Statistical utilitarianism

Pivato, Marcus

Trent University, Department of Mathematics

6 September 2013

Statistical utilitarianism

Marcus Pivato, Trent University

September 6, 2013

Abstract

We show that, in a sufficiently large population satisfying certain statistical regularities, it is often possible to accurately estimate the utilitarian social welfare function, even if we only have very noisy data about individual utility functions and interpersonal utility comparisons. In particular, we show that it is often possible to identify an optimal or close-to-optimal utilitarian social choice using voting rules such as the Borda rule, approval voting, relative utilitarianism, or any Condorcet-consistent rule.

Keywords: utilitarian; relative utilitarian; approval voting; Borda; scoring rule; Condorcet.
JEL class: D63; D71.

1 Introduction

Utilitarianism may be philosophically attractive, but as a practical method for making social decisions, it faces at least four major problems.

(Pr1) Interpersonal comparisons of utility are problematic. Even if we accept that such interpersonal comparisons are meaningful in principle, it is not clear how precise interpersonal comparisons could be made in practice.

(Pr2) It is difficult for the social planner to obtain accurate information about the voters' utility functions. (It is not generally feasible to obtain a precise utility assessment from every voter for every possible social alternative.)

(Pr3) Due to epistemic failures, a voter may have incorrect beliefs about the long-term consequences of various policy alternatives. Furthermore, people fail to correctly predict their own future utility level, even in apparently straightforward decision problems (Loewenstein and Schkade, 1999). Indeed, there is ample empirical evidence that people's beliefs about their own past, present, and future happiness are surprisingly unreliable, and subject to systematic biases, errors, and illusions (Kahneman et al., 1999). In short: voters do not even correctly perceive their *own* utility functions.

(Pr4) Voters may strategically misrepresent their utility functions (e.g. exaggerate the intensity of their preferences) to manipulate a utilitarian social planner.

However, in this paper, we will show that these problems almost disappear in large populations of voters satisfying certain mild statistical assumptions. In Section 2, we show that averaging utility data (even noisy or miscalibrated data) from a large population of voters will yield a good approximation of utilitarianism with high probability, despite problems (Pr1)-(Pr3). In particular, in Section 3, we show that the relative utilitarian social choice rule yields such an approximation. In Section 4, we show that, if the voters' utility functions are related to their ordinal preferences through a plausible stochastic model, then scoring rules (such as the Borda rule) will yield a good approximation of utilitarianism with high probability, when the population is large. In Section 5, we provide a similar result for approval voting. In Section 6, we first observe that, under a weak assumption (called "reasonability") about the statistical distribution of voter's preferences, the Condorcet winner will be the utilitarian social choice. We then show that reasonability holds for several plausible statistical models of voter preferences. Finally, in the conclusion, we briefly discuss how, for the perspective taken in this paper, problem (Pr4) can be solved through *virtual implementation*.

Related literature. Lerner (1944, pp.29-32) was perhaps the first to deploy statistical aggregation to obviate the technical difficulties of utilitarianism. Under the plausible assumption that all agents have diminishing marginal utility for wealth, Lerner famously argued that, even in a state of *total ignorance* about the precise structure of people's cardinal utility functions, an egalitarian distribution of wealth would maximize the *expected* aggregate utility for society, because the expected utility gains of the poor under such a wealth redistribution would more than cancel the expected utility losses of the rich. Lerner's original argument was obscure and generated much confusion; it was later clarified by Breit and Culbertson Jr. (1970).[1] For Lerner's argument to work, his vague hypothesis of "total ignorance" about individual utility functions must be formalized in terms of quite specific assumptions about the probability distribution of the utility functions. The results of Section 2 can be seen as an extension of this approach (although we draw no conclusions about wealth redistribution).

Later, Bordley (1983) and Merrill (1984) used computer simulations to estimate the expected value of the utilitarian social welfare function for various voting rules. More recently, Caragiannis and Procaccia (2011) estimated the "distortion" of the plurality, approval, and antiplurality voting rules —that is, the worst-case ratio between the utilitarian social welfare of the optimal alternative, and the utilitarian social welfare of the alternative which actually wins, where the worst case is computed over all possible profiles of "normalized" utility functions. (A utility function is "normalized" if it is positive and the utilities sum to one.) Procaccia and Caragiannis were particularly interested in the asymptotic growth rate of this distortion ratio as the number of voters and/or alternatives becomes

[1] See also the responses by Lerner (1970), Breit and Culbertson Jr. (1972), McCain (1972), and McManus et al. (1972).

large. They showed that, if voters randomly convert their cardinal utility functions into voting behaviours in a plausible way, then the expected distortion ratio grows surprisingly slowly. Their intended application was preference aggregation in a cooperating group of artificially intelligent agents (e.g. Mars rovers), but their results are also applicable to more traditional social welfare problems.

Meanwhile Apesteguia et al. (2011) considered scoring rules (e.g. the plurality, antiplurality, and Borda rules) in a setting with three alternatives and honest voters; they investigated the efficiency of these rules with respect to a one-parameter family of social welfare functions (SWFs) which includes the utilitarian, maximin and maximax SWFs. Their analytical results focus mainly on societies with *two* voters, whose utilities for the three alternatives are independent, identically distributed (i.i.d.) random variables drawn from a one-parameter family of probability distributions on the interval $[0, 1]$ (including the uniform distribution). They also presented numerical simulation results for somewhat larger populations. Based on these simulations, they conjectured that, in the large population limit, the Borda rule is the optimal scoring rule with respect to the utilitarian SWF. We verify this conjecture in Section 4 (see Corollary 5).

Giles and Postl (2012) have conducted a similar investigation for (A, B)*-voting rules*, a two-parameter family of rules introduced by Myerson (2002), which includes approval vote as well as all scoring rules. Like Apesteguia et al., Giles and Postl suppose there are three alternatives, whose utilities for each voter are privately known, i.i.d. random variables on the interval $[0, 1]$. But unlike Apesteguia et al., they focus only on the utilitarian SWF, and they allow for strategic voting. Giles and Postl first analytically characterize the symmetric Bayesian Nash equilibrium (BNE) for the N-player strategic voting game for any $N \geq 2$. Then they numerically compute the expected value of the utilitarian SWF at the three-player BNE for various $(A, B) \in [0, 1]^2$ (where the three players' utilities are i.i.d. random variables drawn from either a uniform distribution or a beta distribution on $[0, 1]$). In a very similar environment (three alternatives, independent uniformly distributed random utilities, but many players), Kim (2012) constructs incentive-compatible voting rules which, in terms of the utilitarian SWF, are superior to any ordinal rule (in particular, any scoring rule), but which utilize only a limited amount of cardinal utility information from the voters.

Like Kim (2012) and Giles and Postl (2012), this paper focuses only on the utilitarian SWF, and considers a larger class of voting rules than Apesteguia et al. (2011). However, unlike any of these three other papers, we work with an arbitrarily large set of alternatives, and a very wide variety of statistical models of voter preferences. Instead of two- or three-voter societies, our analytic results concern the asymptotic behaviour of societies with hundreds or thousands of voters. Also, while these other papers are mainly concerned with showing that some rules are better than others, or with finding the best rule within a certain class (in terms of the expected value of the utilitarian SWF), we show that, under certain conditions, certain rules actually approach *perfect* agreement with utilitarianism, as the population becomes large. Finally, while the aforementioned papers suppose that cardinal interpersonal utility comparisons are unproblematic, we allow the possibility that these interpersonal comparisons are themselves highly ambiguous in practice (although still meaningful in principle). The results on scoring rules in Section 4 can thus be seen as complementary to the results of Apesteguia et al. (2011), while the results on approval voting in Section 5 are complementary to the findings of Giles and Postl (2012).

In a setting with only two alternatives, Schmitz and Tröger (2012) have shown that "weak" majority voting rules yield the highest expected value for the utilitarian SWF amongst all dominant-strategy rules.[2] They also review a series of much earlier papers starting with Rae (1969), which investigated the utilitarian efficiency of qualified majority rules in a two-alternative setting. In a similar spirit, in Section 6, we note that the Condorcet winner will always maximize the utilitarian SWF for any "reasonable" profile of utility functions, and any number of alternatives. We then show that such a "reasonable" profile will occur with very high probability in large populations of voters with random utility functions drawn from certain classes of probability distributions.

This paper presents asymptotic probabilistic results for large populations. It is not possible here to adequately summarise the vast and growing literature on the large-population asymptotic probabilistic analysis of voting rules. Instead, we will only briefly touch on two strands of this literature. The first strand is the Condorcet Jury Theorem (CJT) and its many generalizations.[3] Like the CJT literature, the results of the present paper say that, under certain statistical assumptions, a large population using a certain voting rule is likely to arrive at the "correct" decision. But the goal for the CJT literature is to find the objectively correct answer to some factual question, whereas the goal in the present paper is to maximize social welfare.

The second strand is the literature on strategic voting and/or strategic candidacy in large populations with some kind of randomness or uncertainty in voters' preferences. This literature is mainly concerned with characterizing the Nash equilibria of certain large election games. These equilibria occasionally have surprising social welfare properties. For example, Ledyard (1984), Lindbeck and Weibull (1987, 1993), Coughlin (1992; Theorem 3.7 and Corollary 4.4), Banks and Duggan (2004; §4) and McKelvey and Patty (2006) have all shown that, in certain election games, there is a unique Nash equilibrium (sometimes called a "political equilibrium") where all the candidates select the policy which maximizes a utilitarian SWF. But these utilitarian SWFs are based on somewhat peculiar systems of interpersonal utility comparisons. In these models, voter behaviour is described by a stochastic device: the probability that voter i votes for candidate C (or in some cases, the probability that i votes at all) is a function of the difference between the cardinal utility which i assigns to C and the cardinal utility she assigns to other candidates. Although the different models use different stochastic devices and seek to capture different phenomena (e.g. random private costs for voting, or random private shocks to the utility functions, or random individual errors due to bounded rationality, or other exogenous perturbations),

[2]Strictly speaking, Schmitz and Troger show that these rules yield the highest *ex ante* expected utility for each individual *voter*, before she learns her true preferences. However, in their model, all voters have identically distributed random utility functions *ex ante*, so this is equivalent to maximizing the expected value of the utilitarian SWF, and also equivalent to *ex ante* Pareto efficiency, as Schmitz and Troger observe in footnote 13 of their paper.

[3]See Nitzan (2009, Ch.11-12) or Pivato (2013) for surveys of this literature.

each model assumes that utility functions are translated into voting probabilities *in the same way* for every voter. In this way, each model smuggles in a system of "implicit" interpersonal utility comparisons via the stochastic device. As observed by Banks and Duggan (2004, p.29), this makes the normative significance of the "utilitarianism" emerging from these political equilibria somewhat unclear.

In contrast, this paper assumes that there is a pre-existing, normatively meaningful system of cardinal interpersonal utility comparisions, explicitly described by a set of "calibration constants" which exist independently of the voting rule and any other random factors in the model. The social planner does not know the exact values of these calibration constants, but regards them as random variables; the main result of Section 2 states that it is still possible to closely approximate the utilitarian social choice, even with this kind of uncertainty. On the other hand, unlike the political equilibrium literature described in the previous paragraph, this paper does not grapple with strategic issues, except in the conclusion. Also, unlike the political equilibrium literature, this paper treats the social alternatives as exogenous, rather than endogenizing them as the result of political candidates competing for popularity.

2 Statistical interpersonal comparisons

First we will fix some notation which will be maintained throughout the paper. Let $\mathbb{R}$ denote the set of real numbers, and let $\mathbb{R}_+$ be the set of positive real numbers. Let $\mathcal{A}$ denote a set of social alternatives (either finite or infinite), let $\mathcal{I}$ be a set of voters, and let $I := |\mathcal{I}|$. (We will typically suppose that I is very large.) For every i in $\mathcal{I}$, let $u_i : \mathcal{A} \longrightarrow \mathbb{R}$ be a cardinal utility function for voter i, and let $c_i > 0$ be a "calibration constant", which we will use to make cardinal interpersonal utility comparisons. We suppose that the functions $c_i\, u_i$ and $c_j\, u_j$ are interpersonally comparable for all voters i and j in $\mathcal{I}$. In other words, for any alternatives a, b, c, and d in $\mathcal{A}$, if $c_i\, u_i(b) - c_i\, u_i(a) = c_j\, u_j(d) - c_j\, u_j(c)$, then the welfare that voter i gains in moving from alternative a to alternative b exactly the same as the welfare that voter j gains in moving from c to d. We would therefore like to maximize the utilitarian social welfare function $U_{\mathcal{I}} : \mathcal{A} \longrightarrow \mathbb{R}$ defined by

$$U_{\mathcal{I}}(a) \quad := \quad \frac{1}{I} \sum_{i \in \mathcal{I}} c_i\, u_i(a), \qquad \text{for every alternative } a \text{ in } \mathcal{A}. \tag{1}$$

Let $\operatorname{argmax}_{\mathcal{A}}(U_{\mathcal{I}})$ denote the set of elements of $\mathcal{A}$ which maximize $U_{\mathcal{I}}$ —we will refer to these as *utilitarian optima*. A utilitarian social planner wants to find a utilitarian optimum, but she may not have enough information to do this, because of the aforementioned problems (Pr1)-(Pr3). We can formalize these informational problems with two assumptions:

(U1) The interpersonal calibration constants $\{c_i\}_{i \in \mathcal{I}}$ are unknown. We regard $\{c_i\}_{i \in \mathcal{I}}$ as random variables taking values in $\mathbb{R}_+$, which are independent, but not necessarily identically distributed. We assume there is some constant $\sigma_c^2 > 0$ such that, for every voter i in $\mathcal{I}$, the random variable c_i has a variance less than σ_c^2, and an expected value of 1.[4]

(U2) The utility functions $\{u_i\}_{i\in\mathcal{I}}$ are not precisely observable. Instead, for each i in $\mathcal{I}$, we can only observe a function $v_i := u_i + \epsilon_i$, where $\epsilon_i : \mathcal{A}\longrightarrow\mathbb{R}$ is a random "error" term. For each alternative a in $\mathcal{A}$, we assume that the random errors $\{\epsilon_i(a)\}_{i\in\mathcal{I}}$ are independent.[5] These random errors are not necessarily identically distributed, but we assume they all have an expected value of 0 and a finite variance less than or equal some constant $\sigma_\epsilon^2 > 0$.

Finally, we assume that the random variables $\{c_i\}_{i\in\mathcal{I}}$ are independent of the random functions $\{\epsilon_i\}_{i\in\mathcal{I}}$.

