WORKING PAPERS

N° 122

Does universalization ethics justify participation in large elections?

Konrad Dierks, Ingela Alger and Jean-François Laslier

Institute for Advanced Study in Toulouse

Does universalization ethics justify participation in large elections? *

Konrad Dierks† Ingela Alger‡ Jean-François Laslier§

April 29, 2024

Abstract

What drives voters' decisions to participate in large elections under costly voting, despite the rational expectation that this has no impact on the outcome? We propose a new model of ethical voters, by positing that they have Kantian or semi-Kantian preferences. With such preferences, voters evaluate their behavior in light of what the outcome would be, should a fraction of the other voters choose the same course of action. The "other voters" can be either the entire population ("non-partisan ethics") or the individuals with same interest ("partisan ethics"). In a model with two candidates and a continuum of voters, we find that turnout is strictly positive as soon as the evaluation by the voters of the political outcome is not strictly of the "winner-take-all" kind. Moreover, the equilibrium turnout rates depend on the specifics of the election at hand, such as the relative stake of the election for the two supporter groups and the presence of core constituent groups.

Keywords: voter turnout, ethical voter, universalization, *Homo moralis*, Kantian morality

---

*The authors are listed by age, in ascending order. This simple rule would make a fairer norm than alphabetical order, since we all age at the same speed. We urge everyone to adopt this rule, "willing that it become a universal law" for signature in economics papers.

I.A. acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 789111 - ERC EvolvingEconomics), as well as IAST funding from the French National Research Agency (ANR) under grant ANR-17-EURE-0010 (Investissements d'Avenir program). J.-F. L. thanks the EUR grant ANR-17-EURE-0001. We are grateful for comments from David Austen-Smith, Gregory Ponthière, Roberto Sarkisian, Arthur Silve, Alberto Simpser, and Jörgen Weibull, as well as from audiences at Bilkent University, the COMSOC online seminar, Paris School of Economics, the Social Choice and Welfare conference in Seoul, Toulouse School of Economics, and the UC Louvain Hoover Chair. Olivier Lision provided excellent research assistance.

†PhD candidate, Toulouse School of Economics, University of Toulouse Capitole. konrad.dierks@tse-fr.eu

‡Toulouse School of Economics, CNRS, University of Toulouse Capitole, and Institute for Advanced Study in Toulouse, France. ingela.alger@tse-fr.eu

§CNRS, Paris School of Economics. jean-francois.laslier@ens.fr

1 Introduction

Election outcomes depend not only on voters' preferences, but also on whether voters actually vote, since turnout rates may be correlated with party preferences. To wit, in the 20th century in the U.S.A., turnout rates in the three groups of registered voters (Democrats, Republicans, Independents) shifted over time (DeNardo, 1980; Nagel and McNulty, 1996).So turnout matters, and it has been found to vary not only over time but also across countries, and to be correlated with macro-economic factors, the type of election, and even the weather on the day of the election, to name just a few of the variables that have been examined (see, e.g., Blais and Daoust, 2020; Cancela and Geys, 2016; Frank and Martínez i Coma, 2023, and references therein). Empirical studies have further detected correlates between individual turnout decisions and individual characteristics such as age, sex, education, occupation, residential and marital status, etc (see the meta-study by Smets and Van Ham, 2013, and references therein).

Understanding these patterns requires understanding the individual turnout decision-making process. This process has been the subject of a host of theories. Perceived benefits and costs of voting at the individual level are at the center of any rational voter theory of turnout (Downs, 1957). A variety of factors have been invoked to explain turnout, such as a desire to express allegiance to the political system, to participate in the democratic process, to express an opinion, to affirm loyalty to a party, to fulfill one's duty, or to comply with a social norm (Riker and Ordeshook, 1968; Fiorina, 1976; Morton, 1987; Schuessler, 2000; Feddersen, Gailmard, and Sandroni, 2009; D. K. Levine and Mattozzi, 2020). While such factors have been shown to matter empirically (Blais, 2000; Blais and Achen, 2019; Gerber, Green, and Larimer, 2008; Rogers, Green, Ternovski, and Young, 2017), it is not clear how they could explain the patterns evoked above.

The aforementioned patterns suggest that there is a relation between the material impact of the election outcome and the individual decision to vote, at least for part of the electorate. A satisfactory theory should deliver predictions about this relation. While models with voters driven by purely instrumental concerns do deliver such predictions, such concerns are expected to influence participation decisions only if voters can expect to be pivotal (Krishna and Morgan, 2015; Ledyard, 1984; Myerson, 2000; Palfrey and Rosenthal, 1985).[1] But since the probability of being pivotal is essentially nil in elections with large enough electorates, instrumental motives should not matter in such elections, contradicting the patterns evoked

[1]See also the survey by Dhillon and Peralta, 2002 and the literature discussion in Coate and Conlin, 2004, as well as the books by Aytaç and Stokes, 2019 and Blais and Daoust, 2020.

above. We propose a novel theory of turnout in large elections.

The driving force is a form of ethical preferences—dubbed *Homo moralis*—which can be interpreted as capturing a form of universalization: when contemplating a course of action, a *Homo moralis* evaluates what his material payoff would be if, hypothetically, a share κ of the population to which he belongs would follow the same course of action, where κ is the individual's *degree of universalization*.[2] This idea is reminiscent of Kant's categorical imperative (Kant, 1785), although philosophers warn that a Kantian "maxim" is not a course of action (Braham and van Hees, 2020). In any case, the logic of universalization spreads over moral theories (Gravel, Laslier, and Trannoy, 2000) and also guides actual moral judgments (S. Levine, Kleiman-Weber, Schultz, and Cushman, 2020). One obtains the standard materialistic *Homo oeoconomicus* for $\kappa = 0$, the Kantian model of Laffont, 1975 or Roemer, 2019 for $\kappa = 1$, while values of κ between 0 and 1 trigger partial universalization (Alger and Weibull, 2013).

Given these ethical preferences, in our model each voter will be seen to act on information about the material consequences of the election, despite a positive cost to vote. The action is rational in the sense that it maximises a well-defined utility function. It is compatible with the awareness that no single vote has any impact on the election outcome. Furthermore, at equilibrium the beliefs about the other voters' behaviors are taken to be correct.

Specifically, we evaluate the consequences of these preferences in a standard political model. As in most models of turnout, there are two candidates (or parties, or referendum proposals), A and B. We take B to be the (known) underdog. Some voters always turn out to vote, perhaps because of a deep sense of duty, a long-held habit, a strong wish to signal support of democracy, etc. They are the core voters (DeNardo, 1980), which we will refer to as the candidate's *base*. Our model is about the other voters, who do *not* systematically turn out to vote. Their cost of voting is uncertain at the individual level, although the cost distributions are known, and each such *cost-sensitive voter* decides on a threshold strategy: she votes if and only if the realized cost falls short of this threshold (like in Coate and Conlin, 2004; Feddersen and Sandroni, 2006).

The distribution of expressed votes across the two candidates determines the political outcome. At this level, we keep the familiar zero-sum pattern of electoral competition: the

[2]These preferences emerged from analysis of the evolutionary foundations of preferences (Alger and Weibull, 2013). Our approach is thus close in spirit to Conley, Toossi, and Wooders, 2006, who base their voters' motivation to participate in elections on evolutionary arguments. See also the book by Hatemi and McDermott, 2011, which *inter alia* cites evidence of intriguing correlations between biological factors such as genes on the one hand, and political preferences and even turnout on the other hand. For experimental evidence on behavior consistent with *Homo moralis* preferences, see Van Leeuwen and Alger, forthcoming.

outcome is positive for one side and negative for the other side. However, it need not be a winner-take-all election. Our gain-loss function also encompasses institutional settings where the vote share itself matters and political power is shared so that "the loser gets some". This is meant to capture a range of possible institutional settings ranging from the pure majoritarian case (that will be obtained as a limit case of our model) to a kind of "random dictatorship", or "proportional two-party system"[3]

At the level of the voters we break the symmetry and introduce a parameter $\rho \geq 1$ that we call *the stake of the election*: this parameter represents the importance of the political benefit obtained through the election as perceived by the underdog's supporters relative to the other group's supporters. As an illustrating example, if the underdog tends to represent low-income households, the stake is expected to be higher, the stronger are the redistributive consequences of the election. We will say that the stake is neutral if $\rho = 1$. To summarise, the material consequences of the election for each individual depend on the political outcome, the stake, and the individual's cost (that may be incurred or not).

By definition, universalization ethics implies a reference to a group, the population to which the individual belongs. We examine two settings: the *partisan* setting and the *non-partisan* one. In the non-partisan setting, the reference group can be interpreted as the set of all (independent) voters, while in the partisan setting there are two distinct populations, one for each candidate. In the partisan setting, the voter applies the universalization argument to the set of co-partisans, by evaluating what the outcome of the election would be if—hypothetically—a share κ of the co-partisans were to choose the same threshold as the voter himself. By contrast, in the non-partisan setting, the voter applies the universalization argument to the set of independent voters, and chooses two rather than one cost threshold, by taking into account the expected benefits and costs over the two possible preference realizations, "behind a veil of ignorance".[4]

Our objective is to characterize rational behavior in these (one or two) population games, assuming that all voters have the same degree of universalization κ. The rich setting enables us to address a host of questions: is turnout positive in equilibrium, and how do turnout

[3]Modifying this way the political benefit function in a two-party model in order to contrast proportional representation with winner-take-all is used, for instance by Lizzeri and Persico, 2001. Studying turnout, Herrera, Morelli, and Nunnari, 2016 similarly modify the outcome function.

[4]This setting does not appear in the existing literature. Existing models are therefore not suited to explain turnout rates of voters who may change their ranking over parties between elections, and who are humble enough to realize that the information on which they base their party preference may be wrong, or who consider democratic participation as a norm for all citizens, not only for those who happen to be on their side.

rates depend on the primitives (the way in which the relative margins affect the material benefits accruing to the parties, the candidates' bases, the stake of the election)? Are there equilibria in which the underdog wins the election? Finally, do equilibria exist, and if so, can there be multiple equilibria?

Prior to summarizing our findings, we compare our formalization of ethically driven voters to existing ones. An early formalization (Harsanyi, 1980) of ethical voters posits that voters are *rule utilitarians*: in the words of Harsanyi, 1977 such a voter does not look at the various issues from a partisan point of view but from the standpoint of an impartial but humane and sympathetic observer.[5] Furthermore, Harsanyi defines the moral behavior of rule-utilitarian individuals as "involving a firm commitment [...] to a specific moral strategy" (p.115 in Harsanyi, 1980), where the moral strategy maximizes the sum of individual utilities. Our non-partisan setting is in line with Harsanyi's view of voters as "impartial observers", but our formalization of ethically driven voters does not amount to altruistic utilitariansim, but instead to a self-centered universalization thought experiment: a voter considers each course of action in the light of what *her* material well-being would be if some share of the others voters were to choose the same course of action.[6]

Our formalization of ethical voters should not be confused either with group-based voter participation models in which strategic decisions are made at the collective level and an ethical voter applies a decision rule —that is a cost threshold, like in our model— which maximizes the group's aggregate material well-being, given the other group's cost threshold.[7] By adopting this dutiful behavior, such a voter receives a constant payoff $D > 0$. For each group, the equilibrium cost threshold optimally trades off the probability of winning against the group-aggregate expected cost of voting, given the other group's threshold. A counter-intuitive feature of these models is that, at the individual level, each ethical voter would be perfectly happy to incur any positive voting cost, since D is assumed to exceed

[5] The full quote is: "In any social situation, each participant will tend to look at the various issues from his own, self-centered, partisan point of view. In contrast, if anybody wants to assert the situation from a *moral* point of view in terms of some standard of justice and equity, this will essentially amount to looking at it from the the standpoint of an impartial but humane and sympathetic observer." (p.623 in Harsanyi, 1977)

[6] Alger and Weibull, 2017 and Laslier, 2023 study the relation between (partial) universalisation ethics and (partial) Beckerian altruism.

[7] In Coate and Conlin, 2004 or Herrera et al., 2016 voters are like in our partisan setting. By contrast, the model adopted by Feddersen and Sandroni, 2006 can be viewed as a mix of our non-partisan and partisan settings: their voters face no uncertainty regarding their party preference, but they care about the expected cost of voting for the supporters of both parties. See also the group-based models by Morton, 1991 and Bierbrauer, Tsyvinski, and Werquin, 2022, which have both endogenous turnout and endogenous party platforms.

the largest possible cost realization. The duty that such a voter feels obliged to fulfill thus consists of reducing the aggregate cost of voting, by abstaining from voting when the realized cost is above the equilibrium threshold: some voters "receive a [duty] payoff for not voting" (Feddersen and Sandroni, 2006, p. 1272). By contrast, in our model a voter's utility depends directly on the material benefit she would enjoy if some share of the others also applied the same cost threshold, and she considers only own cost when evaluating cost thresholds. In other words, in our model the decision to incur a positive voting cost is individually rational. In particular, a voter would be willing to incur a positive cost to vote even if no other voter voted, whereas in the above mentioned models an ethical voter votes only because she knows that a share of the other voters will apply the same decision rule .

We now turn to the description of our results. They heavily depend on the shape of the political outcome function that captures the institution, being more or less power-sharing, from random dictatorship to pure majoritarian. Since some effects vanish when the institution become close to winner-take-all, we first describe the results for the general power-sharing cases.

With power-sharing, we can establish existence generally in the non-partisan setting, while in the partisan setting equilibria sometimes fail to exist. We find that for any positive degree of universalization κ, in any equilibrium aggregate turnout is strictly positive (except in two knife-edge cases). This is because the universalization thought experiment makes each voter act *as if* their decision had a real weight on the outcome, and because the smallest possible cost realizations are close to zero. The result that positive voting costs are incurred in any equilibrium is qualitatively similar to results found in existing models on ethical voting, However, the driver is very different, as already mentioned. We further derive results which shed light on aspects that have hitherto been neglected in the literature.

Firstly, the inclusion in the model of core constituents, or bases, is novel. We show that the relative size of the candidates' bases are crucial. In particular, if the underdog's base exceeds that of the leader, there may exist equilibria in which the underdog wins. This result is explained by the cost advantage that a large base confers on the cost-sensitive voters: the base enables them to reach higher turnout levels at a lower cost. This contrasts sharply with the results in other models with known underdogs (Feddersen and Sandroni, 2006; Herrera et al., 2016), where the underdog gets the smallest (expected) vote share in the unique equilibrium (in Feddersen and Sandroni, 2006, the underdog may win due to the assumed uncertainty about the share of ethical voters).

Secondly, analysis of the non-partisan setting, in which voters select their participation strategy behind the veil of ignorance as to which candidate they will support, is also novel to the literature. In this setting the voters take into account the effects of both cost thresholds on the expected utility, thereby internalizing the externalities they generate across the two groups of cost-sensitive voters. We show that as a result the only candidate that obtains a turnout among its cost-sensitive supporters is that with the highest expected net benefit from voting. By contrast, in the partisan setting there is no internalization of externalities, and the cost-sensitive voters of both groups incur a positive expected voting cost in any equilibrium.

Thirdly, by contrast to most of the literature, we do not impose assumptions that guarantee equilibrium existence and uniqueness. To get a sense of how common non-existence and multiplicity is, we provide illustrating examples, as well as an online tool that enables the reader to explore other parameter sets. We also derive sufficient conditions for existence and uniqueness.

Without power sharing, that is in the majoritarian, winner-take-all case, the incentive to vote provided by the κ-unviversalization reasoning vanishes so that, except for some knife-edge values of the parameters, costly participation is nil for both sides in the partisan case. In the non-partisan case, either costly participation is nil for both sides leaving the front-runner to win, or is positive for one and only one side, leading to a tied outcome.

In the next section we describe the political model, and in the following two sections we analyze the partisan and the non-partisan settings, assuming that voters have *Homo moralis* preferences. A final section provides a summary of the results.

2 The political model

2.1 Institutional setting and political outcome

An election is taking place with two candidates, A and B (or more generally, two alternatives, such as political parties, two proposals in a referendum, etc.). The electorate is formalized as a continuum, divided in two groups: a group of size $\bar{a}$ supporting A, and a group of size $\bar{b} < \bar{a}$ supporting B. Since the B-supporters are less numerous in the population, we will refer to their group as the *underdog supporters*, and the group of A-supporters as the *leader supporters*.

Each voter either votes for their preferred candidate or abstains, and candidates A and B receive $a \leq \bar{a}$ and $b \leq \bar{b}$ votes, respectively, which generates the relative margins α and

$\beta = -\alpha$:

$$\alpha = \frac{a-b}{a+b}, \quad \beta = \frac{b-a}{a+b}. \tag{1}$$

The outcome of the election generates some material (instrumental) benefit to the voters. The material benefit is given by a strictly increasing and twice differentiable function $h : \mathbb{R} \to \mathbb{R}$ of the relative margin of one's candidate. We assume that h is symmetric around 0, i.e., $h(-x) = -h(x)$ and $h'(x) = h'(-x)$, and that $h''(0) < 0$ for all $x > 0$. This assumption captures the idea that the competition between the two candidates is zero-sum, and that the marginal impact on the material benefit is the largest at $x = 0$, the threshold value of x above which the candidate wins the election by securing a greater total turnout than the other candidate. In particular, our setting includes functions for which the slope of h at 0 is arbitrarily large, while it is close to 0 elsewhere; this limit case of our model thus approximates the classical winner-take-all setting. However, by including h-functions such that the slope is sizeable everywhere, our model also encompasses situations where voters care about the margin of victory: this assumption is natural for parliamentary elections, where margins determine the number of seats obtained. The other limit case, opposed to the winner-take-all setting, consists in taking h to be linear. A possible interpretation is that the decision will be taken by one side or the other, yielding outcomes +1 and -1, with a probability that is precisely equal to the proportion of votes obtained by A and B. This has a flavor of proportional representation and could be called "random dictatorship among participants".[8]

We further assume that there is a parameter ρ, which we call the *stake of the election*, such that the material benefit to A-supporters is

$$h(\alpha), \tag{2}$$

while that to B-supporters is

$$\rho \cdot h(\beta). \tag{3}$$

If $\rho > 1$ the election is more important for the underdog supporters than for the leader supporters. We will say that the stake is neutral if $\rho = 1$.

