Toward a 50%-majority equilibrium when voters are symmetrically distributed.

Hervé Crès[*] M. Utku Ünver[†]

NYUAD *Boston College*

Abstract

Consider a two-dimensional spatial voting model. A finite number m of voters are randomly drawn from a (weakly) symmetric distribution centered at O. We compute the exact probabilities of all possible Simpson-Kramer scores of O. The computations are independent of the shape of the distribution. The resulting expected score of O is used as a proxy for an upper-bound to the min-max score.

Keywords: Spatial voting, super majority, random point set, stochastic geometry.

JEL Codes: D70, D71

[*]New York University, Abu Dhabi Saadiyat campus, PoBox 129188, Abu Dhabi, United Arab Emirates; herve.cres@nyu.edu

[†]Boston College, Department of Economics, 140 Commonwealth Ave., Chestnut Hill, MA, 02467 USA; unver@bc.edu

Introduction

Consider a society comprising a potentially small number m of individuals, what is the smallest rate of (super) majority for which there exists an equilibrium? This question is important for anyone interested in the governance of such institutions as corporations or partnership, political parties, associations etc., if only to optimize the design of the institution's charter or constitution.

Let us be more precise. We consider a society $\mathcal{I}$ of m $(= \sharp\, \mathcal{I})$ individuals having preferences over a set of alternatives $\mathcal{X}$. An alternative $x \in \mathcal{X}$ is an equilibrium for the rate $\rho \in [0, 1]$ if there is no alternative $x' \in \mathcal{X}$ that rallies the vote of more than ρm individuals against x. Let ρ^* denote the smallest rate for which there exists an equilibrium, called the min-max rate of society $\mathcal{I}$ (Simpson, 1969; Kramer, 1973).

The social choice literature offers several upper bounds to ρ^*. Black's (1953) median voter theorem states that if alternatives can be ordered along a one-dimensional space, and preferences are single-peaked over this space, then $\rho^* \leq 0.5$: there is an equilibrium for the (simple) majority rule. However, we know since the seminal work of Plott (1967) that this does not extend to higher dimensions: McKelvey and Wendell (1976), McKelvey (1979), Rubinstein (1979), Schofield (1983), McKelvey and Schofield (1987), Banks (1995), Saari (1997), and Banks et al. (2006) have shown that the set of configurations for which $\rho^* \leq 0.5$ has measure 0 in three or more dimensions, and in two dimensions when the number of voters is odd. Tovey (2010a) completes the study of the two-dimensional case by proving that, when the number of voters is even and voters are sampled i.i.d. from any nonsingular distribution, the measure of the latter set converges to 0 exponentially rapidly.

Among other important contributions, Greenberg (1979) shows that if $\mathcal{X}$ is a convex and compact set of dimension d, and individuals have convex and continuous preferences over $\mathcal{X}$, then $\rho^* \leq 1 - \frac{1}{d+1}$. Caplin & Nalebuff (1988) show that when voters have Euclidian preferences, and the distribution of their ideal points is concave over a convex support, then $\rho^* \leq 0.64$. One of the beautiful attributes of the latter result (compared with Greenberg's) is that it is independent of the dimension of the space of alternatives, another one is that the super majority rate is not too high. One of the strengths of Greenberg's result is that it is independent of the distribution of ideal points.

Our study is geared toward societies with potentially small number of individuals. As an illustration, consider $m = 4$ and $d = 2$. The set of alternatives is the plane: $\mathcal{X} = \mathbb{R}^2$ and each voter $i \in \{1, 2, 3, 4\}$ has a most preferred alternative $x_i \in \mathbb{R}^2$, with (Euclidian) indifference curves being circles around x_i. There can be only two geometric configurations[1], depending

[1]In the Educational Times of 1864, J. J. Sylvester proposed what became known as his four-point problem: what is the chance of a reentrant quadrilateral when the four points are taken at random in an indefinite plane. (This problem laid the ground to the field of stochastic geometry.) Sylvester himself admitted that

on whether the quadrilateral that they form is convex or reentrant, as Figure 1 shows.