Assumption (U1) encodes problem (Pr1), while assumption (U2) encodes both (Pr2) and (Pr3). Note that, while we assume $\{v_i\}_{i\in\mathcal{I}}$ and $\{c_i\}_{i\in\mathcal{I}}$ are random variables, we make no assumptions about the mechanism generating the profile of utility functions $\{u_i\}_{i\in\mathcal{I}}$. These utility functions might be fixed in advance, or they might themselves be generated by some other random process.[6] If they *are* randomly generated, then we do *not* need to assume that $\{u_i\}_{i\in\mathcal{I}}$ are identically distributed, or assume that the random variables $u_i(a)$ and $u_i(b)$ are independent for a given voter i in $\mathcal{I}$ and distinct alternatives a and b in $\mathcal{A}$. (Indeed, this would be highly unrealistic.) However, we will assume the utility profile $\{u_i\}_{i\in\mathcal{I}}$ satisfies the following boundedness condition:

(U3) There is a constant $M > 0$ such that, for every alternative a in $\mathcal{A}$, we have $\frac{1}{I}\sum_{i\in\mathcal{I}} u_i(a)^2 < M^2$.

For example, if $|u_i(a)| < M$ for every alternative a in $\mathcal{A}$ and every voter i in $\mathcal{I}$, then (U3) is clearly satisfied. Alternately, suppose that, for each a in $\mathcal{A}$, the utility values $\{u_i(a)\}_{i\in\mathcal{I}}$ are independent, identically distributed random variables drawn from a distribution with mean zero and variance less than M^2. If I is large, then the Law of Large Numbers says that (U3) will hold with very high probability.

For every alternative a in $\mathcal{A}$, we define the "observed" social welfare function

$$V_{\mathcal{I}}(a) \quad := \quad \frac{1}{I}\sum_{i\in\mathcal{I}} v_i(a). \tag{2}$$

Thus, $V_{\mathcal{I}}$ is based on observable data (the functions $\{v_i\}_{i\in\mathcal{I}}$), and does not require the true values of the utility functions $\{u_i\}_{i\in\mathcal{I}}$ or the interpersonal calibration constants $\{c_i\}_{i\in\mathcal{I}}$. Therefore, the social planner can compute $V_{\mathcal{I}}$, and identify the alternative in $\mathcal{A}$ which

[4]The assumption that $\{c_i\}_{i\in\mathcal{I}}$ all have expected value 1 involves no loss of generality. If there was some i in $\mathcal{I}$ such that $\mathbb{E}[c_i] = \overline{c}_i \neq 1$, then we could replace the utility function u_i with $\widetilde{u}_i = \overline{c}_i\, u_i$ and replace the random variable c_i with $\widetilde{c}_i = c_i/\overline{c}_i$; we would then have $\widetilde{c}_i\,\widetilde{u}_i = c_i\, u_i$, but $\mathbb{E}[\widetilde{c}_i] = 1$, as desired.

[5]Note that we do *not* assume that, for a fixed voter i in $\mathcal{I}$, the random errors $\epsilon_i(a)$ and $\epsilon_i(b)$ are independent for different alternatives a and b in $\mathcal{A}$.

[6]In this case, Theorem 1 (below) should be interpreted as a statement which holds for any specific realization of these random utility functions.

maximizes $V_{\mathcal{I}}$. Our first result says that, if the population is sufficiently large, then $V_{\mathcal{I}}$ is a good approximation of $U_{\mathcal{I}}$, so that an alternative in $\mathrm{argmax}_{\mathcal{A}}(V_{\mathcal{I}})$ will also maximize (or almost maximize) the value of $U_{\mathcal{I}}$.

Before stating this result, we must introduce some more notation. Let $\delta > 0$ represent a "social suboptimality tolerance", and let $p > 0$ represent the probability that this tolerance will be exceeded (we want both of these to be small). For any values of δ and p, we define

$$\overline{I}(\delta, p) \quad := \quad 8\,\frac{M^2\,\sigma_c^2 + \sigma_\epsilon^2}{p\,\delta^2}. \tag{3}$$

Our first result says that, for any voter population larger than $\overline{I}(\delta, p)$, any $V_{\mathcal{I}}$-maximizing social alternative will produce a social welfare within δ of the theoretical optimum, with probability at least $1 - p$.

Theorem 1 *For every voter i in $\mathcal{I}$, let $u_i : \mathcal{A} \longrightarrow \mathbb{R}$ be a utility function. Suppose that the profile $\{u_i\}_{i \in \mathcal{I}}$ satisfies rule* (U3)*, and suppose $\{c_i\}_{i\in\mathcal{I}}$ and $\{v_i\}_{i\in\mathcal{I}}$ are randomly generated according to rules* (U1) *and* (U2)*. Define $U_{\mathcal{I}}, V_{\mathcal{I}} : \mathcal{A} \longrightarrow \mathbb{R}$ as in equations* (1) *and* (2)*, and let $U^* := \max\{U_{\mathcal{I}}(a);\ a \in \mathcal{A}\}$ (the optimum utilitarian social welfare). Then for any $\delta > 0$ and any a_V^* in $\mathrm{argmax}_{\mathcal{A}}(V_{\mathcal{I}})$, we have*

$$\lim_{I\to\infty} \mathrm{Prob}\left[U_{\mathcal{I}}(a_V^*) \geq U^* - \delta\right] \quad = \quad 1.$$

To be precise, for any $p > 0$ and any $I \geq \overline{I}(\delta, p)$, we have $\mathrm{Prob}\left[U_{\mathcal{I}}(a_V^*) < U^* - \delta\right] \ < \ p$.

For example, for every alternative a in $\mathcal{A}$, suppose the utilities $\{u_i(a);\ i \in \mathcal{I}\}$ are independent, uniformly distributed random variables ranging over some interval of length at most 10, contained within the interval $[-9, 9]$ (with perhaps different subintervals of $[-9, 9]$ for different alternatives in $\mathcal{A}$). Let $M := 5$; then for a large population of voters, condition (U3) will be satisfied with extremely high probability.[7] Suppose $\sigma_c^2 = 1$ and $\sigma_\epsilon^2 = 5$, and let $\delta := 0.2$ (i.e. 1.1% of the total utility range) and $p := 0.01$. Then $\overline{I}(\delta, p) = 600\,000$. Thus, for a polity of six hundred thousand voters, Theorem 1 says that, with 99% probability, the $V_{\mathcal{I}}$-maximizing alternative will yield a $U_{\mathcal{I}}$-value within 1.1% of the theoretical optimum U^*.

Define $\Delta(U_{\mathcal{I}}) := \inf\{U^* - U_{\mathcal{I}}(a);\ \ a \not\in \mathrm{argmax}_{\mathcal{A}}(U_{\mathcal{I}})\}$; this is the minimum "social welfare cost" of failing to choose a utilitarian optimum. If $\mathcal{A}$ is infinite, then $\Delta(U_{\mathcal{I}})$ may be zero. But if $\mathcal{A}$ is finite, then $\Delta(U_{\mathcal{I}}) > 0$. If $\Delta(U_{\mathcal{I}}) > 0$, then Theorem 1 implies that a $V_{\mathcal{I}}$-optimal alternative will not merely be close, but will in fact *be* a utilitarian optimum, in the limit of a large population:

Corollary 2 *Assume the same hypotheses as Theorem 1. If $\delta \leq \Delta(U_{\mathcal{I}})$, and $I \geq \overline{I}(\delta, p)$, then* $\mathrm{Prob}\left[\mathrm{argmax}_{\mathcal{A}}(V_{\mathcal{I}}) \subseteq \mathrm{argmax}_{\mathcal{A}}(U_{\mathcal{I}})\right] > 1 - p$.

[7] *Proof:* The variance of any such uniform distribution is at most $8\frac{1}{3} < 9$, and the square of its mean is at most $4^2 = 16$. Thus, its second moment will be at most $9 + 16 = 25 = 5^2$.

3 Relative utilitarianism

Theorem 1 deals with problems (Pr1)-(Pr3), but it does little about (Pr4). Strategic voters will exaggerate their utility functions. One partial solution is to rescale each voter's utility function to range over the interval $[0, 1]$.[8] The resulting social choice rule is called *Relative utilitarianism* (Dhillon, 1998; Dhillon and Mertens, 1999).

Formally, for every voter i in $\mathcal{I}$ let $w_i : \mathcal{A} \longrightarrow \mathbb{R}$ be her "true" utility function. We suppose these utility functions admit one-for-one cardinal interpersonal comparisons. In other words, for any alternatives a, b, c, and d in $\mathcal{A}$, if $w_i(b) - w_i(a) = w_j(d) - w_j(c)$, then the welfare that voter i gains in moving from a to b exactly the same as the welfare that voter j gains in moving from c to d. We therefore want to maximize the utilitarian SWF $U_{\mathcal{I}}$ defined by

$$U_{\mathcal{I}}(a) \quad := \quad \sum_{i \in \mathcal{I}} w_i(a), \quad \text{for every alternative } a \text{ in } \mathcal{A}. \tag{4}$$

Let $\underline{w}_i := \min\{w_i(a);\ a \in \mathcal{A}\}$. By replacing w_i with the function $\widetilde{w}_i := w_i - \underline{w}_i$ if necessary, we can suppose that $\min\{w_i(a);\ a \in \mathcal{A}\} = 0$, for every voter i in $\mathcal{I}$. Clearly this does not affect the maximizer of (4).

Next, let $c_i := \max\{w_i(a);\ a \in \mathcal{A}\}$, and then define $u_i(a) := w_i(a)/c_i$, for every voter i in $\mathcal{I}$ and every alternative a in $\mathcal{A}$. Then formula (4) is clearly equivalent to formula (1). The scaling constants $\{c_i\}_{i \in \mathcal{I}}$ represent the "preference intensities" of the voters, which we assume are unknown to the social planner; we will treat these as independent random variables, as in assumption (U1). Note that each u_i ranges over the interval $[0, 1]$. The *relative utilitarian* social welfare function $RU : \mathcal{A} \longrightarrow \mathbb{R}$ is defined:

$$RU(a) \quad := \quad \sum_{i \in \mathcal{I}} u_i(a), \quad \text{for every alternative } a \text{ in } \mathcal{A}.$$

For simplicity, we will suppose that the voters have perfect knowledge of their own utilities, and that the planner is able to obtain accurate reports from them (i.e. we will neglect issues (Pr2) and (Pr3)). In terms of assumption (U2), this means we suppose $\epsilon_i(a) = 0$, so that $v_i(a) = u_i(a)$ for every voter i in $\mathcal{I}$ and every alternative a in $\mathcal{A}$. As in Section 2, we define $U^* := \max\{U_{\mathcal{I}}(a);\ a \in \mathcal{A}\}$.

Proposition 3 *Suppose the preference intensities $\{c_i\}_{i \in \mathcal{I}}$ are independent random variables, as in assumption* (U1). *Let $a^*_{\mathrm{RU}} \in$* $\mathrm{argmax}_{\mathcal{A}}(RU)$ (the relative utilitarian social choice). *Then for any $\delta > 0$, we have*

$$\lim_{I \to \infty} \mathrm{Prob}\left[U_{\mathcal{I}}(a^*_{\mathrm{RU}}) \geq U^* - \delta\right] \quad = \quad 1.$$

To be precise: for any $p > 0$, if $I > 8\sigma_c^2/p\,\delta^2$, then $\mathrm{Prob}\left[U_{\mathcal{I}}(a^*_{\mathrm{RU}}) < U^* - \delta\right] < p$.

[8] Obviously such a rescaling is not a complete solution to strategic voting. We will return to this issue in Sections 5 and 7.

If $\delta < \Delta(U_{\mathcal{I}})$, then one can also obtain a result similar to Corollary 2.

Relative utilitarianism prevents voters from strategically exaggerating their utility functions, but strategic voting is still possible. In the game of strategic voting associated with the relative utilitarian rule, each voter's best response is to assign a utility of either 0 or 1 to each alternative in $\mathcal{A}$ (Núñez and Laslier, 2013). In this case, relative utilitarianism is *de facto* equivalent to *approval voting*. We will analyse the utilitarian efficiency of approval voting in Section 5 below.

4 Scoring rules

Let $N := \mathcal{A}$. Let $s_1 \leq s_2 \leq \cdots \leq s_N$ be real numbers, and define $\mathbf{s} := (s_1, s_2, \ldots, s_N)$. The **s**-*scoring rule* on $\mathcal{A}$ is defined as follows:

1. For every voter i in $\mathcal{I}$, let $\succ_i$ denote her ordinal preferences on $\mathcal{A}$.

2. For every alternative a in $\mathcal{A}$, if a is ranked kth place from the bottom with respect to $\succ_i$, then voter i gives a the score s_k. (In particular, i gives the score s_1 to her least-prefered alternative, and the score s_N to her most prefered alternative.)

3. For each alternative in $\mathcal{A}$, add up the scores it gets from all voters.

4. The alternative with the highest total score is chosen.

For example, the *Borda rule* is the scoring rule with $\mathbf{s} = (1, 2, 3, \ldots, N)$. The standard *plurality rule* is the scoring rule with $\mathbf{s} = (0, 0, \ldots, 0, 1)$. Recently, Apesteguia et al. (2011) have investigated the utilitarian efficiency of scoring rules when $N = 3$ and $|\mathcal{I}|$ is small (e.g. $|\mathcal{I}| = 2$). In this section, we will show that, for *any* finite N, if the profile $\{u_i\}_{i\in\mathcal{I}}$ arises from a fairly large class of statistical models, then there exists a scoring rule which will come arbitrarily close to selecting a utilitarian optimum, with very high probability as $I \to \infty$.

Let ρ be a probability measure on $\mathbb{R}$ with finite variance. For every voter i in $\mathcal{I}$, let $\succ_i$ be voter i's ordinal preference relation on $\mathcal{A}$, and suppose voter i's cardinal utility function $u_i : \mathcal{A} \longrightarrow \mathbb{R}$ is randomly generated as follows:

(R1$_\rho$) Let $\{r^i_1, r^i_2, \ldots, r^i_N\}$ be a sample of N independent, ρ-random variables. Rearrange this sample in increasing order, to obtain $r^i_{(1)} \leq r^i_{(2)} \leq \cdots \leq r^i_{(N)}$ (these are called the *order statistics* of the sample).

(R2$_\rho$) If $\mathcal{A} = \{a_1, a_2, \ldots, a_N\}$ and $a_1 \prec_i a_2 \prec_i \cdots \prec_i a_N$, then set $u_i(a_1) := r^i_{(1)}$, $u_i(a_2) := r^i_{(2)}$, ..., and $u_i(a_N) := r^i_{(N)}$.

For example, suppose $\mathcal{A} = \{a, b, c, d\}$, and voter i has ordinal preferences $a \succ_i b \succ_i c \succ_i d$. Suppose we generate four independent, ρ-random values: $r^i_1 = 0.14$, $r^i_2 = -2.62$, $r^i_3 = -0.36$, and $r^i_4 = 1.47$. Then we would set $u_i(a) := 1.47$, $u_i(b) := 0.14$, $u_i(c) := -0.36$, and $u_i(d) := -2.62$.