[8]Herrera et al., 2016 posit the material benefit function $a^\gamma/(a^\gamma + b^\gamma)$ for the A-supporters and $b^\gamma/(a^\gamma + b^\gamma)$ for the B-supporters, where the parameter $\gamma \in [1, +\infty)$ captures the power sharing rule. In Appendix A we identify a function h that is a linear transformation of the material benefit term in Herrera et al., 2016. Rescaling the cost accordingly, this shows that our model benefit term is more general.

2.2 Voting costs and strategies

Some voters always turn out to vote—they may be driven by a strong sense of civic duty, a strong social pressure, a habit, or any other motivation outside of this model. There is a mass $0 < a_0 < \bar{a}$ of such voters who vote for A, and a mass $0 < b_0 < \bar{b}$ of such voters who vote for B. We will refer to a_0 as *A's base* and to b_0 as *B's base*. The model examines the behavior of the remaining voters, a mass $a_v = \bar{a} - a_0$ of which are A-supporters, and a mass $b_v = \bar{b} - b_0$ of which are B-supporters. These voters are cost-sensitive: each of them faces a positive random cost of voting, and their turnout decision depends on their realized cost. Formally, the function $f_A : \mathbb{R}_+ \to \mathbb{R}_{\geq 0}$ maps each cost to the probability density for a cost-sensitive A-supporter to have that voting cost. We assume that

1. the support of f_A, i.e. $\{x \in \mathbb{R}_{>0} : f_A(x) > 0\}$, is either an interval $(0, \bar{c}]$, for some $\bar{c} \in \mathbb{R}_+$, or $\mathbb{R}_+$;

2. f_A is continuous on its support.

We use $F_A(c)$ to denote the proportion of cost-sensitive A-supporters whose cost realization falls short of c:

$$F_A(c) = \int_0^c f_A(t)\, dt, \tag{4}$$

so that $1 = F(+\infty)$. The same assumptions apply to group B, with notation f_B, and F_B.

We will study two different models of how the voting strategies are chosen: the partisan and the non-partisan one, but in both cases, we restrict attention to threshold strategies. A cost-sensitive A-supporter i chooses a threshold $s_A^i \in \mathbb{R}_{\geq 0} \cup \{\infty\}$, and votes (for A) if the cost realization c_A^i does not exceed s_A^i and abstains otherwise. Likewise, a cost-sensitive B-supporter j chooses a threshold $s_B^j \in \mathbb{R}_{\geq 0} \cup \{\infty\}$. Voters have correct beliefs about the voting cost distributions.

We restrict attention to type-homogenous strategy profiles, in which all voters with the same preference over the candidates choose the same strategy. At a type-homogenous strategy profile $s = (s_A, s_B)$, the realized turnouts are

$$a(s_A) = a_0 + a_v F_A(s_A) \tag{5}$$

and

$$b(s_B) = b_0 + b_v F_B(s_B), \tag{6}$$

respectively, for the two candidates, and the following relative vote margins obtain:

$$\alpha(s) = \frac{a(s_A) - b(s_B)}{a(s_A) + b(s_B)}, \quad \beta(s) = -\alpha(s). \tag{7}$$

2.3 Example

Throughout we will use the following specification for h to provide numerical examples to illustrate our general results and discussions:

$$h(x) = \frac{\arctan(mx)}{\arctan(m)}. \tag{8}$$

The parameter $m \in \mathbb{R}_+$ changes the slope of h: the larger is m, the larger is the marginal benefit for small margins and the smaller is the marginal benefit for large margins. The linear case obtains (by continuity) for $m = 0$ and the step function for $m \to \infty$, giving m a similar role as the power-sharing parameter γ in the Herrera et al., 2016 benefit term. This is illustrated in Figure 1.

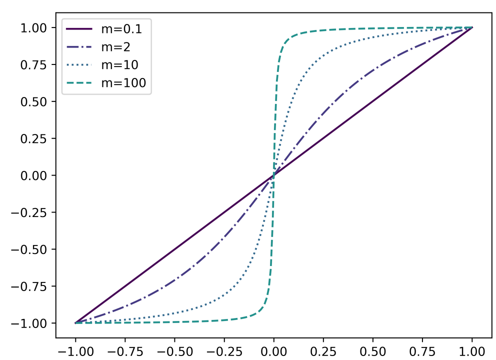


Figure 1: h as defined in equation (8) for different values of m

The examples will always use uniform cost distributions with support $[0, \theta_A]$ and $[0, \theta_A]$ for A- and B-supporters. The model is fully specified with the parameters m, ρ, θ_A, θ_B, $\bar{a}$, $\bar{b}$, a_0, b_0.

3 Partisan ethics (*ex post* setting)

Our formalization of an ethical voter amounts to assuming that their utility function belongs to the *Homo moralis* preference class (Alger and Weibull, 2013). Following Alger and Laslier, 2022, who also study a model with a continuum of voters with such preferences, we posit that *Homo moralis* preferences induce each voter to evaluate any strategy in the light of the material benefit that would realize if—hypothetically—a fraction $\kappa \in [0,1]$ of the other voters were also to play this strategy instead of the strategies they are actually using. The parameter κ is the *degree of universalization*, here assumed to be common to all cost-sensitive voters.

3.1 Partisan ethics: payoff computations

Under partisan ethics, the reference group is taken to be the other cost-sensitive voters who have same preferences over the two candidates; the participation strategy is decided *ex post*, once the voter's affiliation is known. Thus, each voter i in group A (resp. each voter j in group B) evaluates any strategy s_A^i (resp. s_B^j) in the light of the material benefit that would realize if—hypothetically—a fraction $\kappa \in [0,1]$ of the other cost-sensitive A-supporters (resp. B-supporters) were also to play s_A^i (resp. s_B^j) instead of the strategies they are actually using.

At a type-homogenous strategy profile $s = (s_A, s_B)$, each A-supporter obtains expected net material benefit

$$EU_A(s) = h(\alpha(s)) - \int_{c=0}^{s_A} cf_A(c)\,\mathrm{d}c \tag{9}$$

and each B-supporter obtains expected net material benefit

$$EU_B(s) = \rho h(\beta(s)) - \int_{c=0}^{s_B} cf_B(c)\;\mathrm{d}c. \tag{10}$$

Homo moralis preferences induce each A-supporter i to consider the hypothetical number of votes

$$a^\kappa(s_A, s_A^i) \quad = \quad a_0 + (1-\kappa)a_v F_A(s_A) + \kappa a_v F_A(s_A^i) \tag{11}$$

in favor of A, with the corresponding relative vote margin

$$\alpha^\kappa(s, s_A^i) = \frac{a^\kappa(s_A, s_A^i) - b(s_B)}{a^\kappa(s_A, s_A^i) + b(s_B)}, \tag{12}$$

and this defines the voter's expected utility

$$EU_A^\kappa(s, s_A^i) = h(\alpha^\kappa(s, s_A^i)) - \int_{c=0}^{s_A^i} cf_A(c)\,\mathrm{d}c. \tag{13}$$

Likewise, each B-supporter j considers the hypothetical number of votes

$$b^\kappa(s_B, s_B^j) \quad = \quad b_0 + (1-\kappa)b_v F_B(s_B) + \kappa b_v F_B(s_B^j), \tag{14}$$

in favor of B, with the corresponding relative vote margin

$$\beta^\kappa(s, s_B^j) = \frac{b^\kappa(s_B, s_B^j) - a(s_A)}{a(s_A) + b^\kappa(s_B, s_B^j)}, \tag{15}$$

and obtains expected utility

$$EU_B^\kappa(s, s_B^j) = \rho h(\beta^\kappa(s, s_B^j)) - \int_{c=0}^{s_B^j} c f_B(c)\, \mathrm{d}c. \tag{16}$$

These equations reveal the main driver of the behavior of voters under partisan ethics in our model. Consider Equation 13. When $\kappa > 0$ an increase in s_A^i affects the vote share in favor of A and thus, given our assumption on h, the political benefit of an A-supporter. But, at $s_A^i = 0$, the marginal effect on the expected payed cost of an increase in s_A^i is $0 \times f_A(0) = 0$. It follows that setting a 0 threshold is never a best response for an A-supporter.[9]

3.2 A change of variables

Before going further, we proceed to a change of variables that simplifies the analysis. Seeing from (5) and (6) that the threshold s_A that yields turnout a is $F_A^{-1}\left(\frac{a-a_0}{a_v}\right) \in [0, \infty]$ and the threshold s_B that yields turnout b is $F_B^{-1}\left(\frac{b-b_0}{b_v}\right) \in [0, \infty]$, we write the expected utilities in (13) and (16) as

$$EU_A\left(a, b, a^i\right) = h\left(\alpha^\kappa\left(a, b, a^i\right)\right) - C_A\left(a^i\right) \tag{17}$$

and

$$EU_B\left(a, b, b^j\right) = \rho h\left(\beta^\kappa\left(a, b, b^j\right)\right) - C_B\left(b^j\right), \tag{18}$$

where

$$\alpha^\kappa\left(a, b, a^i\right) = \frac{(1-\kappa)\, a + \kappa a^i - b}{(1-\kappa)\, a + \kappa a^i + b} \tag{19}$$

$$\beta^\kappa\left(a, b, b^j\right) = \frac{(1-\kappa)\, b + \kappa b^j - a}{(1-\kappa)\, b + \kappa b^i + a}, \tag{20}$$

[9] This reasoning will, however, be seen to break down in the limit case that we use to examine winner-take-all elections, because the political outcome function h is then locally flat.

and

$$
\begin{aligned}
C_A\left(a^i\right) &= \int_0^{F_A^{-1}\left(\frac{a^i - a_0}{a_v}\right)} c f_A(c)\, dc \\
C_B\left(b^j\right) &= \int_0^{F_B^{-1}\left(\frac{b^j - b_0}{b_v}\right)} c f_B(c)\, dc.
\end{aligned}
\tag{21}
$$

Henceforth, the strategy of A-supporter i is thus a "turnout" $a^i \in [a_0, \bar{a}]$, and that of B-supporter j a "turnout" $b^j \in [b_0, \bar{b}]$, although it should be clear to the reader that what these voters are really choosing are the cost thresholds that would yield these turnout levels. This change of variables facilitates analysis because the functions C_A and C_B are strictly convex, for any cost distributions F_A and F_B satisfying our assumptions.

Lemma 1. *Both C_A and C_B are strictly convex and strictly increasing.*

Proof. By a substitution $z = F_A(c)$

$$
C_A(a) = \int_0^{\frac{a-a_0}{a_v}} F_A^{-1}(z)\, dz. \tag{22}
$$

Then,

$$
C_A'(a) = \frac{1}{a_v} F_A^{-1}\left(\frac{a - a_0}{a_v}\right) > 0 \tag{23}
$$

for $a > a_0$, with $C_A'(a_0) = 0$, and

$$
C_A''(a) = \frac{1}{a_v^2} \frac{1}{f_A\left(F_A^{-1}\left(\frac{a-a_0}{a_v}\right)\right)} > 0 \tag{24}
$$

for $a \geq a_0$. The same argument applies to C_B. □

Unless stated otherwise, in the numerical examples below we use the functions

$$
\begin{aligned}
C_A(a) &= \frac{\theta_A}{2}\left(\frac{a-a_0}{a_v}\right)^2 \\
C_B(b) &= \frac{\theta_B}{2}\left(\frac{b-b_0}{b_v}\right)^2,
\end{aligned}
\tag{25}
$$

which correspond to a uniformly distributed cost on $[0, \theta_A]$ for the A-supporters and a uniformly distributed cost on $[0, \theta_B]$ for the B-supporters.

3.3 Partisan ethics: comparison with group-based models

We are now in a position to provide a detailed comparison between our formalization of ethical voters and that adopted in group-based models (first formalized by Coate and Conlin, 2004, and Feddersen and Sandroni, 2006). In these models an ethical voter gets a "duty payoff" D from "doing their part" (Feddersen and Sandroni, 2006), where D exceeds the highest possible cost realization. A decision simply based on a "duty to vote" would thus lead all the ethical voters to vote. Dependence of an ethical voter's turnout decision on their cost realization obtains by positing that an ethical voter adopts a cost threshold in order to reduce the aggregate cost; this cost reduction is traded off against the loss with the associated reduced probability of winning. In other words, "doing their part" entails abstaining when the cost realization exceeds the threshold. The predicted turnout rates are obtained as "an equilibrium between two party planners", each of which "looks at the total electoral benefit" for their preferred candidate "net of the total cost incurred by his supporters" (Herrera et al., 2016, p. 612).

By contrast, in our model each voter simply maximizes his own expected utility, and there is no constant duty payoff. Such utility maximization imposes fewer demands on the information that the voter needs in order to select an ethical behavior, compared to the group-based models, in which an ethical voter needs to place herself or himself in the shoes of a social planner to understand which cost threshold she should adopt to obtain the constant duty payoff D. With *Homo moralis* preferences, a voter instead evaluates each possible strategy applying a simple universalization calculus to the benefit, while the expected cost of the deviation is the individual's own true expected cost. Indeed, *Homo moralis* preferences make a voter evaluate a strategy in the light of the expected material utility that would obtain if a share κ of the others also adopted the same strategy; whether or not others adopted a different threshold than they do would be irrelevant for this voter's expected cost of voting.

In spite of this important conceptual difference between our model and group-based models, in some settings the two models are mathematically equivalent. We will begin our analysis with such a setting.

Consider the special case of our model where all ethical voters have degree of universalization $\kappa = 1$. The expected utilities in (17) and (18) then boil down to

$$EU_A\left(a, b, a^i\right) = h\left(\frac{a^i - b}{a^i + b}\right) - C_A\left(a^i\right) \tag{26}$$

and

$$EU_B\left(a,b,b^j\right)=\rho h\left(\frac{b^j-a}{a+b^j}\right)-C_B\left(b^j\right). \tag{27}$$

The strategy profile (a^*,b^*) is a type-homogenous Nash equilibrium if and only if

$$\begin{cases} a^*\in\arg\max_{a^i\in[a_0,\bar{a}]}\ h\left(\frac{a^i-b^*}{a^i+b^*}\right)-C_A\left(a^i\right) \\ b^*\in\arg\max_{b^j\in[b_0,\bar{b}]}\ \rho h\left(\frac{b^j-a^*}{a^*+b^j}\right)-C_B\left(b^j\right). \end{cases} \tag{28}$$

The mathematical equivalence of this setting with group-based models is easy to see, since any (a^*,b^*) satisfying (28) could alternatively be interpreted as representing a Nash equilibrium of a game played by "two party planners", each of which "looks at the total electoral benefit" for their preferred candidate "net of the total cost incurred by his supporters" (Herrera et al., 2016, p. 612); likewise, see Definition 1 in Feddersen and Sandroni, 2006 for the conditions ensuring that the cost threshold of each party maximizes its supporters aggregate expected material payoff, given the other party's cost threshold, and also the description of equilibrium on p.1481 in Coate and Conlin, 2004.[10] We can thus state our first observation:

Remark 1. *If all cost-sensitive voters have Homo moralis preferences with degree of universalization $\kappa=1$, any pair of threholds implemented at a type-homogenous Nash equilibrium of the two-population game is implemented at a Nash equilibrium of the two-player game between two party planners, each of whom seeks to maximize the aggregate material payoffs of their respective constituent groups. The type-homogenous Nash equilibrium implements the cost thresholds in a decentralized manner (in the sense that each voter simply maximizes her own expected utility).*

Having established this mathematical equivalence with group-based models of ethical voters in the special case $\kappa=1$, we note that equilibrium existence is not guaranteed. Indeed, while a sufficient condition for an equilibrium to exist is that both objective functions in (28) are quasi-concave in a^i respectively b^j, the strict convexity of h for negative relative margins, together with the strict convexity of the functions C_A and C_B, implies that quasi-concavity is not guaranteed. Facing the same issue with their benefit function (see Footnote 12), Herrera et al., 2016 identify and impose conditions on the cost distributions that imply equilibrium

[10]There are also other differences between our model and those by Feddersen and Sandroni, 2006 and Coate and Conlin, 2004. For example, in the former each party considers the aggregate societal cost, and not only the aggregate cost of its supporters, while in the latter there is *ex ante* uncertainty about the distribution of voters into A- and B-supporters. However, our focus here is on underlining the mathematical similarity between type-homogenous Nash equilibria in our model with $\kappa=1$ and the characterization of equilibrium in the group-based models.

existence, and even uniqueness. While we will return to the issues of existence and uniqueness in the last subsection, we will first establish properties of equilibria conditional on their existence in the general model.

3.4 Partisan ethics: a never-a-best-response result

In the general model voters may have any degree of universalization $\kappa \in [0,1]$, and have the utilities specified in (17) and (18). Then, the strategy profile (a^*, b^*) is a type-homogenous Nash equilibrium if and only if

$$\begin{cases} a^* \in \arg\max_{a^i \in [a_0, \bar{a}]} \; h\left(\frac{(1-\kappa)a^*+\kappa a^i - b^*}{(1-\kappa)a^*+\kappa a^i + b^*}\right) - C_A\left(a^i\right) \\ b^* \in \arg\max_{b^j \in [b_0, \bar{b}]} \; \rho h\left(\frac{(1-\kappa)b^*+\kappa b^j - a^*}{a^*+(1-\kappa)b^*+\kappa b^i}\right) - C_B\left(b^j\right). \end{cases} \tag{29}$$

By contrast to the special case $\kappa = 1$, where each individual best-responds to the other group's turnout only, when $\kappa < 1$ each individual best-responds to both groups' turnouts.