Figure 1.a   Figure 1.b

As far as the opening question is concerned, the two configurations lead to the same result: $\rho^* \leq 0.5$. In case of a convex quadrilateral (see Figure 1.a), the point at the intersection between the two diagonals ($[x_1, x_3]$ and $[x_2, x_4]$) is stable with respect to a 0.5-majority rule. In case of a reentrant quadilateral, the point in the convex hull of the three others (x_4 in Figure 1.b) is also stable with respect to a 0.5-majority rule.[2]

What more general can be said? The problem being highly complex, we restrict ourselves to a two-dimensional space of issues and moreover we consider that voters' ideal points are selected independently from a *nonsingular* distribution f over $\mathbb{R}^2$ (the mass of f on any line in $\mathbb{R}^2$ is 0), which is furthermore *sign-invariant* (a weak symmetry property): for all $x \in \mathbb{R}^2$, $f(x) = f(-x)$; as in Tovey (1992). A first step is to compute the minimum rate $\rho(O)$ for which O is stable. Since $\rho^* \leq \rho(O)$, we then obtain an additional upper bound to ρ^*.

For a given m-sample, there is no guarantee that O be stable for any rate below 1.[3] So our strategy is to compute the exact probability of the whole range of scores, and consequently the expected value of $\rho(O)$. A strength of our contribution is borrowed from a way, due to Wendel (1962), to account for m-samples drawn from a sign-invariant distribution; hence the probabilities of various scores for O are independent of the shape of the distribution − as long as it remains sign-invariant.

this problem has no determinate solutions, as it was ill-posed, since there is no natural probability measure on the plane (see the historical review of the problem in Pfiefer, 1989). The problem was changed into the following: Choose four points at random (independently and uniformly) from K, where K is a convex set in the plane; what is the probability, $\pi(K)$, that their convex hull is a triangle? Blaschke (1917) showed that for all convex compact bodies $K \subset \mathbb{R}^2$, $\pi(\text{disk}) \leq \pi(K) \leq \pi(\text{triangle})$. (See Barany (2001) for a generalization to higher dimensions.)

[2] On a side note, consider the case of m voters and $d = 2$. Assume that their convex hull is ℓ-lateral, with $\ell \leq m$. We conjecture that the set of alternatives that are stable for the min-max rate is also ℓ-lateral.

[3] E.g., the probability that there exists an alternative *unanimously* preferred to O is $m/2^{m-1}$ (see below).

The Model

There are d measurable criteria of social choice, so that a social alternative can be represented as a d-dimensional vector: $x \in \mathbb{R}^d$. There are m voters in a set $\mathcal{I}$. Each voter is endowed with a *Euclidean preference* relation on $\mathbb{R}^d$: voter i, $1 \leq i \leq m$, has a *ideal point* in the space of social choice, $x_i \in \mathbb{R}^n$, and his/her utility function over the space of social choices is decreasing with the Euclidean distance from his/her ideal point:

$$\forall x \in \mathbb{R}^d \qquad u_i(x) = -\|x_i - x\|$$

A *society* is a m-tuple $X = (x_i)_{i=1}^m$.

We measure the stability of an alternative in a given society through the Simpson-Kramer approach (Simpson, 1969; Kramer, 1973). Given two alternatives $(a, b) \in \mathbb{R}^n \times \mathbb{R}^n$, $\rho(b, a)$ measures the ratio of the electorate that strictly prefers b to a:

$$\rho(b, a) = \frac{\sharp\{i \in \mathcal{I} | u_i(b) > u_i(a)\}}{m}.$$

The *score* of an alternative $a \in \mathbb{R}^n$ is: $\rho(a) = \max_{b \in \mathbb{R}^n} \rho(b, a)$.

The *majority rule* with rate $\rho \in [0, 1]$ states that candidate b is preferred by society X to (or defeats) candidate a if and only if $\rho(b, a) > \rho$. A candidate a is a ρ-*majority equilibrium* in society X if and only if there is no alternative that defeats him/her, i.e., if and only if its score is not larger than ρ: $\rho(a) \leq \rho$.

The case of $d = 1$ is trivial. We know from Black's (1953) median voter theorem that when m is odd (resp., even), the median ideal point is an equilibrium (resp., all points in the segment between the lower and upper median ideal points are equilibria) for the majority rule with rate 0.5. and the second term on the right-hand side is 1/2. We turn now to the case of $d = 2$.