Next, let $\mathcal{P} = \{\succ_i\}_{i\in\mathcal{I}}$ be a profile of ordinal preferences for all voters in $\mathcal{I}$. A $(\mathcal{P}, \rho)$-*random utility profile* is a randomly generated profile of interpersonally comparable cardinal utility functions $\{c_i\, u_i\}_{i\in\mathcal{I}}$ such that for every voter i in $\mathcal{I}$, the utility function u_i is randomly generated from $\succ_i$ and ρ using rules (R1$_\rho$) and (R2$_\rho$), while the interpersonal calibration constants $\{c_i\}_{i\in\mathcal{I}}$ are independent random variables as in assumption (U1) from Section 2. Furthermore, we require all the random variables $\{r^i_n;\ \ i \in \mathcal{I} \text{ and } n \in [1 \ldots N]\}$ appearing in rule (R1$_\rho$) to be jointly independent. In particular, for each voter i in $\mathcal{I}$, we require the random utilities $\{r^i_n\}_{n=1}^N$ to be *independent* of i's predetermined preference order $\succ_i$. (Thus, it is *not* the case that voters with some preference orders systematically have stronger preferences than voters with other preference orders.)

For example, suppose that $\{u_i(a);\ i \in \mathcal{I} \text{ and } a \in \mathcal{A}\}$ is a collection of $I \times N$ independent, ρ-random variables (this is a version of the so-called "Impartial Culture" model). For every voter i in $\mathcal{I}$, let $\succ_i$ be the ordinal preferences corresponding to u_i, and then define $\mathcal{P} = \{\succ_i\}_{i\in\mathcal{I}}$. Set $c_i := 1$ for every voter i in $\mathcal{I}$. Then $\{u_i\}_{i\in\mathcal{I}}$ is a $(\mathcal{P}, \rho)$-random utility profile. However, the model of $(\mathcal{P}, \rho)$-random utility profiles is much more general than this Impartial Culture model, because we make no assumptions about the origins of the ordinal preference profile $\mathcal{P} = \{\succ_i\}_{i\in\mathcal{I}}$. For example, $\mathcal{P}$ might be predetermined, or it might itself be randomly generated by some other (unspecified) stochastic process. Furthermore, if $\mathcal{P}$ *is* randomly generated, then (unlike the Impartial Culture model) we do *not* necessarily suppose that all $N!$ possible ordinal preferences on $\mathcal{A}$ are equally likely to occur in $\mathcal{P}$.[9]

Now, if we take a random sample of N independent random variables drawn from ρ, and compute the order statistics of this sample, then we get N new random variables. Let $s^N_1 \leq s^N_2 \leq \cdots \leq s^N_N$ be the *expected values* of these random variables. We can define a scoring rule on $\mathcal{A}$ using the vector $\mathbf{s} := (s^N_1, s^N_2, \ldots, s^N_N)$; we will call this the *ρ-scoring rule*. Our next result says that this scoring rule provides a surprisingly good approximation of the utilitarian social choice rule for large populations. As in Section 2, we define the utilitarian social welfare function $U_\mathcal{I}$ by formula (1), and let $U^* := \max\{U_\mathcal{I}(a);\ \ a \in \mathcal{A}\}$. As in assumption (U1), suppose the random variables $\{c_i\}_{i\in\mathcal{I}}$ all have variance bounded by some constant $\sigma_c^2 > 0$.

Proposition 4 *Let $\mathcal{A}$ be a finite set, let $\mathcal{P} = \{\succ_i\}_{i\in\mathcal{I}}$ be a profile of preference orders on $\mathcal{A}$, let ρ be a finite-variance probability measure on $\mathbb{R}$, and let $a^*_{\text{scr}} \in \mathcal{A}$ be the result of applying the ρ-scoring rule to $\mathcal{P}$. Let $\{c_i\, u_i\}_{i\in\mathcal{I}}$ be any $(\mathcal{P}, \rho)$-random utility profile. Then for any $\delta > 0$, we have*

$$\lim_{I\to\infty} \text{Prob}\left[U_\mathcal{I}(a^*_{\text{scr}}) \geq U^* - \delta\right] \quad = \quad 1. \tag{5}$$

Furthermore, if the fourth moment of ρ is finite,[10] *then there are constants $C_1, C_2 > 0$*

[9]Indeed, one problem with Impartial Culture models is that, in a large population, all elements of $\mathcal{A}$ end up with roughly the *same* average utility (due to the Law of Large Numbers), so that utilitarianism is effectively indifferent between them, and the use of *any* voting rule is somewhat superfluous. The model described here avoids this unrealistic outcome.

[10]The *fourth moment* of the probability measure ρ is the integral $\displaystyle\int_{-\infty}^{\infty} x^4\ \mathrm{d}\rho[x]$. It is finite if $\mathrm{d}\rho[x]$ decays quickly enough as $|x|\to\infty$. For example, the fourth moment of a normal probability measure is finite.

(determined by ρ and σ_c^2) such that, for any $p > 0$, if $I \geq C_1/p$ and $I \geq C_2/p\,\delta^2$, then $\text{Prob}\,[U_{\mathcal{I}}(a^*_{\text{scr}}) < U^* - \delta] < p$.

If $\delta < \Delta(U_{\mathcal{I}})$, then one can also obtain a result similar to Corollary 2.

It is convenient to "renormalize" $s_1^N, s_2^N, \ldots, s_N^N$ to range over the interval $[-1, 1]$, by defining

$$\widetilde{s}_n^N \quad := \quad \frac{2\,s_n^N - s_N^N - s_1^N}{s_N^N - s_1^N}, \quad \text{for all } n \text{ in } [1 \ldots N].$$

This ensures that $\widetilde{s}_N^N = 1$ and $\widetilde{s}_1^N = -1$. (For example, if $N = 3$, then we have $\widetilde{s}_3^3 = 1$ and $\widetilde{s}_1^3 = -1$, and only the value of $\widetilde{s}_2^3$ remains to be determined.) If ρ is a probability distribution symmetrically distributed about some point in the real line, then the values $\widetilde{s}_1^N, \widetilde{s}_2^N, \ldots, \widetilde{s}_N^N$ will by symmetrically distributed around zero —that is, $\widetilde{s}_k^N = -\widetilde{s}_{N+1-k}^N$ for all k in $[1 \ldots N]$. Thus, if N is odd and $k = (N+1)/2$, then $\widetilde{s}_k^N = 0$. In particular, if $N = 3$, then we must have $\widetilde{s}_2^3 = 0$, while $\widetilde{s}_3^3 = 1$ and $\widetilde{s}_1^3 = -1$. Thus, we get the scoring rule defined by the scoring vector $(-1, 0, 1)$, which is obviously equivalent to the Borda rule. Thus, Proposition 4 implies the next result, which says that the Borda rule is the utilitarian-optimal scoring rule for *any* symmetric measure ρ. This confirms a conjecture made by Apesteguia et al. (2011; §4.4).

Corollary 5 *Suppose $|\mathcal{A}| = 3$, let $\mathcal{P} = \{\succ_i\}_{i \in \mathcal{I}}$ be a profile of preference orders on $\mathcal{A}$, and let $a^*_{\text{Brd}} \in \mathcal{A}$ be the result of applying the Borda rule to $\mathcal{P}$. Let ρ be any symmetric, finite-variance probability distribution on $\mathbb{R}$, and let $\{c_i\, u_i\}_{i \in \mathcal{I}}$, $U_{\mathcal{I}}$, and U^* be as in Proposition 4. Then for any $\delta > 0$, we have* $\lim_{I \to \infty} \text{Prob}\,[U_{\mathcal{I}}(a^*_{\text{Brd}}) \geq U^* - \delta] = 1$.

If $|\mathcal{A}| \geq 4$, then the Borda rule is no longer guaranteed to be the optimal scoring rule; the optimal scoring rule will depend on the expected values of the order statistics for ρ, which depend on the structure of ρ itself. For example, suppose ρ was a normal probability distribution and $|\mathcal{A}| = 7$. Then we get the following expected order statistics (to 5 significant digits).[11]

$$\begin{array}{rcl} s_7^7 & \approx & 1.35218, \\ s_6^7 & \approx & 0.75737, \\ s_5^7 & \approx & 0.35271, \\ s_4^7 & = & 0, \\ s_3^7 & \approx & -0.35271, \\ s_2^7 & \approx & -0.75737, \\ \text{and } s_1^7 & \approx & -1.35218, \end{array} \qquad \text{which renormalize to} \qquad \begin{array}{rcl} \widetilde{s}_7^7 & = & 1, \\ \widetilde{s}_6^7 & \approx & 0.56011, \\ \widetilde{s}_5^7 & \approx & 0.26085, \\ \widetilde{s}_4^7 & = & 0, \\ \widetilde{s}_3^7 & \approx & -0.26085, \\ \widetilde{s}_2^7 & \approx & -0.56011, \\ \text{and } \widetilde{s}_1^7 & = & -1. \end{array}$$

By comparision, the Borda rule uses the scoring vector $(-1, -0.6\overline{6}, -0.3\overline{3}, 0, 0.3\overline{3}, 0.6\overline{6}, 1)$.

Unfortunately, the expected values of order statistics are quite hard to compute for many probability distributions. Harter and Balakrishnan (1996) provide tables of these

[11] Here we suppose for simplicity that ρ is a standard normal distribution. Any other normal distribution would yield the same values for $\widetilde{s}_1^7, \ldots, \widetilde{s}_7^7$ after renormalization.

expected values for most of the common probability distributions (e.g. normal, exponential, Weibull, etc.); from this data it is easy to design the appropriate scoring rule. However, this raises the question: what kind of distribution is ρ? This question must be settled empirically.

5 Polarization and approval voting

If the distribution ρ is strongly concentrated around its mean (e.g. ρ is a normal distribution), and $\{u_i\}_{i \in \mathcal{I}}$ is a $(\mathcal{P}, \rho)$-random profile of utility functions, then $u_i(a)$ will be relatively close to zero for most voters i in $\mathcal{I}$ and most alternatives a in $\mathcal{A}$. In other words, a voter will generally assign utilities of large magnitude only to what she regards as the very best and very worst alternatives, and assign a small-magnitude utility to most of the other alternatives in $\mathcal{A}$. Thus, this model describes a community of voters with "moderate" political opinions.

However, on some issues, voters are highly polarized. They assign very high utilities to some alternatives, and very low utilities to all the rest, with nothing in the middle. We could model this with a $(\mathcal{P}, \rho)$-random profile where ρ is a heavy-tailed distribution (e.g. a Student t-distribution), which has a high probability of producing very large or very small values. But such a symmetric distribution has the unrealistic consequence that all voters will judge roughly half the alternatives to be "good" and the other half to be "bad". If ρ was a positively-skewed distribution (e.g. a Poisson, Weibull, Pareto, or log-normal distibution), then we would end up with a more unbalanced form of polarization, where each voter identifies only one or two alternatives as "good", and regards almost all the rest as being "bad". (The opposite statement holds for negatively skewed distributions.) However, this would still have the unrealistic feature that most voters would identify exactly the same number of "good" (or "bad") alternatives. In a more realistic model of political polarization, there may be many voters who like only 10% of the alternatives, while strongly rejecting the other 90%, but there may also be another large group of voters who are fairly happy with 60% of the alternatives, but strongly reject the other 40%. This sort of scenario cannot be captured with a $(\mathcal{P}, \rho)$-random utility profile, for any ρ.

In this section, we will show that a utilitarian optimum for such a scenario will be identified (with high probability) by *approval voting* (Brams and Fishburn, 1983). The approval voting rule works as follows:

1. Each voter i identifies a subset of alternatives which she regards as "good enough".

2. For each social alternative a, count how many voters regard a as good enough.

3. The alternative which is good enough for the largest number of voters is chosen.

Let γ be a finite-variance probability measure on $\mathbb{R}_+ := (0, \infty)$, and let β be a finite-variance probability measure on $\mathbb{R}_- := (-\infty, 0]$.[12] For every voter i in $\mathcal{I}$, let $\mathcal{G}_i \subseteq \mathcal{A}$ be

[12]Mnemonic: "γ" is for "good enough" and "β" is for "bad". Note that these are *not* assumed to be Gamma or Beta distributions.

represent the set of alternatives which voter i regards as "good enough". If $\mathcal{B}_i := \mathcal{A} \setminus \mathcal{G}_i$ then we suppose i regards all alternatives in $\mathcal{B}_i$ as being "bad". Suppose voter i's cardinal utility function $u_i : \mathcal{A} \longrightarrow \mathbb{R}$ is randomly generated as follows:

(R1$_{\gamma,\beta}$) For all g in $\mathcal{G}_i$, let r^i_g be a γ-random variable. For all b in $\mathcal{B}_i$, let r^i_b be a β-random variable. Assume that the random variables $\{r^i_a;\ a \in \mathcal{A}\}$ are all independent.

(R2$_{\gamma,\beta}$) For every alternative a in $\mathcal{A}$, let $u_i(a) = r^i_a$.

For example, suppose $\mathcal{A} = \{a, b, c, d, e\}$, and $\mathcal{G}_i = \{a, b, c\}$ (so that $\mathcal{B}_i = \{d, e\}$). Suppose $r^i_a = 0.31$, $r^i_b = 0.14$ and $r^i_c = 0.71$ are three independent γ-random variables, and let $r^i_d = -0.67$ and $r^i_e = -0.19$ be two independent β-random variables. Then we would set $u_i(a) := 0.31$, $u_i(b) := 0.14$, $u_i(c) := 0.71$, $u_i(d) := -0.67$, and $u_i(e) := -0.19$.

Next, let $\mathfrak{G} = \{\mathcal{G}_i\}_{i \in \mathcal{I}}$ be a profile assigning a subset of $\mathcal{A}$ to each voter in $\mathcal{I}$ (we will call this an *approval profile* on $\mathcal{A}$). A $(\mathfrak{G}, \gamma, \beta)$*-random utility profile* is a profile of random utility functions $\{c_i\, u_i\}_{i \in \mathcal{I}}$ such that for every voter i in $\mathcal{I}$, the utility function u_i is randomly generated from $\mathcal{G}_i$, γ and β using rules (R1$_{\gamma,\beta}$) and (R2$_{\gamma,\beta}$), while the interpersonal calibration constants $\{c_i\}_{i \in \mathcal{I}}$ are independent random variables as in assumption (U1) from Section 2. Furthermore, we require all the random variables $\{r^i_a;\ i \in \mathcal{I} \text{ and } a \in \mathcal{A}\}$ appearing in rule (R1$_{\gamma,\beta}$) to be jointly independent.

We will suppose that, in approval voting, each voter i votes for all and only the alternatives in the set $\mathcal{G}_i$.[13] Our next result says that approval voting provides a surprisingly good approximation of utilitarian social choice for large populations with utility profiles of this kind. As in Section 2, we define the utilitarian social welfare function $U_{\mathcal{I}}$ by formula (1), and let $U^* := \max\{U_{\mathcal{I}}(a);\ a \in \mathcal{A}\}$.