Proposition 1. *Under partisan ethics,*

- *if $\kappa = 0$, there exists a unique equilibrium, $(a^*, b^*) = (a_0, b_0)$;*
- *if $\kappa \in (0,1]$, $a = a_0$ (resp. $b = b_0$) is never a best response for a voter in group A (resp. B), and thus any equilibrium (a^*, b^*) is such that $\bar{a} \geq a^* > a_0$ and $\bar{b} \geq b^* > b_0$.*

Proof. For any level of turnouts (a, b) from the other voters, A-supporter i's marginal utility of a^i is

$$\frac{\partial}{\partial a^i} U_A^\kappa\left(a, b, a^i\right) = h'(\alpha^\kappa(a, b, a^i)) \frac{2\kappa b}{\left[(1-\kappa)\, a + \kappa a^i + b\right]^2} - C_A'\left(a^i\right), \tag{30}$$

where $h' > 0$. Recalling, from the proof of Lemma 1, that

$$C_A(x) = \int_0^{F_A^{-1}\left(\frac{x-a_0}{a_v}\right)} c f_A(c)\, dc \tag{31}$$

$$C_A'(x) = \frac{1}{a_v} F_A^{-1}\left(\frac{x-a_0}{a_v}\right), \tag{32}$$

we see that $C_A'(a_0) = 0$. It follows that if $\kappa = 0$ the unique best response to any (a, b) is $a^i = a_0$, while if $\kappa > 0$ $a^i = a_0$ is never a best response. Identical arguments apply to any B-supporter. □

With purely instrumentally driven voters ($\kappa = 0$), the cost-sensitive voters are not willing to incur any cost to vote, since with a continuum of voters each individual vote has a nil effect on the election outcome. The second part of the proposition shows that a willingness to incur some cost of voting arises as soon as the degree of universalization is strictly positive. This is because (i) for any $\kappa \in (0, 1]$, the individual voter considers the outcome that would realize if some positive share of the voters voted: in some sense, it is *as if* the individual voter had an impact on the outcome, and (ii) the smallest cost realizations approach zero (as per our assumption on the cost distributions).

3.5 Partisan ethics: further properties of equilibria

By Proposition 1, for any $\kappa > 0$ any equilibrium (a^*, b^*) satisfies the following first-order conditions:

$$\frac{\partial}{\partial a^i} U_A^\kappa \left(a, b, a^i\right)|_{a^i=a^*, b=b^*} = \frac{2\kappa b^* h'(\alpha(a^*, b^*))}{(a^* + b^*)^2} - C'_A\left(a^*\right) \begin{cases} = 0 \text{ if } a^* \in (a_0, \bar{a}) \\ \geq 0 \text{ if } a^* = \bar{a} \end{cases} \tag{33}$$

$$\frac{\partial}{\partial b^i} U_B^\kappa \left(a, b, b^j\right)|_{a=a^*, b^j=b^*} = \frac{2\kappa a^* \rho h'(\beta(a^*, b^*))}{(a^* + b^*)^2} - C'_B\left(b^*\right) \begin{cases} = 0 \text{ if } b^* \in (b_0, \bar{b}) \\ \geq 0 \text{ if } b^* = \bar{b}. \end{cases} \tag{34}$$

Notice that, for any interior equilibrium $(a^*, b^*) \in (a_0, \bar{a}) \times (b_0, \bar{b})$, the two equations and the fact that $\alpha(a, b) = -\beta(a, b)$ and $h'(x) = h'(-x)$ together imply:

$$\frac{b^*}{\rho a^*} = \frac{C'_A(a^*)}{C'_B(b^*)}. \tag{35}$$

A number of results can be obtained from conditions (33)-(35). In order to prepare the ground for this, we first derive some properties of the functions C_A and C_B, which depend on the mass of cost-sensitive voters (a_v and b_v) relative to that of the bases (a_0 and b_0), and the cost distributions F_A and F_B (see (21)).

Lemma 2. *The functions C_A and C_B have the following properties:*

1. *if $a_0 \geq b_0$, $a_v \geq b_v$, and $F_A(c) \geq F_B(c)$ for all $c \in \mathbb{R}_+$, with at least one of the inequalities holding strictly, then $C'_A(x) < C'_B(x)$ for all x;*

2. *if $a_0 < b_0$, then for any cost distributions F_A and F_B, $\exists\, \bar{x} \in (b_0, \bar{b}]$ such that $C'_A(x) > C'_B(x)$ for all $x < \bar{x}$.*

The lemma contains two statements about the derivatives of C_A and C_B. These derivatives have a clear interpretation. For any given turnout x by the cost-sensitive voters in group A, $C_A'(x)$ is the (expected) marginal cost that each of these voters would need to incur to increase this turnout marginally. This means that the first part of the lemma establishes sufficient conditions for the leader supporters to enjoy an absolute cost advantage over the underdog supporters: this occurs when their base and its mass of cost-sensitive voters is at least as large and its cost distribution F_A is more favorable, compared to the corresponding features of the underdog supporters. Note that the underdog supporters cannot enjoy such an absolute cost advantage, since the assumption $\bar{a} > \bar{b}$ implies that if $b_0 \geq a_0$, then $b_v < a_v$. If the underdog's base exceeds that of the leader ($b_0 > a_0$) its supporters can still have a cost advantage over the leader supporters for turnouts close enough to the underdog's base b_0, as shown in the second part of the lemma. This results from the marginal cost for the underdog supporters then being close to zero, while that of the leader supporters is strictly positive for these turnout levels. This property holds whether the cost distributions F_A and F_B favor the underdog or the leader supporters.

With these observations in hand, we first examine settings where the leader supporters enjoy a cost advantage over the underdog supporters, like in the first part of Lemma 2: their base is at least as large, their mass of cost-sensitive voters is at least as large, and the cost distributions favor the leader supporters.

Proposition 2 (Partisan ethics)**.** *Suppose that the leader supporters enjoy a cost advantage over the underdog supporters: their base is at least as large ($a_0 \geq b_0$), their mass of cost-sensitive voters is at least as large ($a_v \geq b_v$), and the cost distributions favor them ($F_A(c) \geq F_B(c)$ for all $c \in \mathbb{R}_+$), with at least one of the inequalities holding strictly. Then:*

1. *if the stake is neutral or almost neutral (i.e., $\rho \geq 1$ is close enough to 1), the leader wins ($\alpha(a^*, b^*) > 0$) at any equilibrium (a^*, b^*);*

2. *for a large enough stake ρ, there may exist equilibria (a^*, b^*) in which the underdog wins ($\alpha(a^*, b^*) < 0$).*

Proof. Suppose, by contradiction, that $b^* \geq a^*$, in which case either both a^* and b^* are interior, or $b^* = \bar{b}$ and a^* is interior. Lemma 2 then implies that $C_A'(a^*) < C_B'(b^*)$. Hence, $b^* C_B'(b^*) > a^* C_A'(a^*)$. If b^* is interior, or if $b^* = \bar{b}$ and the first-order condition (34) holds as an equality, the inequality $b^* C_B'(b^*) > a^* C_A'(a^*)$ contradicts (35) if $\rho = 1$. If $b^* = \bar{b}$ and the first-order condition holds as a strict inequality, i.e., $h'(\beta(a^*, b^*)) \frac{2\kappa a^* \rho}{(a^*+b^*)^2} - C_B'(b^*) > 0$, then this inequality and (33) together imply $b^* C_B'(b^*) < \rho a^* C_A'(a^*)$. A contradiction with the inequality $b^*C_B'(b^*) > a^*C_A'(a^*)$ is reached if $\rho = 1$. By continuity, the contradiction also obtains for $\rho = 1 + \varepsilon$ for some $\varepsilon > 0$. This proves the first statement. For the second statement, note first that the contradiction does not obtain for large values of ρ. A numerical example below (see Figure 2) suffices to prove the statement. □

Example 1. When the leader supporters have an absolute cost advantage, one might expect the leader to always win. The proposition confirms this intuition as long as the underdog supporters' stake ρ is not too large. However, as shown by way of an example in Figure 2, the leader may lose if ρ is large. In this figure, the backward bending curve shows, for each turnout b by the B-supporters, the turnout a which for each cost-sensitive A-supporter i maximizes her expected utility, given that all the other A-supporters also choose a. We will refer to such a value of a as an *A-consistent turnout*. In the figure, the dome-shaped curve shows the *B-consistent turnouts*, defined in a similar manner. A type-homogenous equilibrium (a^*, b^*) is such that a^* is A-consistent given b^*, and b^* is B-consistent given a^*. In other words, any (a^*, b^*) where the two curves intersect is an equilibrium. In the figure we thus see that there exists a unique equilibrium, in which the underdog wins, since (a^*, b^*) is above the dashed line, along which the turnouts are equal.

In sum, this result shows that if voters are equipped with *Homo moralis* preferences, the underdog supporters can overcome a seemingly unsurmountable challenge, as captured by their absolute cost disadvantage, if they perceive a high enough stake. In the example the underdog wins even though its supporters represent only $1.5/3.4 \approx 44\%$ of the electorate. Note further that approximately $1.313/(0.5 + 1.4) \approx 69\%$ of the A-supporters and $1.359/(0.4 + 1.1) \approx 91\%$ of the B-supporters participate in the election, although the degree of universalization is only $\kappa = 0.5$. In other words, full universalization is not necessary for high rates of participation to obtain.

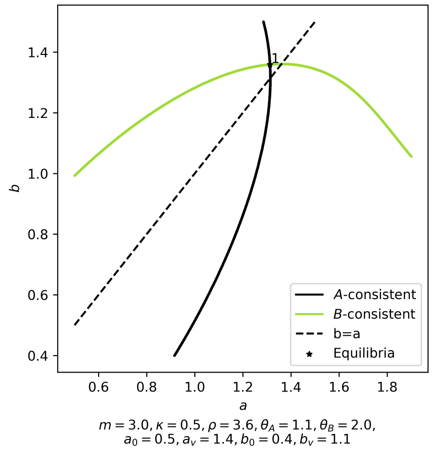


Figure 2: An example where the underdog wins in spite of the leader supporters's cost advantage ($\theta_A < \theta_B$, $a_v > b_v$, and $a_0 > b_0$), thanks to a high enough stake ρ.

We turn now to settings where the underdog supporters enjoy a cost advantage for some turnout rates, thanks to a larger base.

Proposition 3 (Partisan ethics)**.** *Suppose that the underdog's base exceeds that of the leader* ($b_0 > a_0$)*. Then, for some parameter values, there exist equilibria* (a^*, b^*) *in which the underdog wins* ($\alpha(a^*, b^*) < 0$)*, even if the stake is neutral* ($\rho = 1$)*.*

A numerical example proves this result:

Example 2. Figure 3 shows an equilibrium in which the underdog wins even though $\rho = 1$ and the two cost distributions are identical ($\theta_A = \theta_B$). Hence, the underdog's victory is here entirely driven by its greater base, $b_0 = 0.7 > 0.5 = a_0$.

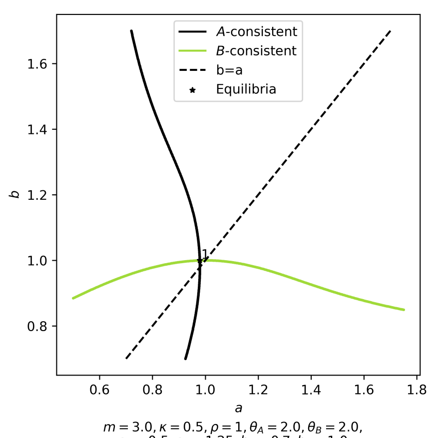


Figure 3: An example where the underdog wins thanks to a larger base ($b_0 > a_0$), in spite of a neutral stake ($\rho = 1$) and identical cost distributions ($\theta_A = \theta_B$).

To get a sense of how large a victory the underdog may obtain thanks to a larger base, a general result can be found when the relative frequency of cost-sensitive voters is the same for groups A and B, that is $a_0/a_v = b_0/b_v$, and all the cost-sensitive voters face the same cost distribution, that is $F_A = F_B = F$. The following proposition provides bounds on how different the equilibrium turnout rates in the two groups can then be, for interior equilibria. These bounds depend only on the stake ρ and are universal with respect to the form of the function h and to the value of the parameter κ. They are therefore valid whether the election is winner-take all, a random dictatorship, or anything in between, and for any degree of morality $\kappa \in (0, 1]$.

Proposition 4 (Partisan ethics)**.** *Suppose that* $\frac{a_0}{a_v} = \frac{b_0}{b_v}$*, that* $F_A(c) = F_B(c) = F(c)$ *for all* $c \in \mathbb{R}_+$*, and that* $\rho > 1$*. Then under partisan ethics, at any interior equilibrium* $(a^*, b^*) \in (a_0, \bar{a}) \times (b_0, \bar{b})$*:*

$$\frac{a^*}{a_v} < \frac{b^*}{b_v} < \rho \cdot \frac{a^*}{a_v}. \tag{36}$$

Proof. Given that a^* and b^* are interior, equation (35) applies, and writes:

$$\frac{b^*}{b_v} F^{-1}\left(\frac{b^*}{b_v} - r\right) = \rho \cdot \frac{a^*}{a_v} F^{-1}\left(\frac{a^*}{a_v} - r\right) \tag{37}$$

for $r = a_0/a_v = b_0/b_v$.

First, suppose that $\frac{b^*}{b_v} \leq \frac{a^*}{a_v}$, then: $F^{-1}\left(\frac{b^*}{b_v} - r\right) \geq \rho \cdot F^{-1}\left(\frac{a^*}{a_v} - r\right)$ and since $\rho > 1$, $F^{-1}\left(\frac{b^*}{b_v} - r\right) > F^{-1}\left(\frac{a^*}{a_v} - r\right)$. Because F is non-decreasing, this implies $\frac{b^*}{b_v} > \frac{a^*}{a_v}$, a contradiction. We conclude that indeed $\frac{b^*}{b_v} > \frac{a^*}{a_v}$.

Next, suppose that $\frac{b^*}{b_v} \geq \rho \frac{a^*}{a_v}$. Then, similar reasoning yields $F^{-1}\left(\frac{b^*}{b_v} - r\right) \leq F^{-1}\left(\frac{a^*}{a_v} - r\right)$ and $\frac{b^*}{b_v} \leq \frac{a^*}{a_v} < \rho \frac{a^*}{a_v}$, a contradiction, showing that indeed $\frac{b^*}{b_v} < \rho \frac{a^*}{a_v}$. □

The first inequality in (36) is similar to the "partial underdog compensation" result found in Herrera et al., 2016 (see the first part of their Proposition 1): the share of cost-sensitive underdog supporters who participate in the election is strictly greater than that of the leader supporters. We thus generalize their result by showing that it holds in a different setting than theirs, since (a) in their model there are only cost-sensitive voters ($a_0 = b_0 = 0$), (b) we do not restrict attention to settings with a unique equilibrium, and (c) we do not restrict attention to a particular functional form for the benefit function h. However, as shown in Proposition 2, the underdog may win if ρ is large enough, in which case a "full underdog compensation" arises. The second inequality in (36) puts an upper bound on this full compensation, which is increasing in the stake ρ. Note further that in the limit as ρ tends to the neutral case ($\rho = 1$), the two inequalities imply that the share of cost-sensitive underdog supporters who participate in the election tends to that of the leader supporters.

3.6 Partisan ethics: issues with equilibrium existence and multiplicity

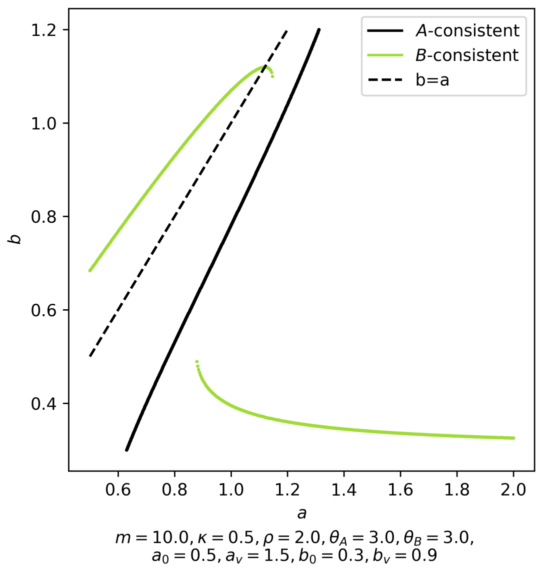
Having derived results on the properties of equilibria, should they exist, we now turn to the questions of existence and uniqueness. By means of examples, we will show that equilibria sometimes fail to exist, and that there sometimes are multiple equilibria.

Example 3. We begin by noting that equilibria may fail to exist. This is illustrated in Figure 4a, where the curve showing the A-consistent turnouts does not intersect with the curve showing the B-consistent turnouts.