The probabilities of all possible scores of O

Consider the following process which is due to Wendel (1962): choose m random points in a disk centered at O: Q_1, Q_2, ..., Q_m. For each i, $1 \leq i \leq m$, we set P_i equal to Q_i or to $-Q_i$ with equal probability 1/2 (without loss of generality, we can choose the Q_i's on the same side of a hyperplane through O as in Figure 2.a below; Figure 2.b corresponds to the configuration: $P_i = Q_i$ for $i = 1, 3, 4$ and $P_i = -Q_i$ for $i = 2, 5$). The points P_1, ..., P_m are again i.i.d. random points in the disk.

The original question answered by Wendel (1962), see also Wagner & Welzl (2001), is: what is the probability that O is not in the convex hull of the P_i's? (In other words, what is the probability that the score of O be 1?) The answer is: $m/2^{m-1}$. Indeed, independently of the choice of the Q_i's (again, we ignore the degenerate configurations, occurring with probability zero, where two vectors Q_i and Q_j would be collinear), there are $2m$ possibilities to choose the signs of the P_i's such that O can be separated from these points by a line (every partition of the Q_i's by a line through O gives two such possibilities). Again, all we used in this line of reasoning was that the original distribution is symmetric about O and that some degeneracies occur with probability zero. Hence the result is independent of the shape of the distribution.

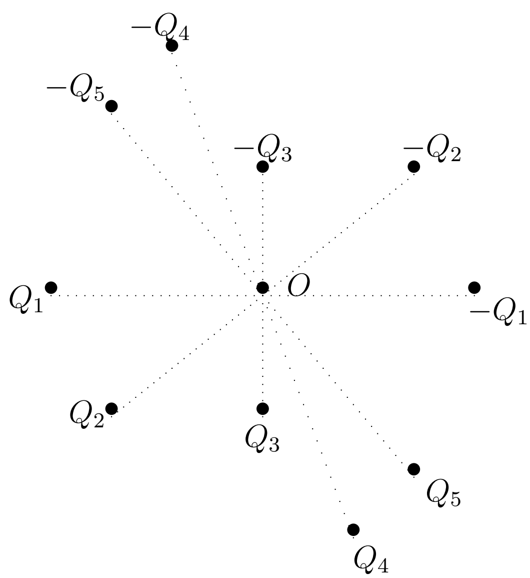

Figure 2.a

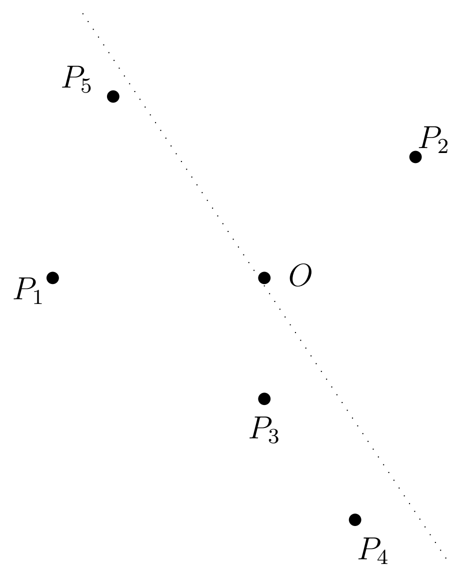

Figure 2.b

We now proceed along this line of reasoning and compute the probability that the score of O be j/m. E.g., in the social choice configuration shown on Figure 2.b, the score of O is 4/5 as shown by the dotted separation line.

As it was already noticed by Tovey (1992, Lemma 1), each social choice configuration can be described as a (q, p)-*sequence* (or random walk) of plus ones and minus ones, according to whether P_i is equal to Q_i (then +) or not (then -), $1 \leq i \leq m$: $\epsilon_1, \ldots, \epsilon_m$, with, say, q plus ones and p minus ones, $q + p = m$. The configuration on Figure 2.b corresponds to the (3,2)-sequence: $+ - + + -$. The partial sum $s_k = \epsilon_1 + \ldots + \epsilon_k$ represents the difference between the number of pluses and minuses occurring at the first k places, $0 \leq k \leq m$, with $s_0 = 0$ and $s_m = q - p$. Define: $\bar{s} = \max_k s_k$ and $\underline{s} = \min_k s_k$.

