does not matter for finding best responses.

A.2 A continuous Gaussian model

It may be surprising that we proved our results without using the approximation of binomial distributions by normal distributions. Indeed an early version of this paper was written systematically using normal approximations (Laslier 2004). This methodology leads, at rather low cost, to the same qualitative results, and is also helpful to get intution on this model. Unfortunately, it turns out that the normal approximation is not mathematicaly correct in the present case. In this appendix, we provide a simple continuous Gaussian, model which looks very much like an approximation of the discrete model and we explain why this continuous model is in fact not a correct approximation.

A.2.1 Magnitude of pivotal events in the Gaussian approximation

Let $\mathbf{b}_k$, for $k = 1, ..., K$, be independent normal random variables. The mean value of $\mathbf{b}_k$ is denoted by a_k and one supposes $a_1 \geq a_2 \geq ... \geq a_K$. The variance of $\mathbf{b}_k$ is σ^2/n, where σ^2 is a fixed parameter and n is a (large) number. All variables $\mathbf{b}_k$ have the same variance[5]. For convenience we denote

$$a_{ij} = \frac{a_i + a_j}{2}.$$

Definition 4 *Given a score vector s, denote by x_i, for $i = 1, ..., K$ the candidates ordered so that $s(x_1) \geq s(x_2) \geq ... \geq s(x_K)$; for $i \neq j$, the* **magnitude** *of the race $\{x_i, x_j\}$ is:*

$$\beta_{i,j} = \lim_{n \to +\infty} \frac{1}{n} \log \Pr[\mathsf{pivot}(n, \{x_i, x_j\})].$$

---

[5] In the exat model of the text, the variance depends on the score: $\text{var} = \frac{\varepsilon(1-\varepsilon)s(x)}{nv}$. The simplification done here is of no consequence for the point we make in this Appendix.

Lemma 3 *The magnitudes of the two-candidate races involving a given candidate x_i are ordered:*

$$\beta_{i,K} < \beta_{i,K-1} < ... < \beta_{i,i+1} < \beta_{i,i-1} < ... < \beta_{i,1}.$$

Moreover, if Y contains three or more candidates then

$$\lim_{n \to +\infty} \frac{1}{n} \log \Pr[\mathsf{pivot}(n, Y)] < \beta_{i,j}$$

for x_i, x_j two of them.

The computations are explained in this appendix under the assumption that for no $i \neq j$ there exists k such that $a_k = a_{ij}$. It is not more difficult to arrive at the conclusion through the same type of computations in the case where for some $i \neq j$ there exists k such that $a_k = a_{ij}$.

Consider the event $\mathsf{pivot}(n, \{i, j\})$ of a race between two candidates:

$$\mathbf{b}_j \in [\mathbf{b}_i - 1/n, \ \mathbf{b}_i + 1/n[\, ,$$
$$\forall k \neq i, j \ , \ \ \mathbf{b}_k + 1/n < \mathbf{b}_i, \mathbf{b}_j$$

We will prove that

$$\lim_{n \to +\infty} \frac{1}{n} \log \Pr\left[\mathsf{pivot}(n, \{i, j\})\right] = -\sum_{k=1}^{k_{ij}} \frac{1}{2\sigma^2} \left(a_k - a_{ij}\right)^2 - \frac{(a_i - a_j)^2}{4\sigma^2}, \quad (4)$$

where k_{ij} is the last integer k such that $a_k > a_{ij}$.

Tie between the two first candidates

To start by the simplest case, consider the event $\mathsf{pivot}(n, \{x_1, x_2\})$: It is the disjoint union of the two events

$$\forall k = 3, ..., K \ , \ \ \mathbf{b}_k < \mathbf{b}_1 - 1/n < \mathbf{b}_2 < \mathbf{b}_1$$

and

$$\forall k = 3, ..., K \ , \ \ \mathbf{b}_k < \mathbf{b}_2 - 1/n < \mathbf{b}_1 < \mathbf{b}_2$$

The probability of the former writes :

$$\int_{b_1=-\infty}^{+\infty} f(b_1; a_1, \frac{\sigma^2}{n}) \int_{b_2=b_1-1/n}^{b_1} f(b_2; a_2, \frac{\sigma^2}{n}) \prod_{k=3}^{K} F(b_1 - \frac{1}{n}; a_k, \frac{\sigma^2}{n}) \, db_2 \, db_1,$$

where f and F denote the normal density and cumulative functions

$$f(t; \mu, \sigma^2) = \frac{1}{\sqrt{2\sigma^2\pi}} \exp -\frac{1}{2\sigma^2}(t - \mu)^2,$$

$$F(t; \mu, \sigma^2) = \int_{u=-\infty}^{t} f(u; \mu, \sigma^2) \, du.$$

Because the inner integral on b_2:

$$\int_{b_2=b_1-1/n}^{b_1} f(b_2; a_2, \sigma^2/n) \, db_2$$

is close to $\frac{1}{n} f(b_1; a_2, \sigma^2/n)$, the probability of the former event is close to $A_{12}/2$, with:

$$A_{12} = \frac{2}{n} \int_{b_1=-\infty}^{+\infty} \prod_{k=3}^{K} F(b_1 - \frac{1}{n}; a_k, \frac{\sigma^2}{n}) \, f(b_1; a_1, \frac{\sigma^2}{n}) \, f(b_1; a_2, \frac{\sigma^2}{n}) \, db_1.$$