Proposition 6 *Let $\mathcal{A}$ be a finite set, let $\mathfrak{G}$ be an approval profile on $\mathcal{A}$, and let $a^*_{\text{appr}} \in \mathcal{A}$ be the result of applying approval voting to $\mathfrak{G}$. Let γ and β be finite-variance probability measures on $\mathbb{R}_+$ and $\mathbb{R}_-$ respectively, and let $\{c_i\, u_i\}_{i \in \mathcal{I}}$ be any $(\mathfrak{G}, \gamma, \beta)$-random utility profile. Then for any $\delta > 0$, we have*

$$\lim_{I \to \infty} \text{Prob}\left[U_{\mathcal{I}}(a^*_{\text{appr}}) \geq U^* - \delta\right] \quad = \quad 1. \tag{6}$$

Furthermore, if the fourth moments of γ and β are finite, then there are constants $C_1, C_2 > 0$ (determined by γ, β, and σ_c^2) such that, for any $p > 0$, if $I \geq C_1/p$ and $I \geq C_2/p\,\delta^2$, then $\text{Prob}\left[U_{\mathcal{I}}(a^*_{\text{appr}}) < U^* - \delta\right] < p$.

If $\delta < \Delta(U_{\mathcal{I}})$, then one can also obtain a result similar to Corollary 2.

[13] This means that we suppose each voter is honest. But one advantage of approval voting is that the extent of strategic voting is generally small; see (Brams and Fishburn, 1983, Ch.2) and (Laslier and Sanver, 2010, Part IV).

6 Condorcet consistent rules

Let $\mathcal{A}$ be a set of alternatives, and let $\mathcal{P} = \{\succ_i\}_{i\in\mathcal{I}}$ be a profile of preference orders on $\mathcal{A}$. Let $a \in \mathcal{A}$. We say that a is a *Condorcet winner* if, for every other $b \in \mathcal{A}$, some majority prefers a over b —that is, $\#\{i \in \mathcal{I};\ \ a \succ_i b\} \geq I/2$. The Condorcet winner is one of the oldest and most well-known solution concepts in social choice theory. A voting rule is called *Condorcet consistent* if it selects a Condorcet winner whenever one exists. Many well-known voting rules are Condorcet consistent, including the Copeland rule, the Simpson-Kramer rule, the Slater rule, the Kemeny rule, and any voting rule which selects the winner through a sequence of pairwise majority votes.

Unfortunately, not all preference profiles admit a Condorcet winner. Furthermore, in general, there is no relationship between Condorcet consistency and social welfare. However, in this section, we will that the Condorcet winner will be the utilitarian optimum, for many plausible statistical models of voter preferences.

Without loss of generality, we can suppose that the utility functions $\{u_i\}_{i\in\mathcal{I}}$ admit one-for-one cardinal interpersonal comparisons. (In the notation of assumption (U1), we suppose $c_i = 1$ for every voter i in $\mathcal{I}$. If this is not the case, then simply replace u_i with $\widetilde{u}_i = c_i\, u_i$ for each i in $\mathcal{I}$.) Thus, we seek the social alternative which maximizes the utilitarian social welfare function $U_{\mathcal{I}}$ defined by

$$U_{\mathcal{I}}(a) \quad := \quad \frac{1}{|\mathcal{I}|}\sum_{i\in\mathcal{I}} u_i(a), \quad \text{for every alternative } a \text{ in } \mathcal{A}. \tag{7}$$

Let a and b be alternatives in $\mathcal{A}$, and for every voter i in $\mathcal{I}$, let $u^i_{a,b} := u_i(a) - u_i(b)$. Thus, $U_{\mathcal{I}}(a) \geq U_{\mathcal{I}}(b)$ if and only if the mean of the set $\mathcal{U}_{a,b} := \{u^i_{a,b}\}_{i\in\mathcal{I}}$ is positive. Meanwhile, a strict majority prefers a over b if and only if the *median* of $\mathcal{U}_{a,b}$ is positive. Thus, a strict majority will choose the $U_{\mathcal{I}}$-maximizing element of the pair $\{a, b\}$ if and only if $\text{sign}[\text{median}(\mathcal{U}_{a,b})] = \text{sign}[\text{mean}(\mathcal{U}_{a,b})]$. In this case, we say that the utility profile $\{u^i\}_{i\in\mathcal{I}}$ is *reasonable* relative to a and b.[14]

Example 7. If $|\text{mean}(\mathcal{U}_{a,b})|$ exceeds the standard deviation of the set $\mathcal{U}_{a,b}$ (i.e. if the social welfare gap between the alternatives a and b is large enough), then the utility profile $\{u^i\}_{i\in\mathcal{I}}$ is (a, b)-reasonable. To see this, note that Chebyshev's inequality implies that $|\text{median}(\mathcal{U}_{a,b}) - \text{mean}(\mathcal{U}_{a,b})| \leq \text{std dev}(\mathcal{U}_{a,b})$. ♢

We say the utility profile $\{u_i\}_{i\in\mathcal{I}}$ is *reasonable* if it is (a, b)-reasonable, for every possible pair $a, b \in \mathcal{A}$. The following observation is immediate.

[14] If I is odd, then $\text{median}[\mathcal{U}_{ab}]$ is the unique point m in $\mathcal{U}_{a,b}$ such that $\#\{i \in \mathcal{I};\ \ u^i_{a,b} \geq m\} > I/2$ and $\#\{i \in \mathcal{I};\ \ u^i_{a,b} \leq m\} > I/2$. However, if I is *even*, then $\text{median}[\mathcal{U}_{ab}]$ is generally an *interval* $[\underline{m}, \overline{m}]$ with $\underline{m} \leq \overline{m}$, such that $\#\{i \in \mathcal{I};\ \ u^i_{a,b} \geq \underline{m}\} \geq I/2$ and $\#\{i \in \mathcal{I};\ \ u^i_{a,b} \leq \overline{m}\} \geq I/2$. In this case, we will say $\text{median}[\mathcal{U}_{ab}]$ is *positive* if $\overline{m} \geq \underline{m} > 0$, and we will say $\text{median}[\mathcal{U}_{ab}]$ is *negative* if $\underline{m} \leq \overline{m} < 0$. If $\underline{m} \leq 0 \leq \overline{m}$, then we consider the "sign" of $\text{median}[\mathcal{U}_{ab}]$ to be undefined (in this case, the voters are evenly split between alternatives a and b). Note that the definition of "reasonable" specifically excludes this last possibility.

Proposition 8 *Let* $\mathcal{U} = \{u_i\}_{i\in\mathcal{I}}$ *be a cardinal utility profile, and let* $\mathcal{P} = \{\succ_i\}_{i\in\mathcal{I}}$ *be the corresponding profile of ordinal preferences. If* $\mathcal{U}$ *is reasonable, then* $\mathcal{P}$ *admits a Condorcet winner. Furthermore, this Condorcet winner maximizes the utilitarian social welfare function* $U_{\mathcal{I}}$ *in equation* (7).

Reasonability may seems like a heroic assumption, but the rest of this section will show that it is actually quite plausible, under certain hypotheses. We will suppose that the voters' utility functions are randomly generated by some stochastic process. Our aim is to show that, in a large population, such a randomly generated utility profile *will* be reasonable, with very high probability. We will establish this result for two plausible families of stochastic processes.

6.1 Random utility functions

Suppose $\mathcal{A}$ is a finite set, so that utility functions correspond to vectors in $\mathbb{R}^{\mathcal{A}}$. Let ρ be a probability measure on $\mathbb{R}^{\mathcal{A}}$. We will use ρ to randomly generate utility functions for the voters. For any distinct alternatives a and b in $\mathcal{A}$, let $\rho_{a,b}$ be the distribution of the quantity $x_a - x_b$, where $\mathbf{x}$ is a ρ-random variable. We will say that the measure ρ is ***reasonable*** if ρ has finite variance, and if mean$[\rho_{a,b}]$ and median$[\rho_{a,b}]$ are nonzero and have the same sign, for all distinct alternatives a and b in $\mathcal{A}$. (For example, suppose ρ is any multivariate normal distribution on $\mathbb{R}^{\mathcal{A}}$ with mean $\mathbf{m} \in \mathbb{R}^{\mathcal{A}}$ such that $m_a \neq m_b$ for any distinct a and b in $\mathcal{A}$. Then ρ is reasonable.) This section's first result says that reasonable measures generate reasonable utility profiles.

Proposition 9 *Let* $\mathcal{A}$ *be a finite set, let* ρ *be a reasonable probability measure on* $\mathbb{R}^{\mathcal{A}}$, *and suppose that the utility functions* $\{u_i\}_{i\in\mathcal{I}}$ *are independent,* ρ*-random variables. Then*

$$\lim_{I\to\infty} \mathrm{Prob}\left(\textit{The utility profile } \{u_i\}_{i\in\mathcal{I}} \textit{ is reasonable}\right) \;=\; 1.$$

To be precise, there are constants $q \in (0,1)$ *and* $C > 0$ *(determined by the structure of* ρ*) such that, if* I *is large enough, then*

$$\mathrm{Prob}\left(\{u_i\}_{i\in\mathcal{I}} \textit{ is not reasonable}\right) \quad\leq\quad \frac{|\mathcal{A}|^2}{2}\left(3\sqrt{I}\,q^I + \frac{C}{I}\right) \quad\xrightarrow[I\to\infty]{}\quad 0.$$

Remark. $q^I \to 0$ very rapidly as $I\to\infty$. Thus, Prob$[\{u_i\}_{i\in\mathcal{I}}$ is not reasonable$]$ is dominated by the term $\frac{C\,|\mathcal{A}|^2}{2I}$. For example suppose $q = 0.98$. If $I \geq 10\,000$, then $\sqrt{I}(0.98)^I \leq 10^{-85}$, so we can ignore it. Suppose $|\mathcal{A}| = 7$ and $C = 10$; Then for $I = 10\,000$, we have

$$\mathrm{Prob}\left(\{u_i\}_{i=1}^{10000} \text{ is not reasonable}\right) \;\leq\; \frac{49}{2}\left(3\sqrt{I}\,(0.98)^I + \frac{10}{I}\right) \;\approx\; \frac{25\cdot 10}{10\,000} \;=\; 0.025.$$

In other words, a ρ-random utility profile $\{u_i\}_{i=1}^{10000}$ will be reasonable with probability at least 97.5%. Thus, with very high probability, the Condorcet winner of such a profile will be the utilitarian optimum. (In particular, this holds if the utility profile $\{u_i\}_{i=1}^{10000}$ is generated from any multivariate normal distribution on $\mathbb{R}^{\mathcal{A}}$ whose mean coordinates are all distinct.)

6.2 Random ideal points

Spatial voting models are very common in the theoretical political science literature.[15] In these models, we regard $\mathbb{R}^N$ as a space of policies described by N distinct parameters. (For example, different coordinates of $\mathbb{R}^N$ might represent interest rates, tax rates, expenditure levels for various public goods or income support mechanisms, and/or the inflation and unemployment rates.) We further suppose that each voter i in $\mathcal{I}$ has some "ideal point" $\mathbf{x}_i$ in $\mathbb{R}^N$, and has a *distance-based* utility function of the form $u_i(\mathbf{a}) = -\phi(\|\mathbf{a} - \mathbf{x}_i\|)$ for some increasing function $\phi : [0, \infty) \longrightarrow \mathbb{R}$. (Here, $\|\bullet\|$ is the Euclidean norm on $\mathbb{R}^N$.) Thus, voter i prefers policy points in $\mathbb{R}^N$ which are closer to her ideal point $\mathbf{x}_i$.

Let ρ be a continuous probability measure on $\mathbb{R}^N$. We will use ρ to randomly generate the ideal points of the voters. For any vector $\mathbf{v}$ in $\mathbb{R}^N$, a $\mathbf{v}$-*median hyperplane* of ρ is any hyperplane $\mathcal{H}_{\mathbf{v}}^{\rho} \subset \mathbb{R}^N$ which is orthogonal to $\mathbf{v}$, and such that at least half the mass of ρ lies on each side of $\mathcal{H}_{\mathbf{v}}^{\rho}$. Such a hyperplane always exists, but it might not be unique for some vectors $\mathbf{v}$ in $\mathbb{R}^N$. However, if there is a $\mathbf{v}$-median hyperplane $\mathcal{H}_{\mathbf{v}}^{\rho}$ which intersects the support of ρ, then $\mathcal{H}_{\mathbf{v}}^{\rho}$ is the *only* $\mathbf{v}$-median hyperplane.[16] (If $N = 1$, then the vector $\mathbf{v}$ is irrelevant, and a median "hyperplane" of ρ is actually a single point —it is any point h in $\mathbb{R}$ such that $\rho(-\infty, h] \geq \frac{1}{2}$ and $\rho[h, \infty) \geq \frac{1}{2}$.)

Now let $\phi : [0, \infty) \longrightarrow \mathbb{R}$ be any convex increasing function (e.g. $\phi(x) = x^p$, for some $p \geq 1$). The ϕ-*median* of ρ is the set of global minima for the function $\Phi_\rho : \mathbb{R}^N \longrightarrow \mathbb{R}$ defined by

$$\Phi_\rho(\mathbf{m}) \quad := \quad \int_{\mathbb{R}^N} \phi\left(\|\mathbf{m} - \mathbf{x}\|\right) \, \mathrm{d}\rho[\mathbf{x}], \quad \text{for all } \mathbf{m} \text{ in } \mathbb{R}^N. \tag{8}$$

For example, if $N = 1$ and $\phi(x) = x$ for all $x \geq 0$, then the ϕ-median of ρ is the classical median of ρ: the point(s) in $\mathbb{R}$ which cut the distribution of ρ into two equal halves. We will say that ρ is ϕ-*balanced* if:

(B1) The function Φ_ρ is well-defined by formula (8);[17]

(B2) The ϕ-median of ρ is a single point, $\mathbf{m}_\rho^\phi$;

(B3) Φ_ρ is rotationally symmetric around $\mathbf{m}_\rho^\phi$; and

(B4) For every $\mathbf{v}$ in $\mathbb{R}^N$, there is a unique $\mathbf{v}$-median hyperplane $\mathcal{H}_\rho^{\mathbf{v}}$, and $\mathbf{m}_\rho^\phi \in \mathcal{H}_\rho^{\mathbf{v}}$.

For example, suppose $\phi(x) = x^2$ for all $x \geq 0$. If ρ has finite variance, then (B1) and (B2) are satisfied, and $\mathbf{m}_\rho^\phi$ is the mean of the distribution ρ. Indeed, a straightforward computation yields $\Phi_\rho(\mathbf{x}) := \mathrm{var}[\rho] + \left\|\mathbf{x} - \mathbf{m}_\rho^\phi\right\|^2$ for any $\mathbf{x}$ in $\mathbb{R}^N$.[18] Thus, condition (B3) is also satisfied. Thus, ρ is ϕ-balanced if and only if the mean of ρ lies in every median hyperplane of ρ. In particular:

---

[15] See e.g. Hinich and Munger (1997) or Enelow and Hinich (2008) for introductions to this literature.

[16] A point $\mathbf{x}$ in $\mathbb{R}^N$ is in the *support* of ρ if $\rho[\mathcal{U}] > 0$ for any open set $\mathcal{U} \subseteq \mathbb{R}^N$ which contains $\mathbf{x}$. Thus, $\mathcal{H}_{\mathbf{v}}^{\rho}$ intersects the support of ρ if and only if $\rho[\mathcal{U}] > 0$ for any open set $\mathcal{U} \subseteq \mathbb{R}^N$ which contains $\mathcal{H}_{\mathbf{v}}^{\rho}$.