Lemma 1 *In the social choice configuration represented by a (q, p)-sequence $(\epsilon_1, \ldots, \epsilon_m)$, the score of O is* $\rho(O) = \dfrac{\max\{q - \underline{s}, p + \bar{s}\}}{m}$.

Proof of Lemma 1: Consider a line which separates the $+Q_i$'s from the $-Q_i$'s and passing through O. It has q of the P_i's on one side and p on the other side. Now turn this line by pivoting at O so that it goes in-between Q_1 and Q_2: it has now $q - s_1$ of the P_i's on one side and $p + s_1$ on the other side. Now turn it by pivoting at O so that it goes in-between Q_2 and Q_3: it has now $q - s_2$ of the P_i's on one side and $p + s_2$ on the other side. And so on. The maximum number of P_i's on one side of a line through O is therefore $\max\{\ldots, q - s_k, \ldots, p + s_k, \ldots\} = \max\{q - \underline{s}, p + \bar{s}\}$. Hence the result. □

To compute the probability that the score of O be j/m, we need to compute the number of (q, p)-sequences such that $\max\{q-\underline{s}, p+\bar{s}\} = j$. To do that, we follow a classical geometric method in the standard orthonormal basis where the x-axis in horizontal and the y-axis is vertical. Following Feller (1968), the sequence $(\epsilon_1, \ldots, \epsilon_m)$ is identified with a *path* from the origin to the point $(m, q - p)$: this path is a polygonal line whose vertices have abscissa $0, 1, \ldots, m$ and ordinates $s_0, s_1, \ldots, s_m = q - p$; $\bar{s}$ is the highest point of the path and $\underline{s}$ the lowest. Obviously there are[4] $\binom{m}{p}$ such paths from the origin to the point $(m, q - p)$: as many as there are ways of choosing the p places for the minuses out of the m possibilities.

Note that for a (q, p)-sequence, $\bar{s} \geq \max\{0, q - p\}$ and $\underline{s} \leq \min\{0, q - p\}$ entail that $\max\{q - \underline{s}, p + \bar{s}\} \geq \max\{q, p\}$, therefore we restrict attention to $j \geq \max\{q, p\}$. And in the case when m is even and $p = q = m/2$, obviously $\max\{q - \underline{s}, p + \bar{s}\} \geq m/2 + 1$ therefore we restrict attention in that case to $j \geq m/2 + 1$. Hence we consider j such that $[m/2] + 1 \leq j \leq m$.

A (q, p)-sequence is such that $\max\{q-\underline{s}, p+\bar{s}\} = j$ whenever the associated path remains in the corridor between the lines $y = j - p$ and $y = q - j$, and hits at least one of them (see Figure 3 drawn for the configuration of Figure 2.b and $j = 4$).

Lemma 2 *Fix j, $[m/2]+1 \leq j \leq m$. The number of (q, p)-paths such that* $\max\{q-\underline{s}, p+\bar{s}\} =$ *j is*

$$a_{m,q,j} = \begin{cases} A_{m,q,j+1} - A_{m,q,j} & \text{if } m - j \leq q \leq j \\ 0 & \text{otherwise} \end{cases}$$

where for $m - j \leq q \leq j$

$$A_{m,q,j} = \sum_k \left[\binom{m}{q + k(2j - m)} - \binom{m}{j + k(2j - m)} \right] \qquad (1)$$

(the series extending over all integers k from $-\infty$ to $+\infty$, but having only finitely many non-zero terms) is the number of (q, p)-paths such that $\max\{q - \underline{s}, p + \bar{s}\} < j$*: those which hit*

[4] By convention, the combination number will be set to zero in case $p < 0$ or $p > m$.