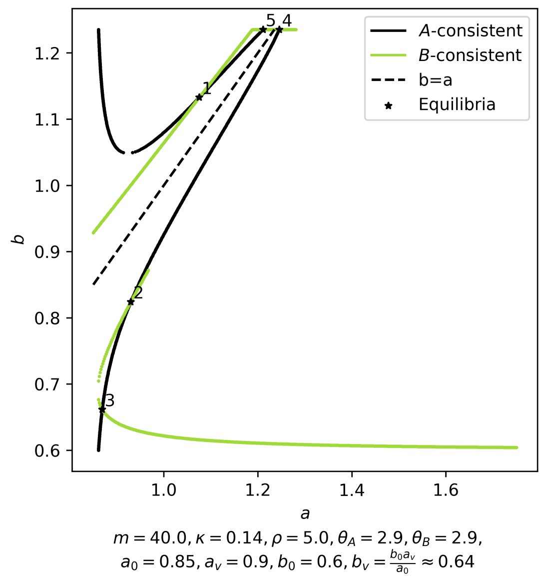
Example 4. However, in other settings there are many equilibria. In Figure 4b we show an example with five equilibria: two of them exhibit low turnouts and a victory by the leader, while the other three exhibit high turnouts and a victory by the underdog.

Figure 5 provides some insight into how the number of equilibria varies with the parameter values. The figures in the first column show the number of equilibria (a number that varies between 0 and 5). The figures in the second (respectively third) column then show, for the same parameter configurations, the number of equilibria in which the underdog (respectively the leader) wins the election. In the first three lines, it is the universalization parameter (κ) that varies along the horizontal axis: on the vertical axis it is the curvature parameter of the benefit function (m) that varies in the first line of figures, while it is the underdog's base (b_0) in the second line and the stake (ρ) in the third line of figures. The parameter values used for the first three lines correspond to the case considered in Proposition 4, and in each group there are almost as many cost-sensitive voters as core voters ($a_0/a_v = b_0/b_v \approx 0.94$). Moreover, in the first two lines, where the stake ρ is fixed, it is high enough to generate equilibria where the underdog wins ($\rho = 5$). Finally, in the fourth line of figures it is the sizes of the cost-sensitive constituencies (a_v on the horizontal axis and b_v on the vertical axis) that vary.

In the three first lines the following patterns appear. First, equilibrium multiplicity (respectively non-existence) appears only for sufficiently low (respectively high) values of κ. Second, the first line shows that when there exists a unique equilibrium, the leader wins when the value of the curvature parameter m is either low enough or high enough, while the underdog wins for the set of values of m in between. We further see that as κ increases, the interval of m-values for which the underdog wins gets closer to 0. Third, the size of the base b_0 has an unambiguously positive effect on the underdog's prospects of winning, as seen in the second line. Fourth, in the third line we see that, at least when the base of the leader exceeds that of the underdog, an increase in the stake ρ does not necessarily lead to an increase in the number of equilibria with a victory for the underdog.

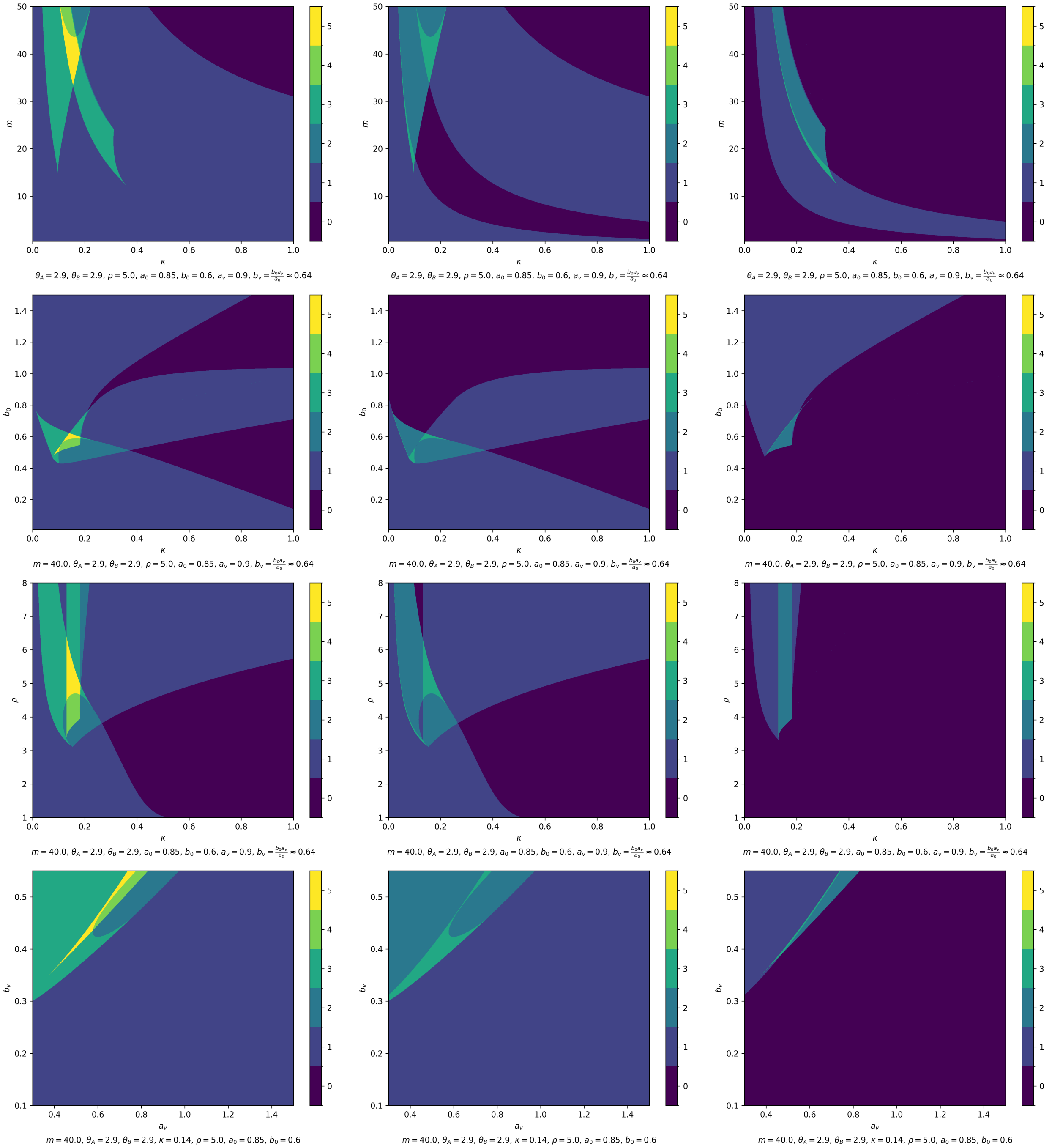


(a) An example of failure of equilibrium existence



(b) An example with five equilibria

Figure 4: Two extreme cases on multiplicity of equilibria



(a) Number of equilibria

(b) Number of equilibria with $a > b$ (A wins)

(c) Number of equilibria with $b > a$ (B wins)

Figure 5: Existence and multiplicity of equilibria, depending on (κ, m) (first line), (κ, ρ) (second line), (κ, b_0) (third line), and (a_v, b_v) (fourth line)

3.7 Partisan ethics: sufficient conditions for equilibrium existence and uniqueness

Here we identify sufficient conditions for there to exist a unique equilibrium (a^{*}, b^{*}) of the population game studied above. To prepare the ground for the statements and proofs, we define the auxiliary functions

$$\begin{aligned} \Phi_A(a,b) &= \kappa h(\alpha(a,b)) - C_A(a), &(38)\\ \Phi_B(a,b) &= \rho\kappa h(-\alpha(a,b)) - C_B(b). &(39)\end{aligned}$$

We also define the *auxiliary game*: this is a simultaneous-move game between two players, call them Alice and Bob, who have strategy sets $[a_0, \bar{a}]$ and $[b_0, \bar{b}]$, and payoff functions Φ_A and Φ_B, respectively.

In the following statements, by single-peaked, we mean that a function defined on $[a_0, \bar{a}]$ (or $[b_0, \bar{b}]$) is strictly increasing up to some $a \in (a_0, \bar{a})$ (or $b \in \left(b_0, \bar{b}\right)$), and strictly decreasing thereafter. In case it is differentiable, this amounts to its first-derivative being first strictly positive, crossing zero once from above, and then being strictly negative.

Assumption 1. *Assume that for all (a,b), $a^i \mapsto EU_A^{\kappa}(a,b,a^i)$ and $b^i \mapsto EU_B^{\kappa}(a,b,b^i)$ are single-peaked. Moreover, assume that for all b, $a \mapsto \Phi_A(a,b)$ is single-peaked, and for all a, $b \mapsto \Phi_B(a,b)$ is single-peaked.*

Lemma 3. *Under Assumption 1, (a^{*}, b^{*}) is an equilibrium of the population game if, and only if, it is a Nash equilibrium of the auxiliary game.*

By reducing the analysis to that of a standard two-player game, this Lemma facilitates identification of sufficient conditions for there to exist a unique equilibrium of the population game.

Assumption 2. *Let $h(x) = \frac{\arctan(mx)}{\arctan(m)}$, and assume that:*

1. $\kappa \in (0,1]$,
2. $\frac{f_k(c)c}{F_k(c)}$, $k = A, B$, *is decreasing,*
3. $m \leq 1$ *or for some* $r < \frac{2m}{(m-1)^2}$,

$$\lim_{c\to 0} \frac{F_A(c)}{c^r} > 0, \tag{40}$$

4. $m \leq 1$ *or for some* $r < \frac{2m}{\rho(m-1)^2}$,

$$\lim_{c \to 0} \frac{F_B(c)}{c^r} > 0, \tag{41}$$

5. $\bar{s}_B \geq \frac{\rho \bar{a}(m^2+1)}{2\kappa b_v m \arctan(m)}$, *and*

6. $\bar{s}_A \geq \frac{\bar{b}(m^2+1)}{2\kappa a_v m \arctan(m)}$.

Proposition 5. *Under Assumption 1, there exists a unique equilibrium of the population game. Moreover, Assumption 1 holds under Assumption 2.*

Conditions 2-6 of Assumption 2 are reminiscent of those that Herrera et al., 2016 adopt to ensure equilibrium existence and uniqueness in their model (see their "decreasing generalized reversed hazard rate (DGRHR) property" and their Condition 1). These conditions ensure that the density for low costs is large enough to avoid possible multiple peaks for low turnouts levels.

Observe that, for arbitrarily large m, part 3 and 4 of Assumption 2 require F_A and F_B to be arbitrarily steep at 0. Therefore, there cannot be a single continuous cost distribution on $[0, \infty)$ that satisfies Assumption 2 for all $m > 0$. A similar observation holds for arbitrarily large γ in Condition 1 of Herrera et al., 2016. Under both our and their assumptions, it is therefore impossible to guarantee existence and uniqueness while taking the limit $m \to \infty$ (or $\gamma \to \infty$, respectively). The next section deals with this limit whilst dropping Assumption 2. In fact, in case $a_0 = b_0$ and $\bar{a} \neq \bar{b}$, part 4 of Proposition 6 in the next section suggests that there cannot be a sufficient condition that implies existence independently of m.

3.8 Partisan ethics: the winner-take-all limit

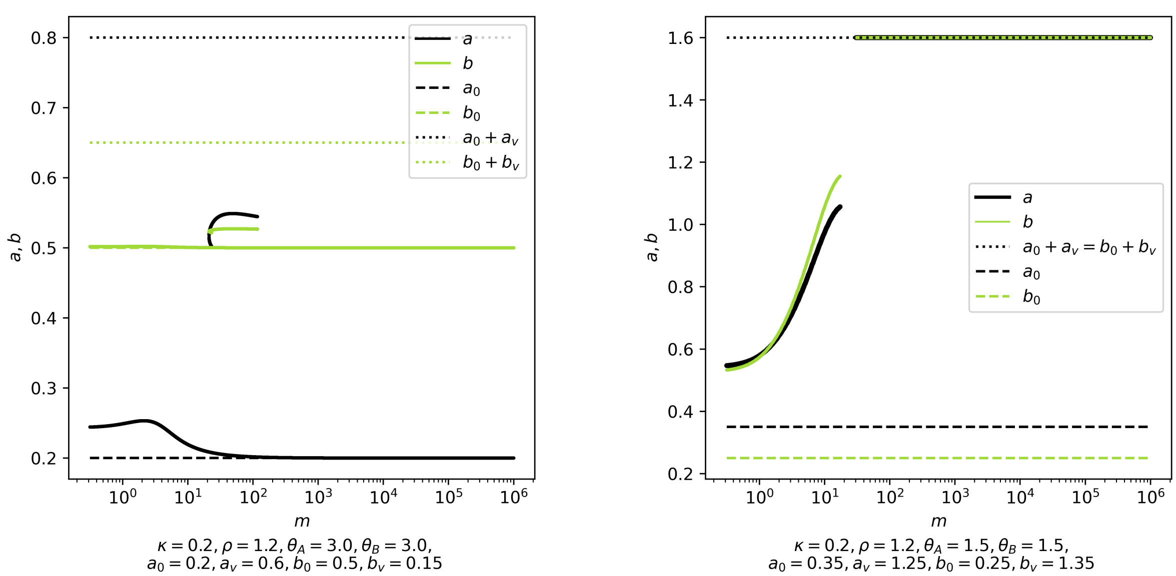
The pure winner-take-all case corresponds to a discontinuous step-function sign$(\cdot)$ in lieu of our function $h(\cdot)$. Our continuous model is not suitable for handling this case. Therefore, we proceed by approximation as follows (the details and proofs of the propositions can be found in the Appendix). A sequence of benefit functions h_t that all satisfy the hypothesis of our model is called an *approximating sequence* if the sequence converges to the winner-take-all benefit function sign$(\cdot)$. An outcome (a, b) is *sustained as a limit winner take all equilibrium* if there exists an approximating sequence h_t and a sequence (a_t, b_t) that converges to (a, b), with (a_t, b_t) an equilibrium for h_t, for all t.

Our first result states that, at equilibrium, only two situations can occur. In the first situation, which is also the generic one, costly participation is (in the limit) nil for both supporter groups, leaving the bases to determine the result of the election. In the second situation, participation is such that the result is (in the limit) tied and such that all (in the limit) of the voters turn out in one of the groups. This second situation can only occur when the two parameters $\bar{a}$ and $\bar{b}$ are equal.

Proposition 6 (Partisan ethics)**.** *Let (a, b) be sustained as a limit equilibrium of the winner-take-all case.*

1. *If $a_0 \neq b_0$ and $\bar{a} \neq \bar{b}$, $(a, b) = (a_0, b_0)$. (The underdog wins if $b_0 > a_0$ while the leader wins if $a_0 > b_0$.)*
2. *If $a_0 \neq b_0$ and $\bar{a} = \bar{b}$ either $(a, b) = (a_0, b_0)$ or $(a, b) = \left(\bar{b}, \bar{b}\right)$.*
3. *If $a_0 = b_0$ and $\bar{a} = \bar{b}$, $(a, b) = \left(\bar{b}, \bar{b}\right)$.*
4. *If $a_0 = b_0$ and $\bar{a} \neq \bar{b}$, no such pair (a, b) exists.*

Example 5. While the proposition does not make any claims about existence, we provide some illustrating examples with existence, for increasingly large values of m. Interestingly, even in the winner-take-all limit, there may exist equilibria in which the underdog wins. We illustrate this in Figure 6a, which also provides numerical evidence for existence in the first case of Proposition 6. In Figure 6b, we illustrate the second case of Proposition 6.



(a) Case 1 of Proposition 6

(b) Case 2 of Proposition 6

Figure 6: Possible equilibrium turnouts a and b depending on m

Our second result shows that in the case examined in Proposition 4, if the stake (ρ) is smaller than the ratio between the sizes of the two groups of cost-sensitive voters, then costly participation is (approximately) nil for both.

Proposition 7 (Partisan ethics)**.** *Suppose that* $\frac{a_0}{a_v} = \frac{b_0}{b_v}$ *and* $F_A(c) = F_B(c) = F(c)$ *for all* $c \in \mathbb{R}_+$. *Let* $r = a_v/b_v = a_0/b_0$. *If* $\rho < r$ *then, in the ex post setting, only* (a_0, b_0) *can be sustained as a limit winner-take-all equilibrium, and the leader wins.*

4 Non-partisan ethics (*ex ante* setting)

We here follow Harsanyi's view by considering non-partisan ethics.[11] In the non-partisan setting a cost-sensitive voter i chooses a strategy which is a pair of thresholds $s^i = (s^i_A, s^i_B) \in [0, \infty]^2$. This strategy means that the voter abstains when her cost for voting is larger than s^i_A if she prefers candidate A, and when her cost for voting is larger than s^i_B if she prefers candidate B ($s^i_A = \infty$ respectively $s^i_B = \infty$ means that she votes independent of the realized cost). Two interpretations are possible. In the first, there is *ex ante* uncertainty regarding the candidate that i prefers, and she selects the strategy behind the veil of ignorance, before this uncertainty is resolved. This may well describe how independent voters reason. In the second interpretation, there is no such uncertainty, but due to her ethical concern the individual adopts the viewpoint of Harsanyi's impartial observer, by inserting a veil of ignorance in her reasoning. Whatever interpretation is chosen, i selects the strategy $s^i = (s^i_A, s^i_B)$ before knowing her actual cost of voting. Voters have correct beliefs about the voting cost distributions, described in Section 2. Each individual votes for A or B, or abstains. Individual i in group A with realized cost c^i_A votes for A if and only if $c^i_A \leq s^i_A$. The corresponding assumptions are made for group B.