[17] This means that $\rho(\mathbf{x}) \rightarrow 0$ fast enough as $\|\mathbf{x}\| \rightarrow \infty$.

[18] This result is sometimes attributed to Christiaan Huygens.

- Any multivariate normal probability measure is ϕ-balanced. (*Proof:* Any one-dimensional projection of a normal probability measure is normal, and in a one-dimensional normal measure, the mean coincides with the median.)

- If ρ is a ϕ-balanced measure on $\mathbb{R}^N$, and $F : \mathbb{R}^N \longrightarrow \mathbb{R}^M$ is an affine transformation, then $F(\rho)$ is a ϕ-balanced measure on $\mathbb{R}^M$. (*Proof:* F maps the mean of ρ to the mean of $F(\rho)$. Meanwhile, the F-preimage of any median hyperplane of $F(\rho)$ is a median hyperplane of ρ.)

- If $N = 1$, then ρ is ϕ-balanced if ρ has finite variance and is symmetrically distributed about some point m contained in the support of ρ. (For example, a uniform distribution on an interval is ϕ-balanced. So is the Laplace double-exponential distribution.)

More generally, the next result says that most rotationally symmetric probability measures are ϕ-balanced.

Proposition 10 *Let ρ be any probability measure on $\mathbb{R}$ which is symmetrically distributed about some point $\mathbf{m}$ in the support of ρ. Or, let $N \geq 2$, and let ρ be any probability measure on $\mathbb{R}^N$ which is rotationally symmetric around some point $\mathbf{m}$ in $\mathbb{R}^N$. Then for every strictly convex increasing function $\phi : [0, \infty) \longrightarrow \mathbb{R}$ satisfying* (B1)*, the measure ρ is ϕ-balanced, with $\mathbf{m}_\rho^\phi = \mathbf{m}$.*

Our last result says that, if any ϕ-balanced measure is used to generate a random collection of ideal points, which in turn is used to obtain a profile of distance-based utility functions, then this utility profile will be reasonable, with very high probability.

Proposition 11 *Let $\phi : [0, \infty) \longrightarrow \mathbb{R}$ be a convex increasing function, and let ρ be a ϕ-balanced probability measure on $\mathbb{R}^N$ with ϕ-median point $\mathbf{m}_\rho^\phi$. Let $\mathcal{A} \subset \mathbb{R}^N$ be a finite set of alternatives, such that $\|\mathbf{a} - \mathbf{m}_\rho^\phi\| \neq \|\mathbf{b} - \mathbf{m}_\rho^\phi\|$ for any distinct $\mathbf{a}, \mathbf{b}$ in $\mathcal{A}$. Finally, let $\{\mathbf{x}_i\}_{i \in \mathcal{I}}$ be a set of independent ρ-random points in $\mathbb{R}^N$. For every voter i in $\mathcal{I}$, suppose her utility function $u_i : \mathcal{A} \longrightarrow \mathbb{R}$ is given by $u_i(\mathbf{a}) = -\phi(\|\mathbf{a} - \mathbf{x}_i\|)$ for all $\mathbf{a}$ in $\mathcal{A}$. Then*

$$\lim_{I \to \infty} \operatorname{Prob}\Big(\textit{The utility profile } \{u_i\}_{i \in \mathcal{I}} \textit{ is reasonable}\Big) \;=\; 1.$$

For example, suppose $\phi(x) = x^2$, so that $u_i(\mathbf{a}) = -\|\mathbf{a} - \mathbf{x}_i\|^2$ for every voter i in $\mathcal{I}$ and every alternative $\mathbf{a}$ in $\mathcal{A}$. (This is a very common spatial voting model.) If $\{\mathbf{x}_i\}_{i \in \mathcal{I}}$ are independent random points drawn from any multivariate normal distribution on $\mathbb{R}^N$, and $|\mathcal{I}|$ is sufficiently large, then Proposition 11 says that the utility profile $\{u_i\}_{i \in \mathcal{I}}$ will be reasonable, with very high probability. Thus, with very high probability, the Condorcet winner of such a profile will be the utilitarian optimum.

6.3 Almost-reasonable profiles

If a utility profile $\mathcal{U}$ is *not* reasonable, then Proposition 8 does not apply; there may be no Condorcet winner, and even if there is, the Condorcet winner is not guaranteed to be a utilitarian optimum. However, if $\mathcal{U}$ is "close" to reasonable, then a suitably chosen Condorcet-consistent voting rule may still have a high probability of selecting a utilitarian optimum. For example, consider the ***Copeland rule***, which chooses the alternative with the highest Copeland score. (The ***Copeland score*** of an alternative a is defined as $\#\{b \in \mathcal{A};$ some majority prefers a over $b\} - \#\{b \in \mathcal{A};$ some majority prefers b over $a\}$.) Suppose that, for every $a, b \in \mathcal{A}$, there is a small probability that the profile $\mathcal{U}$ will fail to be (a, b)-reasonable, and that this probability is decreasing as a function of the average utility gap between a and b (as suggested by Example 7). Also suppose that these reasonability failures are independent random variables. Then the Copeland score of each alternative should be a good estimator of the "true" ranking of that alternative by the utilitarian social welfare order. Thus, the Copeland winner should either be optimal or close-to-optimal with respect to the utilitarian social welfare order. By a similar argument, the ordering of $\mathcal{A}$ determined by the Slater rule should be a good estimate of the ordering of $\mathcal{A}$ determined by the utilitarian social welfare order. These are interesting questions for future research.

7 Conclusion

This paper has neglected strategic voting and implementation issues. That is because these issues have already been adequately addressed in other literature. For example, the Condorcet winner (and thus, the utilitarian optimum, under the hypotheses of Proposition 8) is the outcome of sophisticated voting under *any* binary voting agenda (Miller, 1977, Proposition 8').[19] Meanwhile, the central premise of our results in Sections 2 to 5 is that it is enough to obtain an arbitrarily high *probability* of selecting a utilitarian optimum, rather than certainty. This is exactly the same premise as the theory of *virtual implementation* introduced by Matsushima (1988) and Abreu and Sen (1991). Virtual implementation is an extremely powerful and versatile implementation technology. For example, if the voters have complete information about one another, then *any* social choice rule can be virtually implemented in Nash equilibrium (Abreu and Sen, 1991) or iterated undominated strategies (Abreu and Matsushima, 1992). Even with incomplete information, a very large class of social choice rules can be virtually implemented in Bayesian Nash equilibrium (Serrano and Vohra, 2005), or even robustly virtually implemented (Artemov et al., 2013). Since virtual implementation is the implementation technology most suited to the probabilistic approach taken in this paper, we consider the implementation problem to be essentially solved, for our purposes.

According to conventional wisdom, utilitarianism is a nice idea in theory, but totally impossible to achieve in practice. The results of this paper suggest the opposite conclusion:

[19] See also Bag et al. (2009) and Horan (2013) for more extensive analyses of the implementation of Condorcet-consistent rules via pairwise voting agendas and other multistage elimination procedures.

utilitarianism is not merely possible, but actually fairly easy to achieve —at least as long as we are willing to tolerate a small amount of inefficiency, and as long as we have a sufficiently large population conforming to certain statistical regularities. These statistical regularities are the key assumption, of course. The results on scoring rules and approval voting in Section 4 and Section 5 require the statistical distribution of voters' utility functions to have a fairly specific form. But the Condorcet results of Section 6 are applicable to a much broader class of utility distributions. Which class of probability distributions (if any) best describes the statistical distribution of utility functions in society? Is the class of distributions the same from one social decision to the next? These questions must be answered empirically. If the results of these empirical investigations are affirmative, then "statistical" utilitarianism is indeed feasible. But is it desirable? This is a question of political philosophy.

Acknowledgements. I am grateful to Rohan Dutta, Ori Heffetz, Sean Horan, Jérôme Lang, Christophe Muller, Matías Núñez, and Clemens Puppe for useful discussions and helpful comments on earlier versions of this paper. This research was supported by NSERC grant #262620-2008.

Appendix: Proofs

Lemma A1 *Let $\{u_i\}_{i\in\mathcal{I}}$, $\{v_i\}_{i\in\mathcal{I}}$, and $\{c_i\}_{i\in\mathcal{I}}$ be as in Theorem 1. Let $\delta > 0$, let $p > 0$, and suppose $I \geq \overline{I}(\delta, p)$. For all $a \in \mathcal{A}$, we have* $\text{Prob}\left[|V_\mathcal{I}(a) - U_\mathcal{I}(a)| > \frac{\delta}{2}\right] < \frac{p}{2}$.

Proof. For all $a \in \mathcal{A}$, the quantity $U_\mathcal{I}(a) - V_\mathcal{I}(a)$ is a random variable. We will first compute its expected value and the variance of its distribution.

Claim 1: *For all $a \in \mathcal{A}$, $\mathbb{E}[U_\mathcal{I}(a) - V_\mathcal{I}(a)] = 0$.*

Proof. For all $a \in \mathcal{A}$,

$$\begin{aligned} U_\mathcal{I}(a) - V_\mathcal{I}(a) &\underset{(\diamond)}{=\!=} \frac{1}{I}\sum_{i\in\mathcal{I}} c_i\, u_i(a) - \frac{1}{I}\sum_{i\in\mathcal{I}} v_i(a) \\ &\underset{(\ddagger)}{=\!=} \frac{1}{I}\sum_{i\in\mathcal{I}} c_i\, u_i(a) - \frac{1}{I}\sum_{i\in\mathcal{I}} \Big(u_i(a) + \epsilon_i(a)\Big) \\ &= \frac{1}{I}\sum_{i\in\mathcal{I}} \Big((c_i - 1)\, u_i(a) - \epsilon_i(a)\Big), \end{aligned} \tag{A1}$$

where $(*)$ is by defining equations (1) and (2), and $(\dagger)$ is by assumption (U2). Thus,

$$\begin{aligned}
\mathbb{E}[U_{\mathcal{I}}(a) - V_{\mathcal{I}}(a)] &= \frac{1}{I}\sum_{i\in\mathcal{I}} \mathbb{E}\left[(c_i - 1)\, u_i(a) - \epsilon_i(a)\right] \\
&= \frac{1}{I}\sum_{i\in\mathcal{I}} \Big(\mathbb{E}[(c_i - 1)\, u_i(a)] - \mathbb{E}[\epsilon_i(a)]\Big) \\
&\underset{(*)}{=} \frac{1}{I}\sum_{i\in\mathcal{I}} \Big(u_i(a)\cdot\mathbb{E}[c_i - 1] + 0\Big) \quad \underset{(\dagger)}{=} \quad \frac{1}{I}\sum_{i\in\mathcal{I}} u_i(a)\cdot 0 \;=\; 0.
\end{aligned}$$

as desired. Here, $(*)$ is because, $u_i(a)$ is a constant and $\mathbb{E}[\epsilon_i(a)] = 0$ for all $i \in \mathcal{I}$, while $(\dagger)$ is because $\mathbb{E}[c_i] = 1$ for all $i \in \mathcal{I}$. $\diamond$ `Claim 1`

Claim 2: *For all* $a \in \mathcal{A}$, $\mathrm{var}[U_{\mathcal{I}}(a) - V_{\mathcal{I}}(a)] \leq \dfrac{M^2\cdot\sigma_c^2 + \sigma_\epsilon^2}{I}$.

Proof. For all $a \in \mathcal{A}$,

$$\begin{aligned}
\mathrm{var}[U_{\mathcal{I}}(a) - V_{\mathcal{I}}(a)] &\underset{(\dagger)}{=} \mathrm{var}\left[\frac{1}{I}\sum_{i\in\mathcal{I}}\Big((c_i - 1)\, u_i(a) - \epsilon_i(a)\Big)\right] \\
&= \frac{1}{I^2}\mathrm{var}\left[\sum_{i\in\mathcal{I}}\Big((c_i - 1)\, u_i(a) - \epsilon_i(a)\Big)\right] \\
&\underset{(*)}{=} \frac{1}{I^2}\sum_{i\in\mathcal{I}}\Big(\mathrm{var}\left[(c_i - 1)\, u_i(a)\right] + \mathrm{var}\left[\epsilon_i(a)\right]\Big) \\
&= \frac{1}{I^2}\sum_{i\in\mathcal{I}}\Big(u_i(a)^2\cdot\mathrm{var}\left[c_i\right] + \mathrm{var}\left[\epsilon_i(a)\right]\Big) \quad \underset{(\diamond)}{\leq} \quad \frac{1}{I^2}\sum_{i\in\mathcal{I}}\Big(u_i(a)^2\cdot\sigma_c^2 + \sigma_\epsilon^2\Big) \\
&= \frac{\sigma_c^2}{I}\left(\frac{1}{I}\sum_{i\in\mathcal{I}} u_i(a)^2\right) + \frac{1}{I^2}\left(\sum_{i\in\mathcal{I}}\sigma_\epsilon^2\right) \quad \underset{(\ddagger)}{\leq} \quad \frac{\sigma_c^2\, M^2}{I} + \frac{\sigma_\epsilon^2}{I},
\end{aligned}$$

as desired. Here, $(\dagger)$ is by equation (A1), while $(*)$ is because $\{u_i(a)\}_{i\in\mathcal{I}}$ are constants and $\{\epsilon_i(a)\}_{i\in\mathcal{I}} \cup \{c_i\}_{i\in\mathcal{I}}$ are all jointly independent random variables. Finally, $(\diamond)$ is by assumptions (U1) and (U2), while $(\ddagger)$ is by (U3). $\diamond$ `Claim 2`

In conclusion, for all $a \in \mathcal{A}$, we have

$$\mathrm{Prob}\left[|V_{\mathcal{I}}(a) - U_{\mathcal{I}}(a)| > \frac{\delta}{2}\right] \quad \underset{(*)}{<} \quad \frac{\mathrm{var}[U_{\mathcal{I}}(a) - V_{\mathcal{I}}(a)]}{(\delta/2)^2} \quad \underset{(\dagger)}{\leq} \quad 4\,\frac{M^2\cdot\sigma_c^2 + \sigma_\epsilon^2}{I\,\delta^2} \quad \underset{(\diamond)}{\leq} \quad \frac{p}{2},$$

as desired. Here, $(*)$ is by Claim 1 and Chebyshev's inequality, $(\dagger)$ is by Claim 2, and $(\diamond)$ is because $I \geq \overline{I}(\delta, p)$ by hypothesis. □