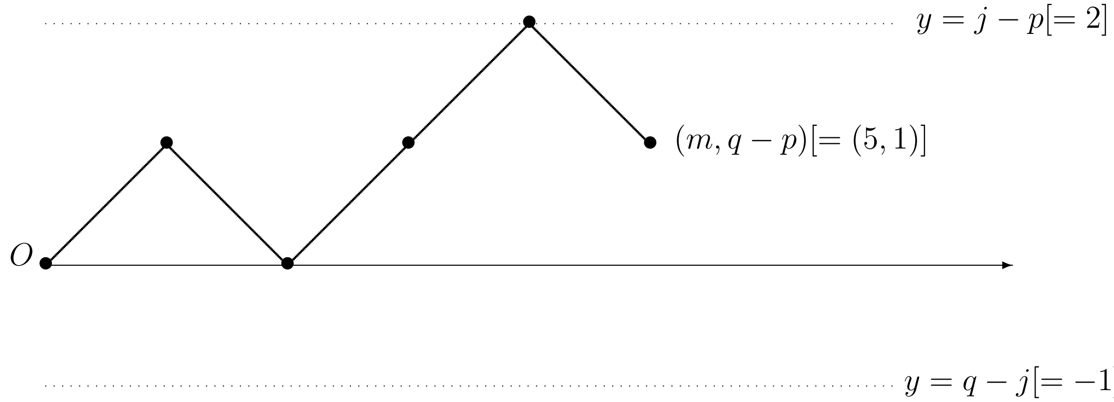

Figure 3

neither $y = j - p$ *nor* $y = q - j$*, and the number of* (q, p) *paths such that* $\max\{q - \underline{s}, p + \bar{s}\} < m + 1$ *is given by*

$$A_{m,q,m+1} = \binom{m}{q}. \tag{2}$$

Proof of Lemma 2: The equation relating the $a_{m,q,j}$'s to the $A_{m,q,j}$'s is immediate.

The numbers $A_{m,q,j}$ of (q, p)-paths such that $\max\{q - \underline{s}, p + \bar{s}\} < j$ remain to be computed. These computations are based on the reflection principle (see, e.g., Feller (1968), Chapter III).

Let $N(m, c) = \binom{m}{\frac{m-c}{2}} = N(m, -c)$ denote the number of paths from $O = (0, 0)$ to (m, c). Let a and b be positive, and $-b < c < a$. By the reflection principle, the number of paths from $(0, 0)$ to (m, c) which touch or cross $y = a$ is equal to the number of paths from $(0, 2a)$ (the reflection of O on the axis $y = a$) to (m, c), i.e., $N(m, 2a - c)$. By the same argument, the number from $(0, 0)$ to (m, c) which touch or cross $y = -b$ is $N(m, c + 2b) = N(m, 2a - c - 2(a + b))$.

Now, by a double application of the reflection principle, a path from $(0, 0)$ to (m, c) which touch or cross $y = a$ and then $y = -b$ (called an '(ab)' path in the sequel) can be first associated to a path from $(0, 2a)$ to (m, c), itself associated to a path from $(0, -2a - 2b)$ to (m, c); hence $N(m, c + 2(a + b))$ of 'ab' paths. A triple application allows through the same line of argument to count the paths which touch or cross $y = a$, then $y = -b$, then $y = a$ again ('$(ab)a$' paths); their number is $N(m, 2a - c + 2(a + b))$. An extension of this method gives:

- $N(m, c + 2k(a + b))$ for the number of paths which touch or cross $y = a$ and then $y = -b$ k times in a row ('$k(ab)$' paths);

- $N(m, 2a - c + 2k(a + b))$ for the number of paths which touch or cross $y = a$ and then $y = -b$ k times in a row and then $y = a$ again ('$k(ab)a$' paths);

- $N(m, c - 2k(a + b))$ for the number of paths which touch or cross $y = -b$ and then $y = a$ k times in a row ('$k(ba)$' paths);

- $N(m, 2a - c - 2(k + 1)(a + b))$ for the number of paths which touch or cross $y = -b$ and then $y = a$ k times in a row and then $y = -b$ again ('$k(ba)b$' paths).

Our aim is to compute the number of paths from $(0, 0)$ to (m, c) which touch or cross neither $y = a$ nor $y = -b$. This comes first by exclusion of paths which touch or cross $y = a$ and paths which touch or cross $y = -b$. But thus '(ab)' and'(ba)' paths are excluded twice and must be re-included once. But then '$(ab)a$' and'$(ba)b$' are excluded twice, then re-included twice, and therefore must be re-excluded once... This standard application of the *inclusion-exclusion principle* (see Comtet, 1974, Chapter IV), leads to the formula:

$$\begin{aligned} N(m, c) &- N(m, 2a - c) + \sum_{k>0} [N(m, c + 2k(a + b)) - N(m, 2a - c + 2k(a + b))] \\ &- \sum_{k>0} [N(m, c - 2k(a + b)) - N(m, 2a - c - 2k(a + b))] \end{aligned}$$

for the concerned number, which can be rewritten:

$$\sum_k [N(m, c + 2k(a + b)) - N(m, 2a - c + 2k(a + b))]$$

(over all integers k from $-\infty$ to $+\infty$, but only finitely many non-zero terms). The formula 1 obtains readily by substitution of the right parameters. ■