4.1 Non-partisan ethics: payoff computations

As in the partisan setting, we are looking for homogenous equilibria, which here means that all the cost-sensitive voters choose the same strategy. At a homogenous equilibrium $s = (s_A, s_B)$, the realized turnouts are $a(s) = a_0 + a_v F_A(s_A)$ and $b(s) = b_0 + b_v F_B(s_B)$,

[11]Most models in the literature stick to partisan ethics. A notable exception is Feddersen and Sandroni, 2006, who adopt a mix of partisan and non-partisan ethics: the group-optimal cost threshold is obtained by maximizing the material benefit that accrues to the group net of the expected cost for both groups.

respectively, for the two candidates, and the following relative vote margins obtain:

$$\alpha(s) = \frac{a(s_A) - b(s_B)}{a(s_A) + b(s_B)}, \quad \beta(s) = -\alpha(s), \tag{42}$$

so that each voter obtains expected net material benefit

$$EU(s) = \frac{a_v}{a_v + b_v} h(\alpha(s)) + \frac{b_v}{a_v + b_v} \rho h(\beta(s)) \tag{43}$$
$$- \frac{a_v}{a_v + b_v} \int_{c=0}^{s_A} c f_A(c) \, \mathrm{d}c - \frac{b_v}{a_v + b_v} \int_{c=0}^{s_B} c f_B(c) \, \mathrm{d}c.$$

With a non-partisan ethic, the voter takes into account both the expected benefits and the expected costs that the thresholds s_A and s_B entail; the benefits are weighed by the relative population shares of the groups, to reflect the *ex ante* perspective that the voter adopts. Henceforth, we will without loss of generality drop the constant positive factor $1/(a_v + b_v)$. Since $h(-x) = -h(x)$, the expected net material benefit can then be rewritten as follows:

$$EU(s) = (a_v - \rho b_v) \, h(\alpha(s)) - a_v \int_{c=0}^{s_A} c f_A(c) \, \mathrm{d}c - b_v \int_{c=0}^{s_B} c f_B(c) \, \mathrm{d}c. \tag{44}$$

This gives the following expression for the expected utility of a voter i with *Homo moralis* preferences, given that all the other voters use strategy s:

$$EU^{\kappa}(s, s^i) = (a_v - \rho b_v) \, h(\alpha^{\kappa}(s, s^i)) - a_v \int_{c=0}^{s_A^i} c f_A(c) \, \mathrm{d}c - b_v \int_{c=0}^{s_B^i} c f_B(c) \, \mathrm{d}c. \tag{45}$$

where

$$\alpha^{\kappa}(s, s^i) = \frac{(1-\kappa)[a(s_A) - b(s_B)] + \kappa[a(s_A^i) - b(s_B^i)]}{(1-\kappa)[a(s_A) + b(s_B)] + \kappa[a(s_A^i) + b(s_B^i)]}. \tag{46}$$

Applying the same change of variables as we did under partisan ethics, we henceforth assume that an individual i's strategy is a pair $(a^i, b^i) \in [a_0, \bar{a}] \times [b_0, \bar{b}]$, and we write (a, b) for the strategy used by the other voters at a homogenous equilibrium, so that the expected utility of i writes

$$EU^{\kappa}(a, b, a^i, b^i) = (a_v - \rho b_v) \cdot h(\alpha^{\kappa}\left(a, b, a^i, b^i\right)) - a_v C_A\left(a^i\right) - b_v C_B\left(b^i\right), \tag{47}$$

where

$$\alpha^{\kappa}\left(a, b, a^i, b^i\right) = \frac{(1-\kappa)\, a + \kappa a^i - (1-\kappa) b - \kappa b^i}{(1-\kappa)\, a + \kappa a^i + (1-\kappa)\, b + \kappa b^i}. \tag{48}$$

Under our assumptions, this is a continuously differentiable function of (a^i, b^i).

4.2 Non-partisan ethics: equilibrium existence

By contrast to the partisan setting, here we can establish general equilibrium existence, thanks to the aforementioned change of variables together with Lemma 1.

Proposition 8 (Non-partisan voters)**.** *An equilibrium always exists.*

Proof. Consider the auxiliary function

$$\Phi(a, b) = \kappa\lambda h(\alpha(a, b)) - a_v C_A(a) - b_v C_B(b) \tag{49}$$

where $\lambda = a_v - \rho b_v$. It takes values in $\mathbb{R} \cup \{-\infty\}$. We begin by showing that (a^*, b^*) is an equilibrium if it is a global maximum of Φ. Thus, let (a^*, b^*) be a point where Φ reaches its maximum, and suppose, by contradiction, that there exists (a', b') such that $EU^\kappa(a^*, b^*, a', b') > EU(a^*, b^*)$, that is:

$$a_v C_A(a^*) + b_v C_B(b^*) - a_v C_A(a') - b_v C_B(b') > \lambda \left[h(\alpha(a^*, b^*)) - h(\alpha(a^\kappa, b^\kappa))\right] \tag{50}$$

for $a^\kappa = (1-\kappa) a^* + \kappa a'$ and $b^\kappa = (1-\kappa) b^* + \kappa b'$. Since (a^*, b^*) maximizes Φ, we have $\Phi(a^\kappa, b^\kappa) \leq \Phi(a^*, b^*)$, which writes:

$$a_v C_A(a^*) + b_v C_B(b^*) - a_v C_A(a^\kappa) - b_v C_B(b^\kappa) \leq \kappa\lambda \left[h(\alpha(a^*, b^*)) - h(\alpha(a^\kappa, b^\kappa))\right]. \tag{51}$$

Combining the two previous equations we find, upon rearranging the terms,

$$(1-\kappa) a_v C_A(a^*) + \kappa a_v C_A(a') + (1-\kappa) b_v C_B(b^*) + \kappa b_v C_B(b') < a_v C_A(a^\kappa) + b_v C_B(b^\kappa). \tag{52}$$

This contradicts the convexity of the functions C_A, C_B (see Lemma 1).

The second part of the proof consists in showing that Φ admits a maximum, which is a sufficient condition for an equilibrium to exist, given the first part of the proof. To see this, first note that $h(\alpha(a, b))$ is continuous on $[a_0, \bar{a}] \times [b_0, \bar{b}]$ and takes values in $\mathbb{R}$. Moreover, in the case where C_A, C_B are continuous functions on $[a_0, \bar{a}] \times [b_0, \bar{b}]$ taking values in $\mathbb{R}$, one can conclude by the extreme value theorem, observing that Φ is continuous and takes values in $\mathbb{R}$. In the remaining case, where $C_A(\bar{a}) = \infty$ or $C_B\left(\bar{b}\right) = \infty$, we have that $C_A(a) \to \infty$ for $a \to \bar{a}$ or $C_B(b) \to \infty$ for $b \to \bar{b}$. Then, since h is bounded on $[a_0, \bar{a}] \times [b_0, \bar{b}]$, one can find some $\varepsilon_A, \varepsilon_B \geq 0$ such that Φ takes real values on $[a_0, \bar{a} - \varepsilon_A] \times [b_0, \bar{b} - \varepsilon_B]$ and such that

$\Phi(a,b) < \Phi\left(\min(a, \bar{a}-\varepsilon_A), \min\left(b, \bar{b}-\varepsilon_B\right)\right)$, which allows to conclude using the extreme value theorem on $[a_0, \bar{a}-\varepsilon_A] \times [b_0, \bar{b}-\varepsilon_B]$. □

4.3 Non-partisan ethics: equilibrium properties

We first show that at equilibrium, exactly one group incurs positive voting costs (except in two knife-edge cases, in which nobody votes).

Proposition 9 (Non-partisan ethics)**.** *If $\kappa = 0$ and/or $a_v = \rho b_v$, then $(a^*, b^*) = (a_0, b_0)$ is the unique equilibrium, while if $\kappa \in (0,1]$ then:*

- *if $a_v > \rho b_v$, any equilibrium is such that $a^* > a_0$ and $b^* = b_0$;*
- *if $\rho b_v > a_v$ any equilibrium is such that $a^* = a_0$ and $b^* > b_0$.*

Proof. Given that all other voters use strategy (a,b), individual i's expected marginal utility from a^i is,

$$\begin{aligned}\frac{\partial}{\partial a^i} EU^\kappa\left(a,b,a^i,b^i\right) &= (a_v - \rho b_v) h'(\alpha^\kappa(a,b,a^i,b^i)) \frac{\partial \alpha^\kappa(a,b,a^i,b^i)}{\partial a^i} - a_v C_A'(a^i) \\ = (a_v - \rho b_v)\, h'\left(\alpha^\kappa(a,b,a^i,b^i)\right) & \frac{2\kappa[(1-\kappa)b + \kappa b^i]}{[(1-\kappa)\,a + \kappa a^i + (1-\kappa)b + \kappa b^i]^2} - a_v C_A'\left(a^i\right) \end{aligned} \tag{53}$$

and the expected marginal utility from b^i is:

$$\begin{aligned}\frac{\partial}{\partial b^i} EU^\kappa\left(a,b,a^i,b^i\right) &= (\rho b_v - a_v) h'(\beta^\kappa(a,b,a^i,b^i)) \frac{\partial \beta^\kappa(a,b,a^i,b^i)}{\partial b^i} - b_v C_B'(b^i) \\ = (\rho b_v - a_v)\, h'\left(\beta^\kappa(a,b,a^i,b^i)\right) & \frac{2\kappa[(1-\kappa)a + \kappa a^i]}{[(1-\kappa)\,a + \kappa a^i + (1-\kappa)b + \kappa b^i]^2} - b_v C_B'\left(b^i\right). \end{aligned} \tag{54}$$

Since $C_A'(a_0) = C_B'(b_0) = 0$, and $C_A(a^i), C_B(b^i) > 0$ for $a_i > 0$ or $b_i > 0$ (see (23)), a best response (a^i, b^i) to (a,b) has $a^i > a_0$ if and only if the first of the two terms in (53) is strictly positive. Since h is strictly increasing, and $b \geq b_0 > 0$, this is true if and only if $\kappa(a_v - \rho b_v) > 0$; otherwise, any best response has $a^i = a_0$. Likewise, a best response (a^i, b^i) to (a,b) has $b^i > b_0$ if and only if $\kappa(\rho b_v - a_v) > 0$; otherwise, any best response has $b^i = b_0$. These arguments prove the proposition. □

This proposition shows that both the degree of universalization (κ) and the stake of the election for the underdog supporters (ρ) matter for the qualitative nature of the set of equilibria. First, if voters are driven solely by instrumental motives ($\kappa = 0$) or if the expected benefit that one group obtains from a positive margin of its candidate exactly outweighs the expected cost that the other group garners from this margin ($\rho = a_v/b_v$), then turnout is confined to the bases a_0 and b_0, in which case the underdog wins if and only if it has a larger base than the leader ($b_0 > a_0$). Second, whenever $a_v \neq \rho b_v$, any positive degree of universalization $\kappa > 0$ triggers participation of a positive mass of cost-sensitive voters. The reason is clear: a $\kappa > 0$ triggers in the individual voter a utility kick from contemplating the margin that her preferred candidate would obtain if all the other cost-sensitive voters selected the same strategy as herself; the voter is willing to incur a positive voting cost to obtain this utility kick. Third, and in stark contrast with the partisan setting, here a voter internalizes the negative externality that voting for one candidate has on the group supporting the other candidate. Hence, she votes only if she belongs to the group that obtains the highest expected benefit from its candidate's margin.

Henceforth we examine only non-trivial settings where $\kappa(a_v - \rho b_v) \neq 0$. To begin, consider the case $a_v > \rho b_v$ and define

$$\begin{aligned} A(a^*) &\equiv \frac{\partial}{\partial a^i} EU^\kappa \left(a, b, a^i, b^i\right)|_{a^i=a^*, b=b^i=b_0} \\ &= (a_v - \rho b_v) h'(\alpha(a^*, b_0)) \frac{2\kappa b_0}{(a^* + b_0)^2} - a_v C_A' \left(a^*\right). \end{aligned} \tag{55}$$

The necessary first-order condition for any equilibrium a^* is thus $A(a^*) \geq 0$, which must hold as an equality if a^* lies in the interior $(a_0, \bar{a})$. The necessary second-order condition for such an interior solution is:

$$\begin{aligned} \frac{\partial^2}{\partial (a^i)^2} EU^\kappa \left(a, b, a^i, b^i\right)|_{a^i=a^*, b=b^i=b_0} &= (a_v - \rho b_v) h''(\alpha(a^*, b_0)) \frac{4\kappa^2 b_0^2}{(a^* + b_0)^4} \\ &- (a_v - \rho b_v) h'(\alpha(a^*, b_0)) \frac{4\kappa^2 b_0}{(a^* + b_0)^3} \\ &- a_v C_A'' \left(a^*\right) \leq 0. \end{aligned} \tag{56}$$

Likewise, for $\rho b_v > a_v$, define

$$\begin{aligned} B(b^*) &\equiv \frac{\partial}{\partial b^i} EU^\kappa \left(a, b, a^i, b^i\right)|_{a=a^i=a_0, b^i=b^*} \\ &= (\rho b_v - a_v) h'(\beta(a_0, b^*)) \frac{2\kappa a_0}{(a_0 + b^*)^2} - b_v C_B' \left(b^*\right), \end{aligned} \tag{57}$$

so that the necessary first-order condition for any equilibrium b^* is $B(b^*) \geq 0$, which must hold as an equality if b^* lies in the interior $(b_0, \bar{b})$. The necessary second-order condition for such an interior solution is:

$$\begin{aligned}\frac{\partial^2}{\partial (b^i)^2} EU^\kappa\left(a,b,a^i,b^i\right)|_{a=a^i=a_0,b^i=b^*} &= (\rho b_v - a_v)h''(\beta(a_0,b^*))\frac{4\kappa^2 a_0^2}{(a_0+b^*)^4} \\ &- (\rho b_v - a_v)h'(\beta(a_0,b^*))\frac{4\kappa^2 a_0}{(a_0+b^*)^3} \\ &- b_v C_B''(b^*) \leq 0. \end{aligned} \quad (58)$$

If there is a unique a^* (respectively b^*) satisfying both the first-order and second-order conditions, then it is the unique equilibrium. The following proposition identifies a sufficient condition for this to obtain.

Proposition 10 (Non-partisan ethics). *Suppose that $\kappa \in (0,1]$. If $a_v > \rho b_v$ and $a_0 \geq b_0$, there is a unique equilibrium (a^*, b_0). At this equilibrium, the leader wins: $\alpha(a^*, b_0) > 0$. Likewise, if $\rho b_v > a_v$ and $b_0 \geq a_0$, there is a unique equilibrium (a_0, b^*). At this equilibrium, the underdog wins: $\beta(a_0, b^*) > 0$.*

Proof. It is sufficient to prove the result for one of the cases, say $\rho b_v > a_v$. We begin by proving the following claim: if $h''(\beta(a_0,b)) \leq 0$ for all $b \in [b_0, \bar{b}]$ there either exists a unique $b^* \in (b_0, \bar{b})$ satisfying $B(b^*) = 0$ and such that (58) holds strictly, or $B(b) > 0$ for all $b \in \left(b_0, \bar{b}\right)$. To see this, note first that if $h''(\beta(a_0,b)) \leq 0$ for all $b \in [b_0, \bar{b}]$, then the first term in $B(b^*)$ is non-increasing in b^*; this term is also strictly positive for $b^* = b_0$ (since $h' > 0$). Since the second term equals 0 for $b^* = b_0$ and is strictly increasing in b^*, the claim follows.

If $b_0 \geq a_0$, the two statements in the proposition then follow immediately from the fact that $\rho b_v > a_v$ implies $b^* > b_0$ and $a^* = a_0$. Indeed, $\beta(a_0, b)$ is thus strictly positive for any $b \in (b_0, \bar{b})$, and our assumptions on h then imply that $h''(\beta(a_0,b)) \leq 0$ for any $b \in (b_0, \bar{b}]$. □

This proposition again underlines the crucial role played by the bases, a_0 and b_0. If the cost-sensitive voters who do participate in the election can rely on a base that is larger than that of the other candidate, then they win the election independent of their turnout rate. This implies a decreasing marginal benefit and an increasing marginal cost of increases in the turnout, which in turn implies equilibrium uniqueness. Proposition 12 provides a more general result on equilibrium uniqueness that relies on the same idea.

We turn now to settings where the group that votes (i.e., the leader supporters if $a_v > \rho b_v$ and the underdog supporters if $\rho b_v > a_v$) has a smaller base than the other group, in which case the marginal benefit is increasing for turnout rates close enough to the base, implying that there may be multiple candidates a^* (respectively b^*) satisfying the first- and second-order conditions. Each such candidate is an equilibrium if there do not exist utility-enhancing global deviations. The main question we investigate is whether equilibrium uniqueness obtains. Examination of the special case of full universalization provides some initial insights.

Proposition 11 (Non-partisan ethics)**.** *Suppose that* $\kappa = 1$*. If there exist multiple equilibria, they all generate the same expected utility to the cost-sensitive voters who do turn out to vote.*

Proof. It is sufficient to prove the result for one of the cases, say $\rho b_v > a_v$. Plugging in $a = a_0$ and $\kappa = 1$ into the expected utility (47), the expected utility becomes independent of b, the turnout rate among the other voters, and thus a function of b^i only:

$$EU^{\kappa}(a_0, b, a_0, b^i) = (\rho b_v - a_v) h \left(\frac{b^i - a_0}{a_0 + b^i} \right) - C_B(b^i). \tag{59}$$

Hence, each individual voter simply chooses some value of b^i that maximizes this expression. If there are multiple solutions, they must yield the same expected utility. □

This proposition suggests that full universalization can generate multiple turnout rates. The following example, for the case $\rho b_v > a_v$, further shows that full universalization can does not guarantee a high turnout rate.

Example 6. As shown in Figure 7, if $b_v = 0.609$ and the other parameter values are as specified in the figure legend, there are two equilibria, shown as stars: the underdog wins at one of them ($b^* \approx 0.6 > a_0$) but loses at the other (b^* is close to $b_0 = 0.2$). The figure further shows that small variations in the parameter values may induce discrete jumps in b^*. Indeed, for b_v slightly above 0.609, there is a unique equilibrium turnout, at which the underdog wins, while for b_v slightly below 0.609, there is a unique equilibrium turnout, close to b_0.