Proof of Theorem 1. Let $b^* \in \operatorname{argmax}_{\mathcal{A}}(U_{\mathcal{I}})$; thus, $U_{\mathcal{I}}(b^*) = U^*$. Then

$$\begin{aligned} 0 &\leq U^* - U_{\mathcal{I}}(a_V^*) \;=\; U_{\mathcal{I}}(b^*) - U_{\mathcal{I}}(a_V^*) \\ &= U_{\mathcal{I}}(b^*) - V_{\mathcal{I}}(b^*) + V_{\mathcal{I}}(b^*) - V_{\mathcal{I}}(a_V^*) + V_{\mathcal{I}}(a_V^*) - U_{\mathcal{I}}(a_V^*) \\ &\leq U_{\mathcal{I}}(b^*) - V_{\mathcal{I}}(b^*) \;+\; V_{\mathcal{I}}(a_V^*) - U_{\mathcal{I}}(a_V^*), \end{aligned}$$

where the last step is because $V_{\mathcal{I}}(b^*) - V_{\mathcal{I}}(a_V^*) \leq 0$ because $a_V^* \in \operatorname{argmax}_{\mathcal{A}}(V_{\mathcal{I}})$. Thus, if $U^* - U_{\mathcal{I}}(a_V^*) > \delta$, then either $U_{\mathcal{I}}(b^*) - V_{\mathcal{I}}(b^*) > \frac{\delta}{2}$ or $V_{\mathcal{I}}(a_V^*) - U_{\mathcal{I}}(a_V^*) > \frac{\delta}{2}$. Thus,

$$\begin{aligned} \operatorname{Prob}\left[U^* - U_{\mathcal{I}}(a_V^*) > \delta\right] &\leq \operatorname{Prob}\left[|V_{\mathcal{I}}(b^*) - U_{\mathcal{I}}(b^*)| > \frac{\delta}{2}\right] + \operatorname{Prob}\left[|V_{\mathcal{I}}(a_V^*) - U_{\mathcal{I}}(a_V^*)| > \frac{\delta}{2}\right] \\ &< \frac{p}{2} + \frac{p}{2} \;=\; p, \end{aligned}$$

where the second inequality is by two invocations of Lemma A1. □

Proof of Proposition 3. Let $v_i(a) := u_i(a)$ for all $i \in \mathcal{I}$ and $a \in \mathcal{A}$. Thus, if $V_{\mathcal{I}}$ is defined as in formula (2), then $V_{\mathcal{I}}(a) = \frac{1}{I}\, RU(a)$ for all $a \in \mathcal{A}$; thus, maximizing RU is equivalent to maximizing $V_{\mathcal{I}}$. The claim now follows from Theorem 1, by setting $\sigma_\epsilon^2 := 0$ (since we assume $\epsilon_i(a) = 0$ for all i and a) setting $M := 1$ (by definition of $\{u_i\}_{i\in\mathcal{I}}$). □

Proof of Proposition 4. We will apply Theorem 1. For all $a \in \mathcal{A}$ and $i \in \mathcal{I}$, define $v_i(a) := s_k^N$ if a is ranked kth from the bottom by $\succ_i$. Define $V_{\mathcal{I}} : \mathcal{A} \longrightarrow \mathbb{R}$ as in equation (2); then clearly $a_{\text{scr}}^* = \operatorname{argmax}_{\mathcal{A}}(V_{\mathcal{I}})$.

If we take a random sample of N independent random variables drawn from ρ, and compute the order statistics of this sample, then we get N new random variables. Let $\sigma_1^2, \ldots, \sigma_N^2$ denote their variances. Since ρ has finite variance, it is easy to check that $\sigma_1^2, \ldots, \sigma_N^2$ are all finite. Define $\sigma_\epsilon^2 := \max\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2\}$. Also, let $S^2 := \max\{(s_1^N)^2, (s_2^N)^2, \ldots, (s_N^N)^2\}$, and choose any $M > \sqrt{S^2 + \sigma_\epsilon^2}$.

Now, let $\{c_i\, u_i\}_{i\in\mathcal{I}}$ be a $(\mathcal{P}, \rho)$-random utility profile.

Claim 1: $\lim_{I\to\infty} \operatorname{Prob}\Big(M \textit{ and the profile } \{u_i\}_{i\in\mathcal{I}} \textit{ satisfy condition } \text{(U3)}\Big) \;=\; 1.$

Proof. Fix $a \in \mathcal{A}$. For all $i \in \mathcal{I}$, if a is ranked kth from the bottom by $\succ_i$, then $u_i(a)$ is a random variable with mean s_k^N and variance σ_k^2. Thus,

$$\mathbb{E}[u_i^2(a)] \;=\; (s_k^N)^2 + \sigma_k^2 \;\leq\; S^2 + \sigma_\epsilon^2 \;<\; M^2. \tag{A2}$$

Thus, for any $a \in \mathcal{A}$, the sum $\frac{1}{I}\sum_{i\in\mathcal{I}} u_i(a)^2$ is an average of I independent random variables, each with expected value smaller than M^2, by inequality (A2). Thus, regardless of how the preferences $\{\succ_i\}_{i\in\mathcal{I}}$ are obtained, the Law of Large Numbers implies that

$$\lim_{I\to\infty} \operatorname{Prob}\left[\frac{1}{I}\sum_{i\in\mathcal{I}} u_i(a)^2 \;<\; M^2\right] \;=\; 1.$$

Thus, since $\mathcal{A}$ is finite, the claim follows. ◇ `Claim 1`

For all $i \in \mathcal{I}$ and $a \in \mathcal{A}$, define $\epsilon_i(a) := v_i(a) - u_i(a)$. Then for each $a \in \mathcal{A}$, the set $\{\epsilon_i(a)\}_{i\in\mathcal{I}}$ is a set of independent random variables, each with expected value 0 and variance at most σ_ϵ^2. Thus, the collection $\{\epsilon_i\}_{i\in\mathcal{I}}$ satisfies condition (U2). Meanwhile, the collection $\{c_i\}_{i\in\mathcal{I}}$ satisfies condition (U1) by hypothesis. Thus, combining Claim 1 with Theorem 1 yields the desired limit equation (5). To obtain the more precise estimate of convergence speed, we need the next observation.

Claim 2: *Suppose the fourth moment of ρ is finite. Then there is some $C_1 > 0$ (determined by ρ) such that, for any $p \in (0,1)$, if $I > C_1/p$, then*

$$\text{Prob}\Big(M \textit{ and } \{u_i\}_{i\in\mathcal{I}} \textit{ violate condition } \text{(U3)}\Big) \quad < \quad \frac{p}{2}.$$

Proof. If the fourth moment of ρ is finite, then there is some $C' > 0$ such that for any $a \in \mathcal{A}$, the fourth moments of each of the random variables $\{u_i(a)\}_{i\in\mathcal{I}}$ is less than C'. In other words, the second moments of each of the random variables $\{u_i(a)^2\}_{i\in\mathcal{I}}$ is less than C'. This implies that there is some $C'' > 0$ such that the variance of each of $\{u_i(a)^2\}_{i\in\mathcal{I}}$ is less than C''. Also, these random variables are independent. Thus,

$$\text{var}\left[\frac{1}{I}\sum_{i\in\mathcal{I}} u_i(a)^2\right] \quad < \quad \frac{C''}{I}. \tag{A3}$$

Next, inequality (A2) says each of $\{u_i(a)^2\}_{i\in\mathcal{I}}$ has expected value less than M^2. Thus,

$$\mathbb{E}\left[\frac{1}{I}\sum_{i\in\mathcal{I}} u_i(a)^2\right] \quad < \quad M^2. \tag{A4}$$

Thus, Chebyshev's inequality and inequalities (A3) and (A4) imply that there is some $C_1 > 0$ (determined by C'') such that, for any $p > 0$, if $I > C_1/p$, then

$$\text{Prob}\left[\frac{1}{I}\sum_{i\in\mathcal{I}} u_i(a)^2 \;>\; M^2\right] \quad < \quad \frac{p}{2|\mathcal{A}|}. \tag{A5}$$

Now, if the profile $\{u_i\}_{i\in\mathcal{I}}$ and M to violate condition (U3), then $\frac{1}{I}\sum_{i\in\mathcal{I}} u_i(a)^2 > M^2$ for some $a \in \mathcal{A}$. Thus, adding together $|\mathcal{A}|$ copies of inequality (A5) proves the claim. $\diamond$ `Claim 2`

For any $\delta > 0$ and $p > 0$, let $\overline{I}(\delta, p)$ be as in equation (3). Finally, define $C_2 := 16(M^2\,\sigma_c^2 + \sigma_\epsilon^2)$. Thus, for any $p, \delta \in (0,1)$, if $I > C_2/p\,\delta^2$, then $I > \overline{I}(\delta, p/2)$, so that Theorem 1 says

$$\text{Prob}\left[U_{\mathcal{I}}(a^*_{\text{scr}}) < U^* - \delta \,\Big|\, M \text{ and } \{u_i\}_{i\in\mathcal{I}} \text{ satisfy (U3)}\right] \quad < \quad \frac{p}{2}. \tag{A6}$$

If $I > C_1/p$ also, then Claim 2 applies. This, together with inequality (A6), implies that $\text{Prob}\,[U_{\mathcal{I}}(a^*_{\text{scr}}) < U^* - \delta] < \frac{p}{2} + \frac{p}{2} = p$, as desired. □

Proof of Proposition 6. We will apply Theorem 1, as in the proof of Proposition 4. Let $\overline{g}$ be the mean value of γ and let $\overline{b}$ be the mean value of β; then $\overline{b} < 0 < \overline{g}$. (Since γ and β have finite variance, the means $\overline{b}$ and $\overline{g}$ are well-defined.) For all $a \in \mathcal{A}$ and $i \in \mathcal{I}$, define $v_i(a) := \overline{b}$ if $a \in \mathcal{B}_i$ and $v_i(a) := \overline{g}$ if $a \in \mathcal{G}_i$. Now define $V_{\mathcal{I}} : \mathcal{A} \longrightarrow \mathbb{R}$ as in equation (2); then clearly $a^*_{\text{appr}} = \text{argmax}_{\mathcal{A}}(V_{\mathcal{I}})$.

Define $\sigma_\epsilon^2 := \max\{\text{var}[\gamma], \text{var}[\beta]\}$; then σ_ϵ^2 is finite. Also, let $S^2 := \max\{\overline{g}^2, \overline{b}^2\}$, and choose any $M > \sqrt{S^2 + \sigma_\epsilon^2}$. Now, let $\{c_i\, u_i\}_{i \in \mathcal{I}}$ be a $(\mathfrak{G}, \gamma, \beta)$-random utility profile. By an argument very similar to the proof of Claim 1 in the proof of Proposition 4, we obtain:

$$\lim_{I \to \infty} \text{Prob}\Big(M \text{ and the profile } \{u_i\}_{i \in \mathcal{I}} \text{ satisfy condition (U3)}\Big) \quad = \quad 1. \qquad \text{(A7)}$$

For all $i \in \mathcal{I}$ and $a \in \mathcal{A}$, define $\epsilon_i(a) := v_i(a) - u_i(a)$. Then for each $a \in \mathcal{A}$, the set $\{\epsilon_i(a)\}_{i \in \mathcal{I}}$ is a set of independent random variables, each with expected value 0 and variance at most σ_ϵ^2. Thus, the collection $\{\epsilon_i\}_{i \in \mathcal{I}}$ satisfies condition (U2). Meanwhile, the collection $\{c_i\}_{i \in \mathcal{I}}$ satisfies condition (U1) by hypothesis. Thus, combining limit equation (A7) with Theorem 1 yields the desired limit equation (6).

Claim 1: *Suppose the fourth moments of γ and β are finite. Then there is some $C_1 > 0$ (determined by ρ) such that, for any $p \in (0,1)$, if $I > C_1/p$, then*

$$\text{Prob}\Big(M \textit{ and } \{u_i\}_{i \in \mathcal{I}} \textit{ violate condition } \text{(U3)}\Big) \quad < \quad \frac{p}{2}.$$

The proof is very similar to the proof of Claim 2 in the proof of Proposition 4.

Now, for any $\delta > 0$ and $p \in (0,1)$, let $\overline{I}(\delta, p)$ be as in equation (3). Finally, define $C_2 := 16(M^2\, \sigma_c^2 + \sigma_\epsilon^2)$. Thus, for any $p, \delta \in (0,1)$, if $I > C_2/p\,\delta^2$, then $I > \overline{I}(\delta, p/2)$, so that Theorem 1 says

$$\text{Prob}\Big[U_{\mathcal{I}}(a^*_{\text{appr}}) < U^* - \delta \,\Big|\, M \text{ and } \{u_i\}_{i \in \mathcal{I}} \text{ satisfy (U3)}\Big] \quad < \quad \frac{p}{2}. \qquad \text{(A8)}$$

If $I > C_1/p$ also, then Claim 1 applies. This, together with inequality (A6), implies that $\text{Prob}\left[U_{\mathcal{I}}(a^*_{\text{appr}}) < U^* - \delta\right] < \frac{p}{2} + \frac{p}{2} = p$, as desired. □

Proof of Proposition 9. For any distinct $a, b, \in \mathcal{A}$, recall that $\rho_{a,b}$ is the distribution of $x_a - x_b$, where $\mathbf{x}$ is a ρ-random variable. Thus, $\rho_{a,b}$ has finite variance, because ρ has finite variance. Let $m_{a,b}$ be the mean value of $\rho_{a,b}$; then $m_{a,b} \neq 0$, because ρ is reasonable. Let $p_{a,b} := \rho_{a,b}(-\infty, 0)$ if $m_{a,b} > 0$, and let $p_{a,b} := \rho_{a,b}(0, \infty)$ if $m_{a,b} < 0$. (Equivalently, $p_{a,b} := \rho\{\mathbf{x} \in \mathbb{R}^{\mathcal{A}};\ \text{sign}(x_a - x_b) = -\text{sign}(m_{ab})\}$.) Then $p_{a,b} < \frac{1}{2}$, because $\text{sign}(m_{a,b}) = \text{sign}(\text{median}[\rho_{a,b}])$, because ρ is reasonable. Let $p := \max\{p_{a,b};\ a, b \in \mathcal{A}\}$; then $p < \frac{1}{2}$ because $p_{a,b} < \frac{1}{2}$ for all $a, b \in \mathcal{A}$, and $\mathcal{A}$ is finite. It follows that $p(1-p) < \frac{1}{4}$ (because the function $f(x) = x(1-x)$ has a unique maximum at $x = \frac{1}{2}$, and $f(\frac{1}{2}) = \frac{1}{4}$). Thus, if we define $q := 2\sqrt{p(1-p)}$, then $q < 1$. (For example, if $p = 0.4$, then $q = 2\sqrt{0.4 \cdot 0.6} \approx 0.98$.) Let $A := |\mathcal{A}|$, and without loss of generality, suppose $\mathcal{A} = \{1, 2, \ldots, A\}$. For all $i \in \mathcal{I}$, let $u_i := (u_1^i, u_2^i, \ldots, u_A^i) \in \mathbb{R}^A$ be the utility function of voter i (a ρ-random vector). For any $a < b \in \mathcal{A}$, let $\mathcal{U}_{a,b} := \{u_a^i - u_b^i\}_{i \in \mathcal{I}}$ (a collection of I independent, real-valued random variables).