Proposition 1 *Consider an m-sample independently drawn from a sign-invariant distribution; the probability that the score of O be j/m, $[m/2] + 1 \leq j \leq m$ is*

$$\bar{a}_{m,j} = \frac{1}{2^m} \sum_{q=m-j}^{j} a_{m,q,j}$$

and the expected value of the score of O is

$$E\rho(O) = \sum_{j=[m/2]+1}^{m} \frac{j \bar{a}_{m,j}}{m} .$$

Proof of Proposition 1: Immediately follows from Lemma 2. ■

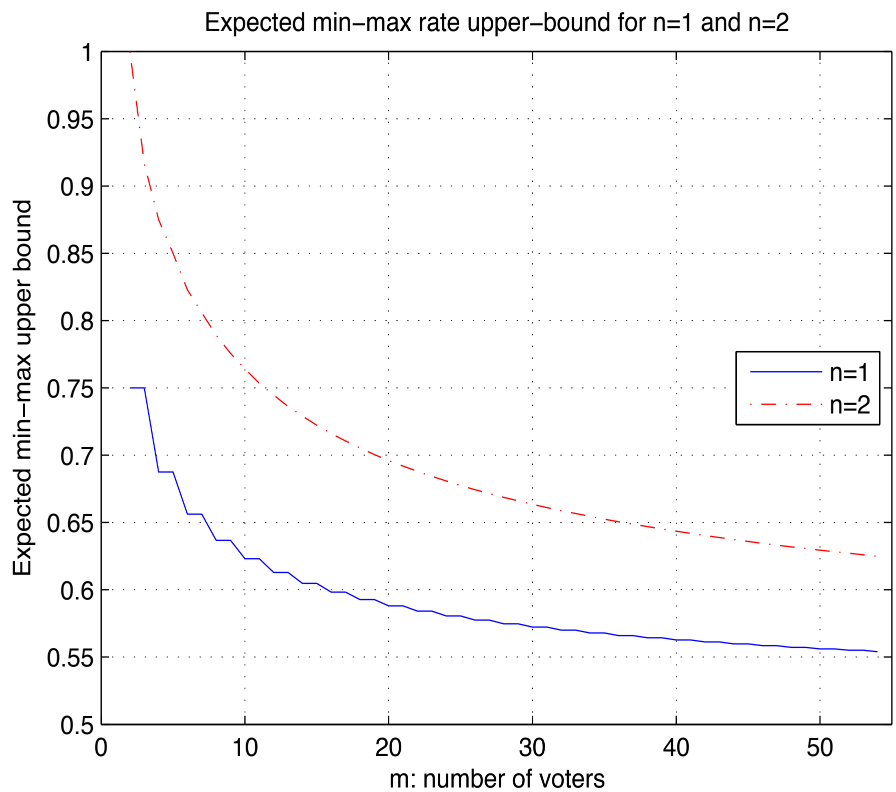

Figure 4

We plot in Figure 4 the expected score of O as a function of the number of voters m. For the sake of benchmarking, we also plot the expected score of O for $d = 1$ − whose exact value is

$$E\rho(O) = \frac{1}{2} + \frac{1}{2^{2\left[\frac{m}{2}\right]+1}} \begin{pmatrix} 2\left[\frac{m}{2}\right] \\ \left[\frac{m}{2}\right] \end{pmatrix}.$$

Obviously, when $m \to \infty$, $\rho(O) \to 0.5$: the score of O converges to $1/2$ − even though we know (see e.g. Tovey, 2010b) that the probability is 0 that O is be equilibrium for the simple majority rule. Figure 4 gives a hint of the speed of convergence toward this lower bound.