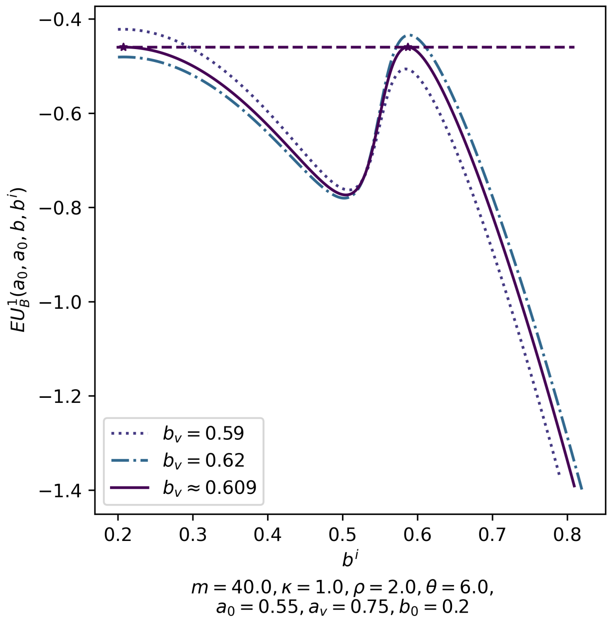


Figure 7: Existence of two equilibria when $\kappa = 1$

While multiplicity of equilibria appears only in knife-edge cases under full universalization ($\kappa = 1$), we will now see that it is a quite common phenomenon under partial universalization ($\kappa \in (0, 1)$).

Example 7. Considering still the case $\rho b_v > a_v$, Figure 8 shows, for $\kappa = 0.4$, an example with two equilibria, indicated by stars. Like in the example under full universalization in Figure 7, here one equilibrium turnout is close to the base, $b^* \approx 0.12$, while the other makes the underdog win, $b^* \approx 0.49 > a_0 = 0.45$. This victory obtains despite the underdog's base being weak compared to that of the leader (compare $b_0 = 0.1$ to $a_0 = 0.45$). We further see in the figure that the high-turnout equilibrium gives a substantially higher expected utility than the low-turnout one.

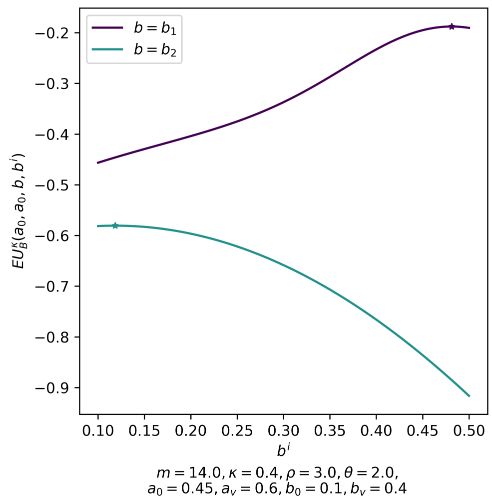


Figure 8: Utility from deviating from each equilibrium candidate

In the preceding example the underdog supporters face a coordination problem, and they would prefer to coordinate on the high-turnout equilibrium. A question of interest is whether such coordination problems — i.e., co-existence of equilibria with substantially different turnouts, where one equilibrium is preferred to the other(s) — are common. We here examine the necessary conditions for global deviations not to exist. This will provide some insights generate insights about parameter values that might give rise to such multiplicity of equilibria. We do this for the case $\rho b_v > a_v$.

Thus, consider some turnout rate $b \in (b_0, \bar{b})$. For b to be an equilibrium, an individual voter must not wish to deviate to any $b' \neq b$. Considering first downward deviations $b' < b$, the following condition must hold:

$$(\rho b_v - a_v)\left[h(\beta(a_0, b)) - h(\beta^\kappa(a_0, b, a_0, b'))\right] \geq C_B(b) - C_B(b') \quad \forall b' \in [b_0, b), \tag{60}$$

where

$$\beta\left(a_0, b\right) = \frac{b - a_0}{a_0 + b} \tag{61}$$

and

$$\beta^\kappa\left(a_0, b, a_0, b'\right) = \frac{(1-\kappa)b + \kappa b' - a_0}{a_0 + (1-\kappa)\, b + \kappa b'}. \tag{62}$$

Any downward deviation $b' < b$ reduces the cost, i.e., the right-hand side of (60) is strictly positive. For any $b' < b$ the left-hand side is equal to zero if $\kappa = 0$, and it is increasing in κ: the utility loss that the voter incurs from a decline in its preferred candidate's margin gets larger as her degree of universalization gets larger. Hence, any value of $\kappa > 0$ imposes an upper bound on the turnout rate that can be sustained in equilibrium. In particular, the voter must not be tempted by abstention ($b' = b_0$), the deviation that would maximize the cost saving $C_B(b) - C_B(b')$, and we note that the deviation to abstention generates a cost saving that is larger the smaller is the base b_0. Taken together, these observations suggest that the underdog supporters can achieve a victory only if κ is large enough, and that this constraint on κ is stronger the weaker is the base b_0. Noting further that C_B is decreasing in the size of the cost-sensitive electorate b_v (see (21)), *ceteris paribus* the constraint on κ is also stronger the smaller is b_v. Finally, (60) clearly implies that the stake for the B-supporters (ρ) must be large for an equilibrium with a higher turnout to be sustained.

Considering now upward deviations $b' > b$, the following condition must hold for an individual not to wish to deviate:

$$C_B(b') - C_B(b) \geq (\rho b_v - a_v)\left[h(\beta^\kappa(a_0, b, a_0, b')) - h(\beta(a_0, b)\right] \quad \forall b' \in (b, \bar{b}]. \tag{63}$$

Any upward deviation $b' > b$ raises the cost, i.e., the left-hand side is strictly positive. But it also raises the utility gain that the voter obtains from an increase in its preferred candidate's margin, as long as her degree of universalization is strictly positive: the right-hand side equals zero if $\kappa = 0$ and is increasing in κ. Hence, for any $\kappa > 0$ there is a lower bound on the turnout rate that can be sustained in equilibrium. In particular, if the underdog has a small base b_0 and κ is close enough to 1 — so that the right-hand side of (63) is large — we should expect existence of equilibria with a turnout rate close to b_0 only if voting costs are high enough. Since C_B is decreasing in the size of the cost-sensitive electorate (b_v), *ceteris paribus* low turnout equilibria also require b_v to be small enough. Finally, (63) clearly implies that low turnout equilibria are more likely to be sustained the smaller is the stake ρ.

Taken together, the preceding arguments suggest that the aforementioned coordination problem should be expected only if κ is neither too large nor too small, $a_0 - b_0$ is large enough, and $\rho b_v - a_v$ is neither too large nor too small.

Recalling that the same arguments apply to the case $a_v > \rho b_v$, the leader's supporters may also face a coordination problem: if the leader's base a_0 is small enough compared to that of the underdog b_0, and κ is moderate, then there may exist two equilibria, one with a high turnout and with a low turnout, and the leader may suffer a sizeable loss in the latter.

Example 8. We illustrate this in Figure 9, which shows the set of equilibrium turnouts for the leader supporters, as a function of a_0, and for three values of κ. For $\kappa = 1$, we see that there is a unique equilibrium for any value of a_0, while for $\kappa = 0.8$ and $\kappa = 0.5$, there are two equilibria for small values of a_0 and a unique equilibrium for large enough values of a_0. The leader wins at equilibria above the dashed line, which corresponds to $a = b_0$. The figure thus shows that if the base a_0 is small, the leader supporters face a coordination problem: they may win or lose. By contrast, a victory for the leader is guaranteed if the base a_0 is large enough.

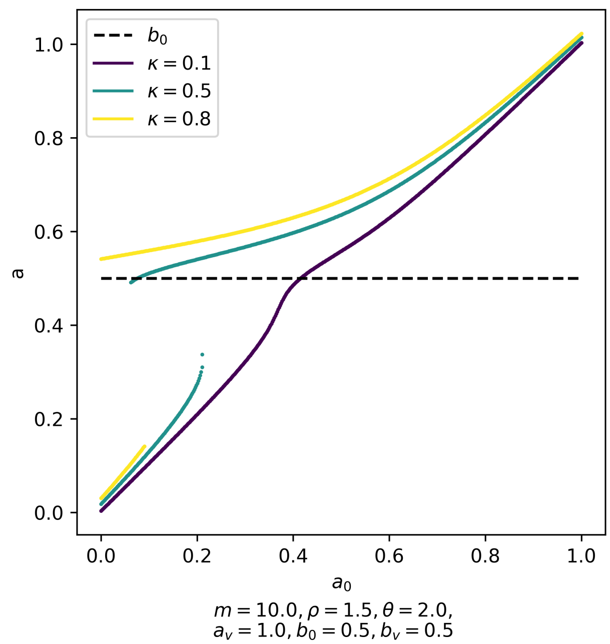


Figure 9: Set of equilibrium turnouts a for different values of a_0

We now examine whether there may be even more than two equilibria.

Example 9. Returning to settings where $\rho b_v - a_v > 0$, Figure 10 shows an example with three equilibria. Like for the example with two equilibria above (recall Figure 8), here the expected utility is higher the higher is the equilibrium turnout.

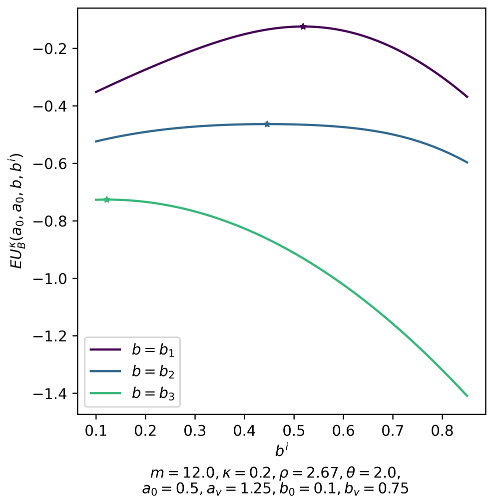


Figure 10: Utility from deviation for each equilibrium candidate

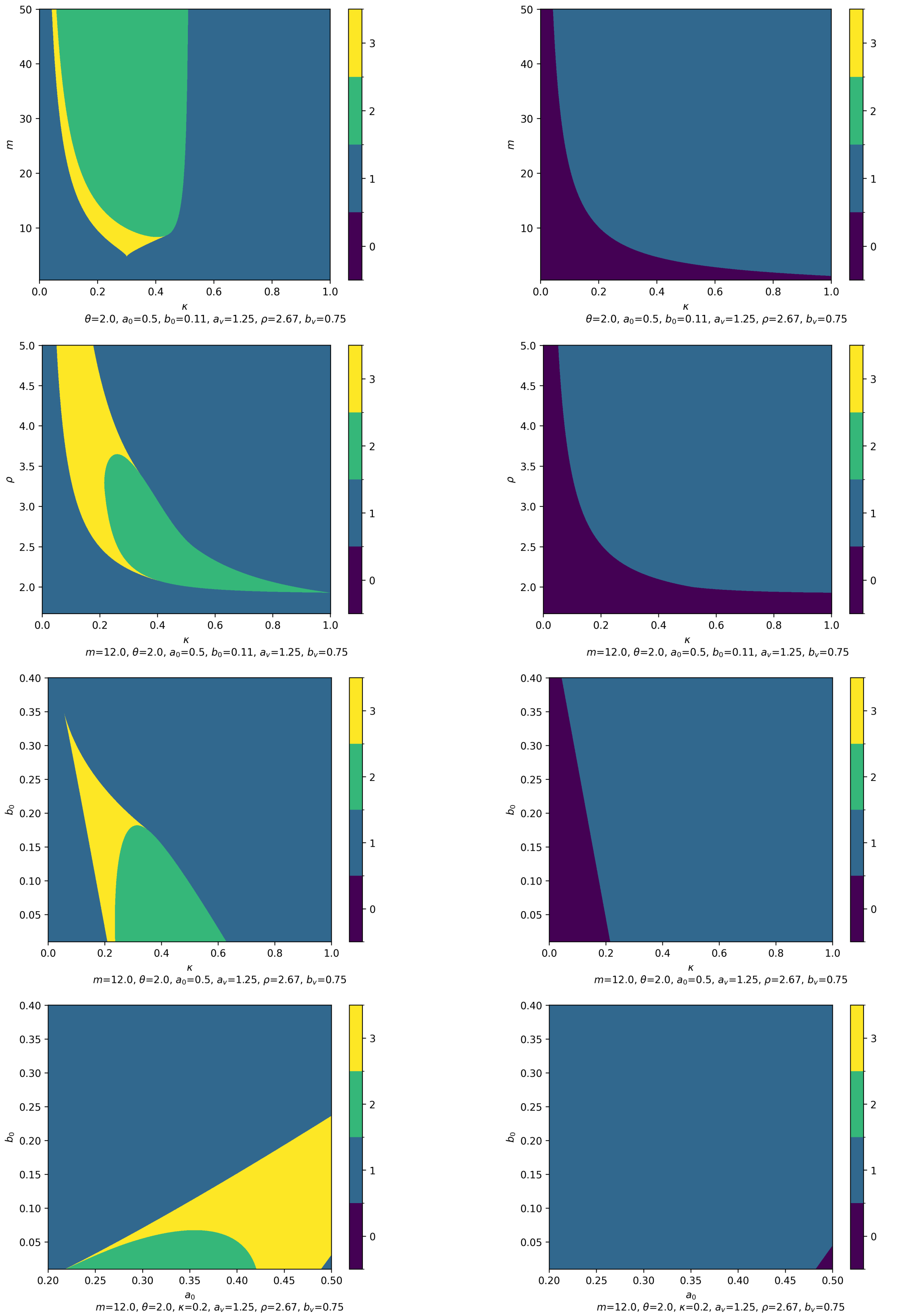
By contrast to the partisan setting, however, in our numerical examples we did not identify any parameter values for which there are more than three equilibria.

Example 10. In Figure 11, we still examine the case $\rho b_v > a_v$ and we vary two parameters at a time. Then, we plot the number of equilibria in panel (a) and the number of equilibria such that the underdog wins in panel (b).

The first line of figures shows how the set of equilibria varies with the degree of universalization κ and the stake ρ. For high (resp. low) enough values of ρ and κ there is a unique equilibrium, in which the underdog wins (resp. loses). The coordination problem appears either if κ is not very high but ρ is, or the reverse, and it is the combination of a modest κ and a high ρ that favors the appearance of more than two equilibria.

The second line of figures then shows how the set of equilibria varies with the degree of universalization κ and m, the curvature parameter for the h function (see (8)). For low enough values of m, the expected utility is concave and equilibrium uniqueness obtains. Multiplicity of equilibria appears for a value of m around 5.

Note that the first two lines of figures confirm one of the conclusions from our analysis of global deviations above: multiplicity of equilibria appears only for values of κ neither too close to 0, nor too close to 1. One exception appears in the second line, however, where a value of ρ slightly below 2 corresponds to a knife-edge case with two equilibria for $\kappa = 1$.



(a) Number of equilibria

(b) Number of equilibria with $b > a_0$ (B wins)

Figure 11: Multiplicity of equilibria, depending on (κ, m) (first line), (κ, ρ) (second line), (κ, b_0) (third line), and (a_0, b_0) (fourth line)

4.4 Non-partisan ethics: sufficient conditions for equilibrium uniqueness

Sufficient conditions for uniqueness are analogous to the partisan case described in Section 3.7. We will show that uniqueness holds if the auxiliary function from the proof of Proposition 8 is single-peaked in a suitable sense. Recall

$$\begin{aligned}\Phi(a,b) &= \kappa a_v h(\alpha(a,b)) + \rho\kappa b_v h(\beta(a,b)) - a_v C_A(a) - b_v C_B(b) \\ &= \kappa(a_v - \rho b_v) h(\alpha(a,b)) - a_v C_A(a) - b_v C_B(b) \\ &= \kappa(\rho b_v - a_v) h(\beta(a,b)) - a_v C_A(a) - b_v C_B(b)\end{aligned} \tag{64}$$

As opposed to the partisan case (see Section 3.7) it is sufficient to assume single-peakedness in the following sense: a univariate function is called single-peaked if it is strictly increasing until reaching its unique maximum, where the latter could be equal to $\bar{b}$. Recall also Proposition 9: if $\rho b_v > a_v$, any equilibrium is such that $a = a_0$, and vice-versa, if $a_v > \rho b_v$, any equilibrium is such that $b = b_0$.

Proposition 12. *Let $\kappa > 0$. If $\rho b_v > a_v$ and $b \mapsto \Phi(a_0, b)$ is single-peaked, there exists a unique equilibrium. Similarly, if $a_v > \rho b_v$ and $a \mapsto \Phi(a, b_0)$ is single-peaked, there exists a unique equilibrium.*

Proof. We restrict ourselves to the case $\rho b_v > a_v$, as the proof in the opposite case goes analogously. By Proposition 9, $a = a_0$, so that any equilibrium is entirely described by b, and $b > b_0$.

Let b be an equilibrium. It satisfies

$$\left.\frac{\partial}{\partial b^i} EU^\kappa\left(a_0, b, a_0, b^i\right)\right|_{b^i=b} = \frac{\partial}{\partial b}\Phi(a_0, b) = 0 \tag{65}$$

if $b \in (b_0, \bar{b})$ and

$$\left.\frac{\partial}{\partial b^i} EU^\kappa\left(a_0, b, a_0, b^i\right)\right|_{b^i=b} = \frac{\partial}{\partial b}\Phi(a_0, b) \geq 0 \tag{66}$$

if $b = \bar{b}$. In both cases, the single-peakedness assumption implies that b is a maximum of $b \mapsto \Phi(a_0, b)$.