Claim 1: *If I is large enough, then for all distinct $a, b \in \mathcal{A}$, we have*

$$\text{Prob}\Big[\text{sign}\left(\text{median}[\mathcal{U}_{a,b}]\right) \neq \text{sign}\left(\text{median}[\rho_{a,b}]\right)\Big] \quad < \quad 3\sqrt{I}\, q^I.$$

Proof. Without loss of generality, suppose $\text{median}(\rho_{a,b}) > 0$. Let J be the smallest integer greater than $I/2$. (That is: $J := (I+1)/2$ if I is odd, whereas $J := (I/2) + 1$ if I is even.) Now, $|\mathcal{U}_{a,b}| = I$, so[20]

$$\Big(\text{median}(\mathcal{U}_{a,b}) < 0\Big) \iff \Big(\text{at least } J \text{ elements of } \mathcal{U}_{a,b} \text{ are in } (-\infty, 0)\Big). \tag{A9}$$

For any $i \in \mathcal{I}$, $\text{Prob}[u_a^i - u_b^i < 0] = p_{a,b}$, and these are independent random events. Thus, for any $n \in [0 \ldots 2J]$,

$$\begin{aligned}
&\text{Prob}\Big(\text{exactly } n \text{ elements of } \mathcal{U}_{a,b} \text{ are in } (-\infty, 0)\Big) \\
&\quad = \binom{2J}{n} p_{a,b}^n (1 - p_{a,b})^{2J-n} \underset{(*)}{\leq} \binom{2J}{n} p^n (1-p)^{2J-n}. \qquad \text{Thus,} \\
&\text{Prob}\Big(\text{at least } J \text{ elements of } \mathcal{U}_{a,b} \text{ are in } (-\infty, 0)\Big) \\
&\quad \leq \sum_{n=J}^{2J} \binom{2J}{n} p^n (1-p)^{2J-n} \underset{(\diamond)}{\leq} \sum_{n=J}^{2J} \binom{2J}{J} p^J (1-p)^J \\
&\quad = J \binom{2J}{J} p^J (1-p)^J \underset{(@)}{<} J \frac{(2J)!}{(J!)^2} \Big(p(1-p)\Big)^{I/2} \\
&\quad \underset{(\dagger)}{=} J \frac{(2J)!}{(J!)^2} \left(\frac{q}{2}\right)^I \underset{(\ddagger)}{\approx} J \sqrt{\frac{2}{\pi I}} \cdot 2^{I+2} \left(\frac{q}{2}\right)^I \\
&\quad = 4J \sqrt{\frac{2}{\pi I}} \cdot q^I \underset{(\P)}{<} 3I \frac{q^I}{\sqrt{I}} = 3\sqrt{I}\, q^I.
\end{aligned} \tag{A10}$$

Here, $(*)$ is because $0 \leq p_{a,b} \leq p < \frac{1}{2}$, and the function $f(x) = x^J(1-x)^J$ is increasing on the interval $\left[0, \frac{1}{2}\right]$. Next, $(\diamond)$ is because $p < \frac{1}{2}$, so the mode of the p-binomial distribution on $[0 \ldots 2J]$ occurs at some $n < J$, so that $\binom{2J}{n} p^n (1-p)^{2J-n} < \binom{2J}{J} p^J (1-p)^J$ for all $n \in [J \ldots 2J]$. Next, $(@)$ is because $J > I/2$, and $(\dagger)$ is because $\sqrt{p(1-p)} = q/2$, so $[p(1-p)]^{I/2} = (\sqrt{p(1-p)})^I = (q/2)^I$. Next, $(\ddagger)$ is via Stirling's approximation of the factorial, which says $n! \approx \sqrt{2\pi\, n}\,(n/e)^n$ as $n \to \infty$. Thus, if J is large enough, then

$$\frac{(2J)!}{(J!)^2} \approx \frac{\sqrt{2\pi\, 2J}\,(2J/e)^{2J}}{[\sqrt{2\pi\, J}\,(J/e)^J]^2} = \frac{2^{2J}}{\sqrt{\pi J}} < \frac{2^{I+2}}{\sqrt{\pi I/2}} = \sqrt{\frac{2}{\pi I}} \cdot 2^{I+2}.$$

[20] See footnote 14 for how to interpret the left-hand side of statement (A9) when I is even.

Finally, (¶) is because $2J \leq I+2$ and $2 \cdot \sqrt{2/\pi} \approx 2.26$, so $4J \cdot \sqrt{2/\pi} \approx (2.26) \cdot (I+2) < 3I$, if I is large enough.

We have assumed $\text{median}(\rho_{a,b}) > 0$. Thus, combining statement (A9) and inequality (A10) yields the claim. ◇ `Claim 1`

Let $C := \max\{\frac{\text{var}[\rho_{a,b}]}{m_{a,b}^2};\ \ a, b \in \mathcal{A}\}$; then $C < \infty$ because $\text{var}[\rho_{a,b}] < \infty$ and $m_{a,b} \neq 0$ for all distinct $a, b \in \mathcal{A}$, and $|\mathcal{A}|$ is finite. For all $a, b \in \mathcal{A}$, Chebyshev's inequality implies that

$$\text{Prob}\left[\text{sign}\left(\text{mean}[\mathcal{U}_{a,b}]\right) \neq \text{sign}(m_{a,b})\right] \quad < \quad \frac{\text{var}[\rho_{a,b}]}{I\, m_{a,b}^2} \quad \leq \quad \frac{C}{I}. \tag{A11}$$

Now, $\text{sign}\left(\text{median}[\rho_{a,b}]\right) = \text{sign}(m_{a,b})$, because ρ is reasonable. Thus, if $\text{sign}\left(\text{median}[\mathcal{U}_{a,b}]\right) = \text{sign}\left(\text{median}[\rho_{a,b}]\right)$ and $\text{sign}\left(\text{mean}[\mathcal{U}_{a,b}]\right) = \text{sign}(m_{a,b})$, then $\text{sign}\left(\text{median}[\mathcal{U}_{a,b}]\right) = \text{sign}\left(\text{mean}[\mathcal{U}_{a,b}]\right)$. Conversely, if $\text{sign}\left(\text{median}[\mathcal{U}_{a,b}]\right) \neq \text{sign}\left(\text{mean}[\mathcal{U}_{a,b}]\right)$, then either $\text{sign}\left(\text{median}[\mathcal{U}_{a,b}]\right) \neq \text{sign}\left(\text{median}[\rho_{a,b}]\right)$ or $\text{sign}\left(\text{mean}[\mathcal{U}_{a,b}]\right) \neq \text{sign}(m_{a,b})$. Thus,

$$\begin{aligned}
&\text{Prob}\left[\text{sign}\left(\text{median}[\mathcal{U}_{a,b}]\right) \neq \text{sign}\left(\text{mean}[\mathcal{U}_{a,b}]\right)\right] \\
&\leq \ \text{Prob}\left[\text{sign}\left(\text{median}[\mathcal{U}_{a,b}]\right) \neq \text{sign}\left(\text{median}[\rho_{a,b}]\right) \ \text{or}\ \text{sign}\left(\text{mean}[\mathcal{U}_{a,b}]\right) \neq \text{sign}(m_{a,b})\right] \\
&\leq \ \text{Prob}\left[\text{sign}\left(\text{median}[\mathcal{U}_{a,b}]\right) \neq \text{sign}\left(\text{median}[\rho_{a,b}]\right)\right] + \text{Prob}\left[\text{sign}\left(\text{mean}[\mathcal{U}_{a,b}]\right) \neq \text{sign}(m_{a,b})\right] \\
&\underset{(*)}{=} \ 3\sqrt{I}\, q^I + \frac{C}{I},
\end{aligned} \tag{A12}$$

where $(*)$ is by Claim 1 and inequality (A11). Thus,

$$\begin{aligned}
&\text{Prob}\left(\text{the profile } \{u_i\}_{i \in \mathcal{I}} \text{ is not reasonable}\right) \\
&= \ \text{Prob}\left(\text{sign}\left(\text{median}[\mathcal{U}_{a,b}]\right) \neq \text{sign}\left(\text{mean}[\mathcal{U}_{a,b}]\right) \text{ for some } a < b \in \mathcal{A}\right) \\
&\underset{(*)}{\leq} \ \sum_{a<b\in\mathcal{A}} \left(3\sqrt{I}\, q^I + \frac{C}{I}\right) \quad = \quad \frac{A(A-1)}{2}\left(3\sqrt{I}\, q^I + \frac{C}{I}\right) \quad \xrightarrow[I\to\infty]{(\dagger)} \quad 0,
\end{aligned}$$

as desired. Here, the inequality $(*)$ follows from inequality (A12), and the limit $(\dagger)$ is a straightforward application of l'Hospital's rule, because $0 < q < 1$. □

Proof of Proposition 10.

Claim 1: *Φ_ρ is strictly convex.*

Proof. For any $\mathbf{x} \in \mathbb{R}^N$, define $\phi_{\mathbf{x}} : \mathbb{R}^N \longrightarrow \mathbb{R}$ by setting $\phi_{\mathbf{x}}(\mathbf{y}) := \phi\left(\|\mathbf{x} - \mathbf{y}\|\right)$ for all $\mathbf{y} \in \mathbb{R}^N$. First observe that $\phi_{\mathbf{x}}$ is strictly convex. To see this, let $\mathbf{y}, \mathbf{z} \in \mathbb{R}^N$, and let $r \in (0, 1)$. Then

$$\begin{aligned}
\|r\,\mathbf{y} + (1-r)\,\mathbf{z} - \mathbf{x}\| \ &= \ \|r\,(\mathbf{y} - \mathbf{x}) + (1-r)\,(\mathbf{z} - \mathbf{x})\| \\
&\leq \ r\,\|\mathbf{y} - \mathbf{x}\| + (1-r)\,\|\mathbf{z} - \mathbf{x})\|,
\end{aligned} \tag{A13}$$

by the triangle inequality. Thus,

$$\begin{aligned}\phi_{\mathbf{x}}\left(r\,\mathbf{y}+(1-r)\,\mathbf{z}\right) &= \phi\left(\|r\,\mathbf{y}+(1-r)\,\mathbf{z}-\mathbf{x}\|\right) \underset{(*)}{\leq} \phi\left(r\,\|\mathbf{y}-\mathbf{x}\|+(1-r)\,\|\mathbf{z}-\mathbf{x})\|\right)\\ &\underset{(\dagger)}{<} r\,\phi\left(\|\mathbf{y}-\mathbf{x}\|\right)+(1-r)\,\phi\left(\|\mathbf{z}-\mathbf{x})\|\right)\\ &= r\,\phi_{\mathbf{x}}(\mathbf{y})+(1-r)\,\phi(\mathbf{z}), \qquad \text{as desired.} \end{aligned} \tag{A14}$$

Here, $(*)$ is by inequality (A13), because ϕ is increasing, while $(\dagger)$ is because ϕ is strictly convex.

Now, for any $\mathbf{y} \in \mathbb{R}^N$, the defining equation (8) says $\Phi_\rho(\mathbf{y}) = \int_{\mathbb{R}^N} \phi_{\mathbf{x}}(\mathbf{y})\ \mathrm{d}\rho[\mathbf{x}]$. Thus, for any $\mathbf{y}, \mathbf{z} \in \mathbb{R}^N$, and any $r \in (0, 1)$, we have

$$\begin{aligned}\Phi_\rho\left(r\,\mathbf{y}+(1-r)\,\mathbf{z}\right) &= \int_{\mathbb{R}^N} \phi_{\mathbf{x}}\left(r\,\mathbf{y}+(1-r)\,\mathbf{z}\right)\ \mathrm{d}\rho[\mathbf{x}]\\ &\underset{(*)}{<} \int_{\mathbb{R}^N} r\,\phi_{\mathbf{x}}(\mathbf{y})+(1-r)\,\phi_{\mathbf{x}}(\mathbf{z})\ \mathrm{d}\rho[\mathbf{x}]\\ &= r\int_{\mathbb{R}^N} \phi_{\mathbf{x}}(\mathbf{y})\ \mathrm{d}\rho[\mathbf{x}]\ +\ (1-r)\int_{\mathbb{R}^N} \phi_{\mathbf{x}}(\mathbf{z})\ \mathrm{d}\rho[\mathbf{x}]\\ &= r\,\Phi_\rho(\mathbf{y})+(1-r)\,\Phi_\rho(\mathbf{z}),\end{aligned}$$

as desired. Here, $(*)$ is by inequality (A13). ◇ Claim 1

Claim 2: $\mathbf{m}$ *is the unique global minimum of* Φ_ρ.

Proof. First suppose $N \geq 2$. Claim 1 implies that the global minimum of Φ_ρ is unique. But if ρ is rotationally symmetric around $\mathbf{m}$, then so is the function Φ_ρ. Thus, so is the set of global minima of Φ_ρ. Thus the (unique) global minimum must be at $\mathbf{m}$.

The argument in the case $N = 1$ is similar, except now "rotationally symmetric around $\mathbf{m}$" is changed to "symmetric under reflection across the point $\mathbf{m}$". ◇ Claim 2

Claim 3: *For every* $\mathbf{v} \in \mathbb{R}^N$, *the measure* ρ *has a unique* $\mathbf{v}$*-median hyperplane* $\mathcal{H}^\rho_{\mathbf{v}}$, *and* $\mathbf{m} \in \mathcal{H}^\rho_{\mathbf{v}}$.

Proof. We will handle the cases $N = 1$ and $N \geq 2$ separately.

In the case $N = 1$, a median "hyperplane" is just a median point of ρ (the vector $\mathbf{v}$ is irrelevant in this case). The theorem hypothesis states that ρ is symmetrically distributed about $\mathbf{m}$. Thus, $\mathbf{m}$ is a median point of ρ. But we also assumed that $\mathbf{m}$ is in the support of ρ; thus, $\mathbf{m}$ is the *only* median point of ρ.