The same approximation is valid for the complementary event, so that the probability of the race $\{i, j\}$ is approximately

$$\Pr\left[\mathsf{pivot}(n, \{i, j\})\right] \simeq A_{12}$$

The product of two normal densities can be written

$$f(b_1; a_2, \sigma^2/n) \, f(b_1; a_1, \sigma^2/n)$$
$$= \frac{n}{2\sigma^2\pi} \exp -\frac{n}{2\sigma^2} \left[(b_1 - a_1)^2 + (b_1 - a_2)^2\right]$$
$$= \frac{n}{2\sigma^2\pi} \exp -\frac{n}{\sigma^2} \left[(b_1 - a_{12})^2 + \frac{(a_1 - a_2)^2}{4}\right]$$
$$= \frac{1}{2}\sqrt{\frac{n}{\sigma^2\pi}} \left(\exp -\frac{n(a_1 - a_2)^2}{4\sigma^2}\right) f(b_1; a_{12}, \frac{\sigma^2}{2n})$$

so that one gets:

$$A_{12} = \alpha_{12} \int_{b_1=-\infty}^{+\infty} \prod_{k=3}^{K} F(b_1 - \frac{1}{n}; a_k, \frac{\sigma^2}{n}) \, f(b_1; a_{12}, \frac{\sigma^2}{2n}) \, db_1,$$

with

$$\alpha_{12} = \frac{1}{\sqrt{n\sigma^2\pi}} \exp -\frac{n(a_1 - a_2)^2}{4\sigma^2}.$$

23

For n large, $F(b_1 - 1/n; a_k, \sigma^2/n)$ tends to 1 if $b_1 > a_k$ and to 0 if $b_1 < a_k$, and the density $f(b_1; a_{12}, \frac{\sigma^2}{2n})\, db_1$ tends to a Dirac mass at point $b_1 = a_{12}$. Here $a_k < a_{12}$, so that the integral in A_{12} tends to 1 and one finally gets:

$$\log A_{12} \simeq -n\, \frac{(a_1 - a_2)^2}{4\sigma^2}.$$

General case

More generally, consider i and j such that $1 < i < j$. The probability of the event $\mathsf{pivot}(n, \{i, j\})$ is approximately

$$A_{ij} = \alpha_{ij} \int_{b_i=-\infty}^{+\infty} \prod_{k \neq i,j} F(b_i - \frac{1}{n}; a_k, \frac{\sigma^2}{n})\, f(b_i; a_{ij}, \frac{\sigma^2}{2n})\, db_i,$$

with

$$\alpha_{ij} = \frac{1}{\sqrt{n\sigma^2\pi}} \exp -\frac{n(a_i - a_j)^2}{4\sigma^2}.$$

One still has that the density $f(b_i; a_{ij}, \frac{\sigma^2}{2n})\, db_i$ tends to a Dirac mass at point $b_i = a_{ij}$, but now $F(a_{ij}; a_k, \sigma^2/n)$ tends to 1 only for those k such that $a_k < a_{ij}$. Denote them by $k = k_{ij} + 1, k_{ij} + 2, ..., K$. For $k = 1, ..., k_{ij}$, one uses the standard approximation of the tail of the normal distribution. Recall that, for $t \gg 1$,

$$
\begin{aligned}
\frac{1}{\sqrt{2\pi}} \int_t^{+\infty} e^{-\frac{1}{2}u^2}\, du \;\; &\simeq \;\; \frac{1}{\sqrt{2\pi}} \int_t^{+\infty} e^{\frac{1}{2}t^2 - tu}\, du \\
&= \;\; \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2}t^2} \int_t^{+\infty} e^{-tu}\, du \\
&= \;\; \frac{1}{t\sqrt{2\pi}} e^{-\frac{1}{2}t^2}.
\end{aligned}
$$

One so gets that for $k = 1, ..., k_{ij}$,

$$F(a_{ij}; a_k, \sigma^2/n) = \frac{1}{(a_k - a_{ij})} \sqrt{\frac{\sigma^2}{2n\pi}} \exp -\frac{n}{2\sigma^2}(a_k - a_{ij})^2.$$

Then

$$\log A_{ij} \simeq -\sum_{k=1}^{k_{ij}} \frac{n}{2\sigma^2}(a_k - a_{ij})^2 - \frac{n(a_i - a_j)^2}{4\sigma^2}$$

and the expression (4) follows. The first part of Lemma 3 is easily deduced from these formulae. The second part of the lemma (about three way ties or more) is easily obtained by the same arguments.