Now, assume for a contradiction that there exist two equilibria $b_1 \neq b_2$. Then, both are maxima of the auxiliary function, contradicting the single-peakedness assumption. □

Note that this proposition generalizes Proposition 10: indeed, if, say, $a_v > \rho b_v$ and $a_0 \geq b_0$, then $\alpha(a, b_0) \geq 0$ for all $a \in [a_0, \bar{a}]$ and therefore $h''(a, b_0) \geq 0$ by assumption. Recalling that the cost term is strictly convex, $a \mapsto \Phi(a, b_0)$ is strictly concave.

Similar to the partisan case (see Section 3.7), we can pin down sufficient conditions for single-peakedness for the specifc function we use in our illustrating examples.

Lemma 4. *Let $\kappa > 0$ and let $h(x) = \frac{\arctan(mx)}{\arctan(m)}$.*

1. *Let $\lambda = a_v - \rho b_v > 0$. If*
   - $\frac{f_A(c)c}{F_A(c)}$ *is decreasing,*
   - *for some $r < \frac{2a_v m}{\lambda(m^2-2m+1)}$,*
   $$\lim_{c\to 0} \frac{F_A(c)}{c^r} > 0, \tag{67}$$
   - *and $\bar{s}_A \geq \frac{\lambda b_0(m^2+1)}{2a_v^2 m \arctan(m)}$,*

   single-peakedness of $a \mapsto \Phi(a, b_0)$ holds.

2. *Let $\lambda = \rho b_v - a_v > 0$. If*
   - $\frac{f_B(c)c}{F_B(c)}$ *is decreasing,*
   - *for some $r < \frac{2b_v m}{\lambda(m^2-2m+1)}$,*
   $$\lim_{c\to 0} \frac{F_B(c)}{c^r} > 0, \tag{68}$$
   - *and $\bar{s}_B \geq \frac{\lambda a_0(m^2+1)}{2b_v^2 m \arctan(m)}$,*

   single-peakedness of $b \mapsto \Phi(a_0, b)$ holds.

Proof. We restrict ourselves to the case $\lambda = a_v - \rho b_v > 0$, as the proof for the other case goes analogously. By an abuse of notation, we write $\Phi(s_A, b_0)$ instead of $\Phi(a_0 + a_v F_A(s_A), b_0)$. Then, writing $M(s_A) = a_0 + a_v F_A(s_A) - b_0$ and $T(s_A) = a_0 + a_v F_A(s_A) + b_0$,

$$\begin{aligned}\frac{\partial}{\partial s_A}\Phi(a_0, s_B) &= \lambda\kappa h'(\alpha(s_A, b_0))\frac{2\kappa b_0 a_v f_A(s_A)}{T(s_A)^2} - a_v s_A f_A(s_A) \\ &= \underbrace{a_v \frac{f_A(s_A)}{T(s_A)^2 + m^2 M(s_A)^2}}_{>0} \lambda\kappa h'(\alpha(s_A, b_0))\left(\frac{2m\kappa\lambda b_0}{\arctan(m)} - \left(T(s_A)^2 + m^2 M(s_A)^2\right)s_A\right).\end{aligned} \tag{69}$$

Note that $a \mapsto \Phi(a, b_0)$ is single-peaked if and only if $s_A \mapsto \Phi(s_A, b_0)$ is single-peaked. Therefore, it is sufficient to show that $\phi(s_A) = \left(T(s_A)^2 + m^2 M(s_A)^2\right)$ is strictly increasing, and eventually greater than $\frac{2m\kappa\lambda b_0}{\arctan(m)}$. This follows from the conditions in the same way as in the proof of Proposition 5. □

4.5 Non-partisan ethics: the winner-take-all limit

Finally, we adopt the same approximation as in the partisan setting to analyze the winner-take-all limit.

Proposition 13. *[Non-partisan ethics] Only two situations can be sustained as limit equilibria in the ex ante, limit "winner-take-all" case. Either cost-sensitive voters incur no cost and the side with the largest base wins, or the side with the largest base pays no cost and the side with the smallest base pays to match the other base.*

Example 11. The two equilibrium types can co-exist. An illustration is provided in Figure 12.

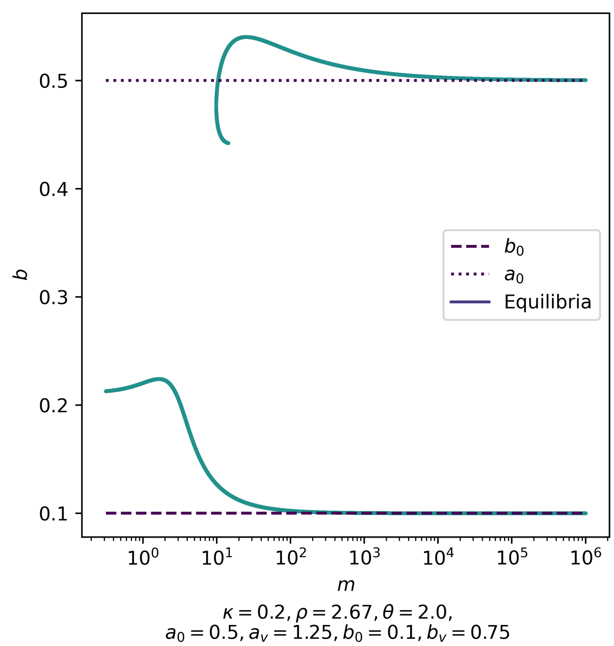


Figure 12: Equilibrium b for varying m in the case $\rho b_v > a_v$

5 Discussion and conclusions

In this paper we have tried to understand what follows if some people base their turnout decisions on an argument of the form “Voting is the right thing to do because there would be bad consequences if too many people abstained,” formalized through the *Homo moralis* preferences (Alger and Weibull, 2013), which capture well such partial universalization. The point is particularly relevant in circumstances where voting is costly and each single vote has a negligible effect on the relative number of votes obtained by the candidates, a feature that we capture by modeling the electorate as a continuum.

We find, first, that any extent of *Homo moralis* universalization ethics, i.e., any positive κ in the model, justifies participation in large electorates in most cases. It is generally true as long as voters perceive some benefit to any increase in the favorite candidate’s margin. This corresponds to the power sharing setting of our model, where both the winning and the losing side stand to gain from further increasing their share of the expressed votes. By contrast, *Homo moralis* universalization ethics do not generally justify participation in large electorates in the winner-take-all limit case of our model, where there is a second reason for why votes have a negligible impact: the marginal benefit from further increasing the share of expressed votes is nil except at the point where the two candidates tie precisely.

Second, our analysis reveals why it is important for a candidate to have a large base, that is, a large share of voters who always turn out to vote for them. The key effect of such a base is that it can motivate the cost-sensitive voters to vote. This occurs when a large base reduces the cost that a cost-sensitive voter needs to incur in order to realize that participation would have a large impact on the benefit. In these cases, the base is a complement to the turnout of cost-sensitive voters. We show that a large enough base for the underdog compared to that of the topdog can even trigger a large enough turnout among the cost-sensitive voters for the underdog to win the election. While this can happen even if the underdog’s supporters do not perceive a particularly high stake in the election (i.e., even if the stake parameter $\rho = 1$), these supporters are even more motivated to incur a cost to vote if the stake is not neutral ($\rho > 1$). That being said, if the base is too large, it becomes a substitute for the cost-sensitive voters’ turnout, since it reduces the marginal benefit of higher turnout rates.

A third pattern that emerges from our analysis is that high values of κ do not necessarily guarantee high turnout rates, because voters may face coordination problems. Indeed, similar to the base of one’s group, an increase in the share of other cost-sensitive voters who are expected to vote dampens the cost that an individual cost-sensitive voter needs to incur to reach a certain benefit. While this explains why there may exist equilibria with high turnout rates, it also explains why such equilibria can sometimes co-exist with equilibria with very low turnout rates.

Going forward, many questions remain. In particular, it would be interesting to allow for heterogeneity in the degrees of universalization. On the empirical side, while several motivations behind turnout decisions have already been documented and studied (Aytaç and Stokes, 2019; Blais, 2000; Blais and Daoust, 2020; Downs, 1957; Gerber et al., 2008; Hatemi and McDermott, 2011), it appears that no study has sought to detect universalization ethics as a driver of turnout decisions.

References

Alger, I., & Laslier, J.-F. (2022). *Homo moralis* goes to the voting booth: Coordination and information aggregation. *Journal of Theoretical Politics*, *34*(2), 280–312.

Alger, I., & Weibull, J. W. (2013). *Homo moralis*—preference evolution under incomplete information and assortative matching. *Econometrica*, *81*(6), 2269–2302.

Alger, I., & Weibull, J. W. (2017). Strategic behavior of moralists and altruists. *Games*, *8*(3).

Aytaç, S. E., & Stokes, S. C. (2019). *Why Bother? rethinking Participation in Elections and Protests*. Cambridge: Cambridge University Press.

Bierbrauer, F., Tsyvinski, A., & Werquin, N. (2022). Taxes and turnout: When the decisive voter stays at home. *American Economic Review*, *112*(2), 689–719.

Blais, A. (2000). *To Vote or Not to Vote? the Merits and Limits of Rational Choice Theory*. Pittsburgh: University of Pittsburgh Press.

Blais, A., & Achen, C. H. (2019). Civic duty and voter turnout. *Political Behavior*, *41*, 473–497.

Blais, A., & Daoust, J.-F. (2020). *The Motivation to Vote: Explaining Electoral Participation*. Vancouver: University of British Columbia Press.

Braham, M., & van Hees, M. (2020). *Kantian* Kantian optimization. *Erasmus Journal for Philosophy and Economics*, *13*, 30–42.

Cancela, J., & Geys, B. (2016). Explaining voter turnout: A meta-analysis of national and subnational elections. *Electoral studies*, *42*, 264–275.

Coate, S., & Conlin, M. (2004). A group rule—utilitarian approach to voter turnout: Theory and evidence. *American Economic Review*, *94*(5), 1476–1504.

Conley, J. P., Toossi, A., & Wooders, M. (2006). Memetics and voting: How nature may make us public spirited. *International Journal of Game Theory*, *35*, 71–90.

DeNardo, J. (1980). Turnout and the vote: The joke's on the democrats. *American Political Science Review*, *74*(2), 406–420.

Dhillon, A., & Peralta, S. (2002). Economic theories of voter turnout. *The Economic Journal*, *112*(480), F332–F352.

Downs, A. (1957). *An Economic Theory of Democracy*. New York: Harper and Row.

Feddersen, T., Gailmard, S., & Sandroni, A. (2009). Moral bias in large elections: Theory and experimental evidence. *American Political Science Review*, *103*(2), 175–192.

Feddersen, T., & Sandroni, A. (2006). A theory of participation in elections. *American Economic Review*, *96*(4), 1271–1282.

Fiorina, M. P. (1976). The voting decision: Instrumental and expressive aspects. *The Journal of Politics*, *38*(2), 390–413.

Frank, R. W., & Martínez i Coma, F. (2023). Correlates of voter turnout. *Political behavior*, *45*(2), 607–633.

Gerber, A. S., Green, D. P., & Larimer, C. W. (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment. *American political Science review*, *102*(1), 33–48.

Gravel, N., Laslier, J.-F., & Trannoy, A. (2000). Consistency between tastes and values: A universalisation approach. *Social Choice and Welfare*, *17*, 129–137.

Harsanyi, J. C. (1977). Morality and the theory of rational behavior. *Social Research: An International Quarterly*, *44*(4), 623–656.

Harsanyi, J. C. (1980). Rule utilitarianism, rights, obligations and the theory of rational behavior. *Theory and decision*, *12*(2), 115–133.

Hatemi, P. K., & McDermott, R. (2011). *Man is by Nature a Political Animal: Evolution, Biology, and Politics.* Chicago: University of Chicago Press.

Herrera, H., Morelli, M., & Nunnari, S. (2016). Turnout across democracies. *American Journal of Political Science*, *60*(3), 607–624.

Kant, I. (1785). *Grundlegung zur Metaphysik der Sitten. [in English: Groundwork of the Metaphysics of Morals (1964).* New York: Harper Torch Books].

Krishna, V., & Morgan, J. (2015). Majority rule and utilitarian welfare. *American Economic Journal: Microeconomics*, *7*(4), 339–375.

Laffont, J.-J. (1975). Macroeconomic constraints, economic efficiency and ethics: An introduction to Kantian economics. *Economica*, *42*(168), 430–437.

Laslier, J.-F. (2023). Universalization and altruism. *Social Choice and Welfare*, *60*, 579–594.

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

Levine, D. K., & Mattozzi, A. (2020). Voter turnout with peer punishment. *American Economic Review*, *110*(10), 3298–3314.

Levine, S., Kleiman-Weber, M., Schultz, L., & Cushman, F. (2020). The logic of universalization guides moral judgment. *PNAS*, *117*, 26158–26169.

Lizzeri, A., & Persico, N. (2001). The provision of public goods under alternative electoral incentives. *American Economic Review*, *91*(1), 225–239.

Morton, R. B. (1987). A group majority voting model of public good provision. *Social Choice and Welfare*, *4*, 117–131.

Morton, R. B. (1991). Groups in rational turnout models. *American Journal of Political Science*, 758–776.

Myerson, R. B. (2000). Large Poisson games. *Journal of Economic Theory*, *94*(1), 7–45.

Nagel, J. H., & McNulty, J. E. (1996). Partisan effects off voter turnout in senatorial and gubernatorial elections. *American Political Science Review*, *90*(4), 780–793.

Palfrey, T. R., & Rosenthal, H. (1985). Voter participation and strategic uncertainty. *American political science review*, *79*(1), 62–78.

Riker, W. H., & Ordeshook, P. C. (1968). A theory of the calculus of voting. *American political science review*, *62*(1), 25–42.

Roemer, J. E. (2019). *How We Cooperate*. New Haven: Yale University Press.

Rogers, T., Green, D. P., Ternovski, J., & Young, C. F. (2017). Social pressure and voting: A field experiment conducted in a high-salience election. *Electoral Studies*, *46*, 87–100.

Schuessler, A. A. (2000). Expressive voting. *Rationality and Society*, *12*(1), 87–119.

Smets, K., & Van Ham, C. (2013). The embarrassment of riches? A meta-analysis of individual-level research on voter turnout. *Electoral studies*, *32*(2), 344–359.

Van Leeuwen, B., & Alger, I. (forthcoming). Estimating social preferences and Kantian morality in strategic interactions. *Journal of Political Economy Microeconomics*.

A Details on functional forms for political outcome

Using our notation, Herrera et al., 2016 posit the material benefit function $a^\gamma/(a^\gamma+b^\gamma)$ for the A-supporters and $b^\gamma/(a^\gamma+b^\gamma)$ for the B-supporters, where the parameter $\gamma\in[1,+\infty)$ captures the power sharing rule. By setting

$$h(\alpha)=\frac{(1+\alpha)^\gamma-(1-\alpha)^\gamma}{(1+\alpha)^\gamma+(1-\alpha)^\gamma}, \tag{70}$$

it is straightforward to show that

$$h(\alpha)=-1+2\frac{a^\gamma}{a^\gamma+b^\gamma}, \tag{71}$$

and similarly for β. Indeed,

$$\begin{aligned}\frac{(1+\alpha)^\gamma-(1-\alpha)^\gamma}{(1+\alpha)^\gamma+(1-\alpha)^\gamma}&=\frac{\frac{2a}{a+b}^\gamma-\frac{2b}{a+b}^\gamma}{\frac{2a}{a+b}^\gamma+\frac{2b}{a+b}^\gamma}\\&=-\frac{(2a)^\gamma+(2b)^\gamma}{(2a)^\gamma+(2b)^\gamma}+2\frac{(2a)^\gamma}{(2a)^\gamma+(2b)^\gamma}\\&=-1+2\frac{a^\gamma}{a^\gamma+b^\gamma}.\end{aligned} \tag{72}$$

It remains to show that the function in (70) fits the assumptions of our model. Clearly, it takes values between -1 and 1 and is symmetric around $\alpha=0$. Moreover, it is continuous and differentiable. Indeed, the first derivative is given by

$$h'(\alpha)=\frac{4\gamma\left((1-\alpha^2)^{\gamma-1}\right)}{((1+\alpha)^\gamma+(1-\alpha)^\gamma)^2}>0, \tag{73}$$

and therefore, the second derivative is given by

$$\begin{aligned}h''(\alpha)&=4\gamma\frac{-2\alpha(\gamma-1)(1-\alpha^2)^{\gamma-2}((1+\alpha)^\gamma+(1-\alpha)^\gamma)^2}{((1+\alpha)^\gamma+(1-\alpha)^\gamma)^4}\\&-4\gamma\frac{2\left((1-\alpha^2)^{\gamma-1}\right)((1+\alpha)^\gamma+(1-\alpha)^\gamma)\gamma\left((1+\alpha)^{\gamma-1}-(1-\alpha)^{\gamma-1}\right)}{((1+\alpha)^\gamma+(1-\alpha)^\gamma)^4}\\&=-8\gamma(\gamma-1)\alpha\frac{(1-\alpha^2)^{\gamma-2}}{((1+\alpha)^\gamma+(1-\alpha)^\gamma)^2}\\&-8\gamma^2\frac{(1-\alpha^2)^{\gamma-1}\left((1+\alpha)^{\gamma-1}-(1-\alpha)^{\gamma-1}\right)}{((1+\alpha)^\gamma+(1-\alpha)^\gamma)^3}\end{aligned} \tag{74}$$

It is now straightforward to check that this h satisfies our assumptions on the derivatives for any $\gamma \in [1, +\infty)$.