Now suppose $N \geq 2$. If ρ is rotationally symmetric around $\mathbf{m}$, then so is support(ρ). Thus, support(ρ) can be written as a union of concentric spheres centred at $\mathbf{m}$. Now let $\mathbf{v} \in \mathbb{R}^N$ be any vector, and define

$$\begin{aligned}\mathcal{C}^-_{\mathbf{v}} &:= \left\{\mathbf{r} \in \mathbb{R}^N\ ;\ \mathbf{v} \bullet \mathbf{r} < \mathbf{v} \bullet \mathbf{m}\right\},\\ \mathcal{H}^\rho_{\mathbf{v}} &:= \left\{\mathbf{r} \in \mathbb{R}^N\ ;\ \mathbf{v} \bullet \mathbf{r} = \mathbf{v} \bullet \mathbf{m}\right\},\\ \text{and}\quad \mathcal{C}^+_{\mathbf{v}} &:= \left\{\mathbf{r} \in \mathbb{R}^N\ ;\ \mathbf{v} \bullet \mathbf{r} > \mathbf{v} \bullet \mathbf{m}\right\}.\end{aligned}$$

Thus, $\mathcal{H}_{\mathbf{v}}^{\rho}$ is the unique hyperplane in $\mathbb{R}^N$ orthogonal to $\mathbf{v}$ and containing $\mathbf{m}$. Note that the halfspace $\mathcal{C}_{\mathbf{v}}^{-}$ can be transformed into $\mathcal{C}_{\mathbf{v}}^{+}$ by rotating 180 degrees through any axis passing through $\mathbf{m}$. Since ρ is rotationally symmetric around $\mathbf{m}$, this implies that $\rho[\mathcal{C}_{\mathbf{v}}^{-}] = \rho[\mathcal{C}_{\mathbf{v}}^{+}]$; thus, $\mathcal{H}_{\mathbf{v}}^{\rho}$ is a $\mathbf{v}$-median hyperplane for ρ. However, we have already noted that support(ρ) is a union of concentric spheres centred at $\mathbf{m}$; thus, $\mathcal{H}_{\mathbf{v}}$ intersects support(ρ). Thus, $\mathcal{H}_{\mathbf{v}}^{\rho}$ is the *only* $\mathbf{v}$-median hyperplane for ρ. This argument works for any $\mathbf{v} \in \mathbb{R}^N$. ◇ `Claim 3`

By hypothesis, ρ satisfies condition (B1). Claim 2 implies that ρ satisfies conditions (B2) and (B3), while Claim 3 implies that it satisfies condition (B4). Thus, ρ is Φ-balanced. □

Proof of Proposition 11. Recall that $\mathcal{A} \subset \mathbb{R}^N$. Let $\mathbf{a}, \mathbf{b} \in \mathcal{A}$. Let $\mathbf{v} := \mathbf{b} - \mathbf{a}$, and define:

$$\begin{aligned} \mathcal{C}_{\mathbf{a}} &:= \left\{\mathbf{r} \in \mathbb{R}^N \; ; \; \|\mathbf{r}-\mathbf{a}\| < \|\mathbf{r}-\mathbf{b}\|\right\}, \\ \mathcal{H}_{\mathbf{a},\mathbf{b}} &:= \left\{\mathbf{r} \in \mathbb{R}^N \; ; \; \|\mathbf{r}-\mathbf{a}\| = \|\mathbf{r}-\mathbf{b}\|\right\}, \\ \text{and } \mathcal{C}_{\mathbf{b}} &:= \left\{\mathbf{r} \in \mathbb{R}^N \; ; \; \|\mathbf{r}-\mathbf{a}\| > \|\mathbf{r}-\mathbf{b}\|\right\}. \end{aligned}$$

Then $\mathcal{C}_{\mathbf{a}}$ and $\mathcal{C}_{\mathbf{b}}$ are two halfspaces separated by $\mathcal{H}_{\mathbf{a},\mathbf{b}}$, which is the hyperplane orthogonal to $\mathbf{v}$, and passing through the point $(\mathbf{a}+\mathbf{b})/2$.

Claim 1: *If* $\mathbf{m}_{\rho}^{\phi} \in \mathcal{C}_{\mathbf{a}}$, *then* $\lim_{I\to\infty} \text{Prob}\left(\textit{A majority of } \{u_i\}_{i\in\mathcal{I}} \textit{ prefer } \mathbf{a} \textit{ over } \mathbf{b}\right) = 1.$

Proof. Let $\mathcal{H}_{\mathbf{v}}^{I} \subset \mathbb{R}^N$ be any $\mathbf{v}$-median hyperplane of the collection $\{\mathbf{x}_i\}_{i\in\mathcal{I}}$ —that is, $\mathcal{H}_{\mathbf{v}}^{I}$ is a hyperplane in $\mathbb{R}^N$ orthogonal to $\mathbf{v}$, such that at least half the points in $\{\mathbf{x}_i\}_{i\in\mathcal{I}}$ lie either in $\mathcal{H}_{\mathbf{v}}^{I}$ or on one side of $\mathcal{H}_{\mathbf{v}}^{I}$, and at least half the points in $\{\mathbf{x}_i\}_{i\in\mathcal{I}}$ lie either in $\mathcal{H}_{\mathbf{v}}^{I}$ or on the other side of $\mathcal{H}_{\mathbf{v}}^{I}$. (Such a hyperplane may not be unique; if it is not unique, then just pick one arbitrarily.)

For any $i \in \mathcal{I}$, we have $u_i(\mathbf{a}) > u_i(\mathbf{b})$ if and only if $\mathbf{x}_i \in \mathcal{C}_{\mathbf{a}}$. It follows that

$$\Big(\text{A majority of } \{u_i\}_{i\in\mathcal{I}} \text{ prefer } \mathbf{a} \text{ over } \mathbf{b}\Big) \quad \Longleftrightarrow \quad \Big(\mathcal{H}_{\mathbf{v}}^{I} \subset \mathcal{C}_{\mathbf{a}}\Big). \tag{A15}$$

Let $\mathcal{H}_{\mathbf{v}}^{\rho}$ be the (unique) $\mathbf{v}$-median hyperplane of ρ; then condition (B4) says $\mathbf{m}_{\rho}^{\phi} \in \mathcal{H}_{\mathbf{v}}^{\rho}$. Thus, $\mathcal{H}_{\mathbf{v}}^{\rho} \subset \mathcal{C}_{\mathbf{a}}$ (because $\mathbf{m}_{\rho}^{\phi} \in \mathcal{C}_{\mathbf{a}}$ and $\mathcal{H}_{\mathbf{v}}^{\rho}$ is parallel to $\mathcal{H}_{\mathbf{a},\mathbf{b}}$). But as $I\to\infty$, the sample median hyperplane $\mathcal{H}_{\mathbf{v}}^{I}$ converges to $\mathcal{H}_{\mathbf{v}}^{\rho}$ in probability (by the Weak Law of Large Numbers). Thus, since $\mathcal{C}_{\mathbf{a}}$ is an open set containing $\mathcal{H}_{\mathbf{v}}^{\rho}$, we have

$$\lim_{I\to\infty} \text{Prob}\left[\mathcal{H}_{\mathbf{v}}^{I} \subset \mathcal{C}_{\mathbf{a}}\right] \quad = \quad 1. \tag{A16}$$

Combining statement (A15) with limit (A16) yields the claim. ◇ `Claim 1`

Let $U_{\mathcal{I}} := \dfrac{1}{I}\sum_{i\in\mathcal{I}} u_i$, as in equation (7).

Claim 2: *If* $\mathbf{m}_{\rho}^{\phi} \in \mathcal{C}_{\mathbf{a}}$, *then* $\lim_{I\to\infty} \text{Prob}\left[U_{\mathcal{I}}(\mathbf{a}) > U_{\mathcal{I}}(\mathbf{b})\right] \;=\; 1.$

Proof. Condition (B3) implies that there is some increasing function $\gamma : [0, \infty) \longrightarrow \mathbb{R}$ such that $\Phi_\rho(\mathbf{x}) = \gamma\left(\left\|\mathbf{x} - \mathbf{m}_\rho^\phi\right\|\right)$ for all $\mathbf{x} \in \mathbb{R}^N$. If $\mathbf{m}_\rho^\phi \in \mathcal{C}_\mathbf{a}$, then $\left\|\mathbf{a} - \mathbf{m}_\rho^\phi\right\| < \left\|\mathbf{b} - \mathbf{m}_\rho^\phi\right\|$; thus, $\Phi_\rho(\mathbf{a}) < \Phi_\rho(\mathbf{b})$. Fix $C \in \mathbb{R}$ with $\Phi_\rho(\mathbf{a}) < C < \Phi_\rho(\mathbf{b})$

Let $\mathbf{x}$ be a ρ-random variable. From equation (8) it is clear that $\Phi_\rho(\mathbf{a})$ is the expected value of $\phi(\|\mathbf{x} - \mathbf{a}\|)$. Meanwhile, $-U_\mathcal{I}(\mathbf{a}) = \frac{1}{I}\sum_{i\in\mathcal{I}} \phi(\|\mathbf{x}_i - \mathbf{a}\|)$ is an empirical estimate of this expected value, based on the sample set $\{\mathbf{x}_i\}_{i\in\mathcal{I}}$. Thus, since $\Phi_\rho(\mathbf{a}) < C$, the Weak Law of Large Numbers says $\lim_{I\to\infty} \text{Prob}[-U_\mathcal{I}(\mathbf{a}) < C] = 1$. By a similar argument, $\lim_{I\to\infty} \text{Prob}[-U_\mathcal{I}(\mathbf{b}) > C] = 1$. Thus, $\lim_{I\to\infty} \text{Prob}\left[U_\mathcal{I}(\mathbf{a}) > -C > U_\mathcal{I}(\mathbf{b})\right] = 1$.
$\diamond$ `Claim 2`

If $\mathbf{m}_\rho^\phi \in \mathcal{C}_\mathbf{a}$, then Claims 1 and 2 together imply that

$$\lim_{I\to\infty} \text{Prob}\Big(\text{The utility profile } \{u_i\}_{i\in\mathcal{I}} \text{ is } \{a,b\}\text{-reasonable}\Big) \quad = \quad 1.$$

We can make a similar argument in the case when $\mathbf{m}_\rho^\phi \in \mathcal{C}_\mathbf{b}$. Finally, it is impossible that $\mathbf{m}_\rho^\phi \in \mathcal{H}_{\mathbf{a},\mathbf{b}}$, because $\left\|\mathbf{a} - \mathbf{m}_\rho^\phi\right\| \neq \left\|\mathbf{b} - \mathbf{m}_\rho^\phi\right\|$ by hypothesis.

This argument holds for any pair $a, b \in \mathcal{A}$. Since $\mathcal{A}$ is finite, we conclude that

$$\lim_{I\to\infty} \text{Prob}\Big(\text{The utility profile } \{u_i\}_{i\in\mathcal{I}} \text{ is reasonable}\Big) \;=\; 1. \qquad \square$$

References

Abreu, D., Matsushima, H., 1992. Virtual implementation in iteratively undominated strategies: complete information. Econometrica 60 (5), 993–1008.

Abreu, D., Sen, A., 1991. Virtual implementation in Nash equilibrium. Econometrica 59 (4), 997–1021.

Apesteguia, J., Ballester, M. A., Ferrer, R., 2011. On the justice of decision rules. Rev. Econ. Stud. 78 (1), 1–16.

Artemov, G., Kunimoto, T., Serrano, R., 2013. Robust virtual implementation: Toward a reinterpretation of the Wilson doctrine. Journal of Economic Theory 148 (2), 424 – 447.

Bag, P. K., Sabourian, H., Winter, E., 2009. Multi-stage voting, sequential elimination and Condorcet consistency. J. Econom. Theory 144 (3), 1278–1299.

Bordley, R. F., 1983. A pragmatic method for evaluating election schemes through simulation. American Political Science Review 77, 123–141.

Brams, S. J., Fishburn, P. C., 1983. Approval voting. Birkhäuser Boston, Mass.

Breit, W., Culbertson Jr., W., June 1970. Distributional equality and aggregate utlity; comment. American Economic Review 60 (3), 435–41.

Breit, W., Culbertson Jr., W. P., June 1972. Distributional equality and aggregate utility: Reply. American Economic Review 62 (3), 501–502.

Caragiannis, I., Procaccia, A. D., 2011. Voting almost maximizes social welfare despite limited communication. Artificial Intelligence 175 (9–10), 1655 – 1671.

Coughlin, P., 1992. Probabilistic Voting Theory. Cambridge Univ. Press, Cambridge.

Dhillon, A., 1998. Extended Pareto rules and relative utilitarianism. Soc. Choice Welf. 15, 521–542.

Dhillon, A., Mertens, J.-F., 1999. Relative utilitarianism. Econometrica 67, 471–498.

Enelow, J. M., Hinich, M. J. (Eds.), 2008. Advances in the Spatial Theory of Voting. Cambridge University Press, Cambridge, UK.

Giles, A., Postl, P., August 2012. Equilibrium and welfare of two-parameter scoring rules. (preprint).

Harter, H. L., Balakrishnan, N., 1996. CRC handbook of tables for the use of order statistics in estimation. CRC Press, Boca Raton, FL.

Hinich, M. J., Munger, M. C., 1997. Analytical Politics. Cambridge University Press, Cambridge, UK.

Horan, S., 2013. Implementation of majority voting rules. (preprint).

Kahneman, D., Diener, E., Schwarz, N. (Eds.), 1999. Well-being: The foundations of hedonic psychology. Russell Sage Foundation, New York, NY.

Kim, S., May 2012. Ordinal versus cardinal voting rules: a mechanism design approach. (preprint).

Laslier, J.-F., Sanver, M. R. (Eds.), 2010. Handbook on approval voting. Studies in Choice and Welfare. Springer, Heidelberg.

Ledyard, J. O., 1984. The pure theory of large two-candidate elections. Public Choice 44 (1), 7–41.

Lerner, A. P., 1944. The Economics of Control. New York.

Lerner, A. P., June 1970. Distributional equality and aggregate utility: Reply. American Economic Review 60 (3), 442–43.

Lindbeck, A., Weibull, J. W., 1987. Balanced-budget redistribution as the outcome of political competition. Public Choice 52 (3), 273–297.

Lindbeck, A., Weibull, J. W., 1993. A model of political equilibrium in a representative democracy. Journal of Public Economics 51 (2), 195–209.

Loewenstein, G., Schkade, D., 1999. Wouldn't it be nice? predicting future feelings. In: Kahneman et al. (1999), Ch. 5, pp. 85–105.

Matsushima, H., 1988. A new approach to the implementation problem. J. Econom. Theory 45 (1), 128–144.

McCain, R., June 1972. Distributional equality and aggregate utility: Further comments. American Economic Review 62 (3), 497–500.

McKelvey, R. D., Patty, J. W., 2006. A theory of voting in large elections. Games Econom. Behav. 57 (1), 155–180.

McManus, M., Walton, G. M., Coffman, R. B., June 1972. Distributional equality and aggregate utility: Further comment. American Economic Review 62 (3), 489–496.

Merrill, S., 1984. A comparison of efficiency of multicandidate electoral systems. American Journal of Political Science 28, 23–48.

Miller, N. R., Nov 1977. Graph-theoretical approaches to the theory of voting. American Journal of Political Science 21 (4), 769–803.

Myerson, R. B., 2002. Comparison of scoring rules in Poisson voting games. J. Econom. Theory 103 (1), 219–251, political science.

Nitzan, S., 2009. Collective Preference and Choice. Cambridge University Press.

Núñez, M., Laslier, J. F., 2013. Preference intensity representation: strategic overstating in large elections. Social Choice and Welfare (to appear).

Pivato, M., February 2013. Voting rules as statistical estimators. Social Choice and Welfare 40 (2), 581–630.

Rae, D., 1969. Decision rules and individual values in constitutional choice. Amer. Polit. Sci. Rev. 63, 40–56.

Schmitz, P. W., Tröger, T., 2012. The (sub-)optimality of the majority rule. Games Econom. Behav. 74 (2), 651–665.

Serrano, R., Vohra, R., 2005. A characterization of virtual Bayesian implementation. Games Econom. Behav. 50 (2), 312–331.