A.2.2   A caveat about the Gaussian model

When approximating a binomial distribution by the normal distribution, one must be careful because the approximation is not valid for arbitrary small events. But, precisely, the problem under consideration here is to compare very rare events. The mistake that one does when directly considering a continuous model is the following.

One considers the binomial variable

$$\mathcal{S}^n \rightsquigarrow \mathcal{B}(n, 1 - \varepsilon)$$

and a rational number α, and wonders how fast is the probability

$$p_n = \Pr\left[\mathcal{S}^n = \alpha n\right]$$

going to 0.

The correct answer to this question is found by noting that

$$p_n(\alpha) = C_n^{\alpha n}(1 - \varepsilon)^{\alpha n}\varepsilon^{(1-\alpha)n}$$

and using Stirling's formula:

$$\lim_{n\to\infty} \frac{1}{n} \log p_n(\alpha) = \alpha \log\left[(1 - \varepsilon)\alpha\right] + (1 - \alpha) \log\left[\varepsilon(1 - \alpha)\right]. \qquad (5)$$

For n large, the variable $\frac{1}{n}\mathcal{S}^n$ is approximately normal with mean $(1 - \varepsilon)$ and variance $\frac{\varepsilon(1-\varepsilon)}{n}$, the density at α of this normal distribution is

$$f_n(\alpha) = \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{\varepsilon(1 - \varepsilon)}} \exp\left[-n\frac{(\alpha - (1 - \varepsilon))^2}{\varepsilon(1 - \varepsilon)}\right]$$

so that

$$\lim_{n\to\infty} \frac{1}{n} \log f_n(\alpha) = -\frac{(\alpha - (1 - \varepsilon))^2}{\varepsilon(1 - \varepsilon)}. \qquad (6)$$

Comparing formulas (5) and (6) for the limits of $p_n(\alpha)$ and $f_n(\alpha)$ one can see that they are different. Reasoning directly in the continuous model amounts to use (6) when one should use (5).

References

[1] Steven J. Brams and Peter C. Fishburn (1983) *Approval Voting*, Birkhauser, Boston.

[2] Steven J. Brams and Peter C. Fishburn (1985) "Comment on 'The problem of strategic voting under Approval Voting"' and "Rejoinder to Niemi" *American Political Science Review* 79: 816-819.

[3] Steven J. Brams and Peter C. Fishburn (2005) "Going from theory to practice: the mixed success of Approval Voting" *Social Choice and Welfare* 25: 457-474.

[4] Steven J. Brams and M. Remzi Sanver (2003) "Voter sovereignty and election outcome" mimeo, Istanbul Bilgi University.

[5] Francesco De Sinopoli, Bhaskar Dutta and Jean-François Laslier (2006) "Approval voting: three examples" mimeo, Ecole Polytechnique.

[6] Jean-François Laslier (2004) "Strategic Approval Voting in a large electorate" Cahier du Laboratoire d'Econométrie de l'Ecole Polytechnique #2004-001.

[7] Jean-François Laslier and Karine Van der Straeten (2006) "A live experiment on Approval Voting" *Experimental Economics* forthcoming.

[8] Roger B. Myerson (1998) "Extended Poisson games and the Condorcet jury theorem" *Games and Economic Behavior* 25: 111-131.

[9] Roger B. Myerson (2000) "Large Poisson games" *Journal of Economic Theory* 94: 7-45.

[10] Roger B. Myerson (2002) "Comparison of scoring rules in Poisson voting games" *Journal of Economic Theory* 103: 219-251.

[11] Roger B. Myerson and Robert J. Weber (1993) "A theory of voting equilibrium" *American Political Science Review* 87: 102-114.

[12] Richard G. Niemi (1984) "The problem of strategic voting under Approval Voting" *American Political Science Review* 78: 952-958.

[13] Donald G. Saari and Jill Van Newenhizen (1988) "The problem of indeterminacy in approval, multiple and truncated voting systems" and "Is approval voting and 'unmitigated evil?' : a response to Brams, Fishburn and Merrill", *Public Choice* 59: 101-120 and 133-147.

[14] Patrick Suppes, David M. Krantz, R. Duncan Luce and Amos Tversky (1989) *Foundations of Measurement. Volume II: Geometrical, Threshoold, and Probabilistic Representations.* Academic Press.

[15] Amos Tversky (1972a) "Elimination by aspects: a theory of choice" *Psychological Review* 79: 281-299.

[16] Amos Tversky (1972b) "Choice by Elimination" *Journal of Mathematical Psychology* 9: 341-367.

[17] Amos Tversky and Shmuel Sattah (1979) "Preference trees" *Psychological Review* 86: 542-573.