Let us compare the present results to the literature. Among other things, Tovey (1992, Theorem 1) computes $\bar{a}_{m,j} = 1/2^{m-2}$ for $j = [m/2] + 1$ when m is odd. The present paper generalizes the latter by giving exact probabilities for the whole range of scores and for all m. Another strategy is to look for asymptotic results: e.g., Schofield & Tovey (1992) shows, for any $d > 1$, the limit of $\bar{a}_{m,j}$ when $m \longrightarrow \infty$ is 0 for $j \geq m/2 + \sqrt{md \log m}$ (when f is weakly centered); although it yields some results for all dimensions d, it cannot be used for small committees.

As a word of conclusion, it is clear that, although the distribution is sign-invariant, it might be that the score of O is a poor upper-bound of our real target, which is the minimum score. Finding the exact result, or at least an improved upper-bound is certainly an object of further research.

References

Banks, J.S. (1995), Singularity theory and core existence in the spatial model. *Journal of Mathematical Economics* **24**, 523-536.

Banks, J.S., Duggan, J. & M. Le Breton (2006), Social choice and electoral competition in the general spatial model. *Journal of Economic Theory* **126**, 194-234.

Barany, I. (2001), A note of Sylvester's four-point theorem. *Studia Scientiarum Mathematicarum Hungarica* **38**, 73-77.

Black, D. (1953), *The theory of committees and elections.* London: Cambridge University Press.

Blaschke, W. (1917), Uber affine Geometrie XI: Losung des "Vierpunktproblems" von Sylvester aus der Theorie der geometrischen Wahrscheinlichkeiten. *Leipz. Ber.* **69,** 436-453.

Caplin A., & B. Nalebuff (1988), On 64%-majority rule. *Econometrica*, **56**, 787-814.

Comtet, L. (1974), *Advanced Combinatorics.* Boston: Reidel 1974.

Feller, W. (1968), *An introduction to probability theory and its applications.* John Wiley and Sons, New York.

Greenberg, J. (1979), Consistent majority rules over compact sets of alternatives. *Econometrica* **47**, 627-636.

Kramer, G.H. (1973), On a class of equilibrium conditions for majority rule. *Econometrica* **41**, 285-297.

McKelvey, R.D. (1979), General conditions for global intransitivies in formal voting models. *Econometrica* **47**, 1085-1112.

McKelvey, R.D. & R.E. Wendell (1976), Voting equilibria in multidimensional choice spaces. *Mathematics of Operations Research*, **1**, 144-158.

McKelvey, R.D. & N.J. Schofield (1987), Generalized symmetry conditions at a core point. *Econometrica,* **55**, 923-933.

Pfiefer, R.E. (1989), The Historical Development of J. J. Sylvester's Four Point Problem. *Mathematics Magazine,* **62**, 309-317.

Plott, C. (1967) A notion of equilibrium and its possibility under majority rule. *American Economic Review*, **57**, 787-806.

Rubinstein, A. (1979), A note about the 'nowhere denseness' of societies having an equilibrium under majority rule. *Econometrica,* **47**, 511-514.

Saari, D. (1997), The generic existence of a core for q-rules. *Economic Theory,* **9**, 219-260.

Schofield, N. (1983), Generic instability of majority rule. *Review Economic Studies*, **50**, 696-705.

Schofield, N. & C. Tovey (1992), Probability and convergence for supra-majority rule with Euclidean preferences. *Mathematical and Computer Modelling,* **16**, 41-58.

Simpson, P.B. (1969), On defining areas of voter choice, *Quarterly Journal of Economics*, **83**, 478-490.

Sylvester, J. J. (1864), *The Educational Times.* London.

Tovey, C. (1992), The probability of an undomiated central voter in 2-dimensional spatial majority voting. *Social Choice and Welfare*, **9**, 43-48.

Tovey, C. (2010a), The probability of majority rule instability in the 2D euclidian model with an even number of voters. *Social Choice and Welfare*, **35**, 705-708.

Tovey, C. (2010b), A Critique of Distributional Analysis in the Spatial Model. *Mathematical Social Sciences,* **59**, 88-101.

Wagner, U. & E. Welzl (2001), A continuous analogue of the upper bound theorem. *Discrete Computational Geometry*, **26**, 205-219.

Wendel, J.G. (1962), A problem in geometric probability. *Math. Scand.*, **11**, 109-111.