Re-scaling the cost accordingly, this shows that our benefit term is more general.

B Computing equilibria with the arctan benefit function and uniform cost

Recall the arctan benefit function, $\frac{1}{\arctan(m)} \arctan(mx)$, that we used throughout our examples. It turns out that equilibria can be computed efficiently if the cost follows a uniform distribution.

Partisan voters

Here, we assume that the cost is distributed according to the uniform distribution on $[0, \theta_A]$ for A-supporters and $[0, \theta_B]$ for B-supporters. We first describe how to compute A-consistent strategies. The computation of these is almost the same as the computation of equilibria for nonpartisan voters.

Computing A- or B-consistent strategies

First, we explain how we may compute A-consistent strategies, where B-consistent strategies can be computed analogously. We look for an A-consistent strategy a given a strategy b played by B-supporters. The first-order condition for a, $0 = \frac{\partial}{\partial a^i} EU_A^\kappa (a, b, a^i)\big|_{a^i=a}$ is given by

$$0 = \frac{m}{\arctan(m)} \frac{1}{1 + \left(m\frac{a-b}{b+a}\right)^2} \frac{2\kappa b}{(b+a)^2} - \theta_A \frac{a - a_0}{a_v^2}. \tag{75}$$

After some algebra, we obtain a polynomial:

$$0 = a^3 + \left(2\frac{1-m^2}{1+m^2}b - a_0\right) a^2 + \left(b^2 - 2\frac{1-m^2}{1+m^2}a_0 b\right) a - b^2 a_0 - \frac{2m\kappa b a_v^2}{(1+m^2)\,\theta_A \arctan(m)} \tag{76}$$

Solving the polynomial gives candidate A-consistent strategies, to which we have to add $\bar{a}$ (but a_0 cannot be A-consistent).

Then, in order to rule out a profitable deviation, it is sufficient to compare the utility level associated to such a candidate a with the utility at any solution of $\frac{\partial}{\partial a^i} U^\kappa \left(a, b^*, a^i\right) = 0$ and with the utility at $\bar{a}$. The equation $\frac{\partial}{\partial a^i} U^\kappa \left(a, b^*, a^i\right) = 0$ can be rewritten as another degree three polynomial equation.

Computing equilibria

In order to find equilibria, we first find pairs (a, b) that simultaneously solve the first order conditions, i.e. $0 = \frac{\partial}{\partial a^i} EU_A^\kappa \left(a, b, a^i\right)\Big|_{a^i=a}$ and $0 = \frac{\partial}{\partial b^i} EU_B^\kappa \left(a, b, a^i\right)\Big|_{b^i=b}$. As above, these can both be rewritten as polynomials in (a, b). We can solve the system of polynomials numerically using the resultant method, where we use SymPy to compute the resultant and NumPy to compute roots.

Other candidate equilibria are $\left(\bar{a}, \bar{b}\right)$, $\left(a, \bar{b}\right)$ with a a candidate A-consistent strategy given $\bar{b}$ and $(\bar{a}, b)$ with b a candidate B-consistent strategy given $\bar{a}$. For each candidate equilibrium, we check for profitable deviations as described above.

Non-partisan voters

Here, we assume that the cost is distributed according to the uniform distribution on $[0, \theta]$. We may also assume, without loss of generality, that $\rho b_v - a_v > 0$: indeed, we have established that in this case, $a = a_0$. Finding an equilibrium therefore amounts to finding $b > b_0$. In case the opposite inequality holds, one needs to find the equilibrium a for $b = b_0$.

Let us write out the first-order condition $0 = \frac{\partial}{\partial b^i} EU^\kappa \left(a, b, a^i, b^i\right)|_{b^i=b, a^i=a=a_0}$:

$$0 = \frac{m\left(\rho b_v - a_v\right)}{\arctan\left(m\right)} \frac{1}{1 + \left(m \frac{b-a_0}{b+a_0}\right)^2} \frac{2\kappa b}{\left(b + a_0\right)^2} - \theta \frac{b - b_0}{b_v}. \tag{77}$$

It is then straightforward to rewrite this equation as a polynomial equation in b. Finding roots of the polynomial yields candidate equilibria.

Finally, for some candidate equilibrium b, one can write out $\frac{\partial}{\partial b^i} EU^\kappa \left(a, b, a^i, b^i\right)|_{a^i=a=a_0}$, observe that it can be rewritten as a polynomial in b^i (we use SymPy) and check if any solution or $\bar{b}$ is associated with a higher expected utility, in order to rule out profitable deviations.

C Proofs of the winner-take-all limit results

C.1 Approximating sequences

Definition 1. Let $(h_t)_{t=1,2...}$ be a sequence of functions that all satisfy the hypothesis of the model and such that, for any $x \in [-1,1]$,

$$\lim_{t\to\infty} h_t(x) = \text{sign(x)}.$$

Such a sequence $(h_t)_{t=1,2,...}$ will be called an *approximating sequence.*

An example for such an approximating sequence is $\left(\frac{1}{\arctan(m)} \arctan(mx)\right)_{m=1,2,...}$.

To prepare the proofs, let us first establish that for any $\varepsilon > 0$, h converges uniformly to 1 on $[\varepsilon, 1]$ (and by symmetry, it converges uniformly to -1 on $[-1, -\varepsilon]$). To see this, notice that h_t is increasing. Therefore, for any $x \in [\varepsilon, 1]$, $h_t(\varepsilon) \leq h_t(x) \leq 1$. Thus, $\sup_{x\in[\varepsilon,1]} |h_t(x) - h(x)| \leq 1 - h_t(\varepsilon)$ and the result follows by the pointwise convergence of h_t to one at ε.

The concavity of h_t on positive numbers implies that for any $x > 0$,

$$\lim_{t\to\infty} h_t'(x) = 0.$$

To see this, note that on intervals $[y, x]$ and $[x, z]$, concavity of h_t implies $\frac{h_t(z)-h_t(x)}{z-x} < h_t'(x) < \frac{h_t(x)-h_t(y)}{x-y}$ and apply the sandwich lemma. The same result is likewise obtained for $x < 0$.

Since h_t' is decreasing for $x > 0$, h_t' converges uniformly to 0 on $[\varepsilon, 1]$, and likewise on $[-1, -\varepsilon]$.

Definition 2. We say that a pair (a, b) is *sustained as a limit winner-take-all equilibrium* under partisan (resp. non partisan) ethics if there exists a sequence $(a_t, b_t, h_t)_{t\in\mathbb{N}}$ such that

- $(h_t)_{t=1,2...}$ is an approximating sequence,
- for all t, (a_t, b_t) is an equilibrium of the partisan (resp. non partisan) game when the political outcome function is h_t, and
- $\lim_{t\to\infty} (a_t, b_t) = (a, b)$.

C.2 Partisan ethics

Proof of Proposition 6: Let (a, b) be sustained as a limit winner-take-all equilibrium under partisan ethics by a sequence (a_t, b_t, h_t). Because a_t and b_t are larger than $a_0 > 0$ and $b_0 > 0$, continuity of the function α implies that, if $a \neq b$, then $\alpha(a_t, b_t)$ converges to $\alpha(a, b) \neq 0$, hence $h'_t(a_t, b_t)$ tends to 0, using the uniform convergence established above. It follows from the equilibrium conditions (Equations 33 and 34) that $C'_A\left(\frac{a_t - a_0}{a_v}\right)$ tends to 0. By continuity of the function C'_A this implies that a_t tends to a_0. The same argument holds for B.

Let us now assume that we have $a = b$ with (a, b) such that $a_0 \leq a < a_0 + a_v$ and $b_0 \leq b < b_0 + b_v$.

Notice that there must be infinitely many (a_t, b_t) such that $b_t \leq a_t$ or infinitely many (a_t, b_t) such that $a_t \leq b_t$ (both can be the case). We will assume that there are infinitely many $b_t \leq a_t$ since the argument that will follow works analogously with infinitely many $a_t \leq b_t$. By extracting a subsequence, we will assume that the whole sequence is such that $b_t \leq a_t$.

Let $\delta > 0$ and consider a deviation $\hat{b}_t$ such that

$$\kappa \hat{b}_t + (1 - \kappa)\, b_t = a + \delta, \tag{78}$$

i.e.

$$\hat{b}_t = \frac{a + \delta - (1 - \kappa)\, b_t}{\kappa} \to a + \frac{\delta}{\kappa}. \tag{79}$$

Since $b_t \to a$ and since we assumed that $a = b < b_0 + b_v$, as long as δ is small enough, there exists a $T \in \mathbb{N}$ large enough such that $\hat{b}_t \in [b_0, b_0 + b_v]$, i.e. such that $\hat{b}_t$ is a feasible deviation for all $t > T$.

Then, recalling $(a_t, b_t) \to (a, a)$, and by possibly increasing T, we can ensure that $\beta\,(a_t, a + \delta) \in [\varepsilon, 1]$ for all $t > T$, for some $\varepsilon > 0$ small enough.

We will now show that $\hat{b}_t$ is a profitable deviation for $t > T$. Indeed, by the uniform convergence of h_t towards 1 on $[\varepsilon, 1]$ and the continuity of C_B,

$$U^{\kappa}_{B,t}\left(a_t, b_t, \hat{b}_t\right) - U^{\kappa}_{B,t}\,(a_t, b_t, b_t) = \rho \underbrace{h_t\,(\beta\,(a_t, a + \varepsilon))}_{\to 1} - \rho \underbrace{h_t\,(\beta\,(a_t, b_t))}_{<0} + \underbrace{C_B\,(b_t) - C_B\left(\hat{b}_t\right)}_{\to C_B(a) - C_B\left(a + \frac{\delta}{\kappa}\right)} \tag{80}$$

is strictly positive for t large enough as well as δ small enough. The continuity of C_B is used twice in the argument: first to establish the convergence of the cost terms, and second to argue that the cost difference is arbitrarily small for small enough δ.

Having shown that there exist profitable deviations in the approximating sequence of equilibria for the individuals supporting at least one of the parties, we have reached a contradiction, i.e. there cannot be a limit winner-take-all equilibrium (a, a) with $a_0 \leq a < a_0 + a_v$ and $b_0 \leq b < b_0 + b_v$.

Let us now assume that $\left(\bar{b}, \bar{b}\right)$ is sustained as a limit winner-take-all equilibrium and that $\bar{a} > \bar{b}$ (where by assumption, $\bar{a} \geq \bar{b}$, but the proof does work analogously if one were to allow $\bar{b} > \bar{a}$ in the model). Let us assume that $h_t\left(\beta\left(a_t, b_t\right)\right)$ does not converge to 1. Then, there exists a ρ such that there exists a subsequence of (a_t, b_t) which satisfies $h_t\left(\beta\left(a_t, b_t\right)\right) < 1 - \rho$. Similarly to the sequence of eventually profitable deviations we constructed above for B, we can then construct a sequence of eventually profitable deviations for A so that we can conclude that $h_t\left(\beta\left(a_t, b_t\right)\right) \to 1$. This, in turn, implies that for t large enough, it is profitable for B-supporters to deviate downwards to zero effort since the benefit of keeping a turnout of $\bar{b}$ vanishes as $t \to \infty$.

Therefore, except in the case $\bar{a} = \bar{b}$, there cannot be a limit winner-take-all equilibrium $\left(\bar{b}, \bar{b}\right)$.

□

Proof of Proposition 7 . Suppose that another pair $(a, b) \neq (a_0, b_0)$ is justified. Following the previous proposition, let $l = a = b$ be the common limit participation.

Recall that, by assumption, $\bar{b} < \bar{a}$; if $l < \bar{b}$ then Equation (35) applies for interior equilibria and writes at the limit:

$$\rho = r \cdot \frac{F^{-1}\left(\frac{l - a_0}{a_v}\right)}{F^{-1}\left(\frac{l/r - a_0}{a_v}\right)}. \tag{81}$$

Since F^{-1} is non-decreasing, and $r > 1$ the ratio in the above equation is larger than 1, hence $r \leq \rho$. Hence the result.

If the limit is $l = \bar{b}$ then (33) holds with an equality and (34) with an inequality ($\geq$) so that the above equality (81) must be replaced by the inequality $\rho \geq r \cdot \frac{F^{-1}\left(\frac{l - a_0}{a_v}\right)}{F^{-1}\left(\frac{l/r - a_0}{a_v}\right)}$, leading to the same conclusion. □

C.3 Non-partisan ethics

Proof of Proposition 13: The other parameters being fixed, let (a_t, b_t) be a sequence of scores sustained at equilibrium for h_t. In view of Proposition 9 we concentrate on the case where $\kappa > 0$ and $a_v \neq \rho b_v$.

Case 1. Suppose first that $a_v > \rho b_v$. In view of Proposition 2, for all t, $a_t > a_0$ and $b_t = b_0$, and a_t satisfies the first-order condition (see equation 53):

$$0 \leq (a_v - \rho b_v) h_t'\Big(\frac{a_t - b_0}{a_t + b_0}\Big) \frac{2\kappa b_0}{a_t + b_0} - a_v C_A'(a_t), \tag{82}$$

where $<$ may only hold for $a_t = \bar{a}$.

1. If $a_0 \geq b_0$. Suppose that for some $\varepsilon > 0$ the sequence (a_t) has an infinite number of points $a_t > a_0 + \varepsilon$, then one can extract a sequence a_τ that converges to some $\tilde{a} > a_0$, then $\frac{a_\tau - b_0}{a_\tau + b_0}$ tends to $\frac{\tilde{a} - b_0}{\tilde{a} + b_0} > 0$, hence $h_t'\Big(\frac{a_\tau - b_0}{a_\tau + b_0}\Big)$ tends to 0. But $C_A'(a_\tau)$ tends to $C_A'(\tilde{a}) > 0$, a contradiction. We conclude that, in this case,
$$\lim_{t \to \infty} a_t = a_0.$$

2. If $a_0 < b_0$. Reasoning as previously rules out accumulation points $\tilde{a}$ that would be strictly larger than b_0 or in the open interval (a_0, b_0), leaving only the two possibilities $\tilde{a} = a_0$ or $\tilde{a} = b_0$.

   (a) Suppose first that $\tilde{a} = a_0$. Consider the response $(s_A^i, 0)$ of player i such that, taking the Homo Moralis effect into account, i will exactly make the score of A match the score (b_0) of B. That is s_a^i is such that $a_0 + \kappa a_v F_A(s_A^i) = b_0$ or, in other terms:
   $$s_A^i = F_A^{-1}\left(\frac{b_0 - a_0}{\kappa a_v}\right). \tag{83}$$
   Such cost threshold exists if and only if
   $$b_0 - a_0 \leq \kappa a_v. \tag{84}$$

Suppose, for the moment, that this condition is met. The expected payoffs are:

$$EU^{\kappa}(a_t, b_0, a_t, b_0) = (a_v - \rho b_v) \cdot h_t\left(\frac{a_t - b_0}{a_t + b_0}\right) - a_v C_A(a_t) \tag{85}$$

$$EU^{\kappa}(a_t, b_0, a_0, b_0) = (a_v - \rho b_v) \cdot h_t(0) - a_v \int_{c=0}^{s_A^i} cf(c)dc \tag{86}$$

When t tends to infinity, a_t tends to a_0 hence $C_A(a_t)$ tends to 0 and $h_t\left(\frac{a_t - b_0}{a_t + b_0}\right)$ tends to -1 so that the equilibrium payoff tends to $-(a_v - \rho b_v)$. Since $h_t(0) = 0$, the deviation payoff is always equal to $-a_v \int_{c=0}^{s_A^i} cf(c)dc$, so the deviation not being profitable implies

$$a_v - \rho b_v \leq a_v \int_{c=0}^{s_A^i} cf(c)dc. \tag{87}$$

In equation 83 we remark that s_A^i does not depend on ρ, so that the above equation is better written as:

$$\rho \geq \rho^* = \frac{a_v}{b_v}\left(1 - \int_{c=0}^{s_A^i} cf(c)dc\right). \tag{88}$$

We conclude that if ρ is small, there is a profitable deviation, the possibility $\tilde{a} = a_0$ is ruled out and the only limit equilibrium is $\tilde{a} = b_0$.

If the condition 84 is not satisfied, which occurs if κ is small, then any deviation in s_A^i will induce a score a smaller than and bounded away from b_0, so the limit electoral payoff will be -1, the cost strictly positive, and the outcome $\tilde{a} = a_0$ will not be de-stabilized.

(b) Suppose that $\tilde{a} = b_0$. Consider the response $(s_A^i, s_B^i) = (0, 0)$ of player i to the equilibrium that yields the score a_t for h_t. The score of A, as perceived by i is no longer a_t but $a'_t = (1 - \kappa)a_t + \kappa a_0$. The payoffs at equilibrium and under this response are respectively:

$$EU^{\kappa}(a_t, b_0, a_t, b_0) = (a_v - \rho b_v) \cdot h_t\left(\frac{a_t - b_0}{a_t + b_0}\right) - a_v C_A(a_t) \tag{89}$$

$$EU^{\kappa}(a_t, b_0, a_0, b_0) = (a_v - \rho b_v) \cdot h_t\left(\frac{a'_t - b_0}{a'_t + b_0}\right) - 0 \tag{90}$$

When t tends to ∞, a_t tends to $\bar{a} = b_0$, hence a'_t tends to $(1 - \kappa)b_0 + \kappa a_0$ and, because this number is strictly lower than b_0, the electoral payoff $h_t\left(\frac{a'_t - b_0}{a'_t + b_0}\right)$ tends to -1. Hence the payoff for deviating tends to $-(a_v - \rho b_v)$. Because the deviation is not profitable, it must be that:

$$-(a_v - \rho b_v) \leq \liminf_{t\to\infty} \left[(a_v - \rho b_v) \cdot h_t\left(\frac{a_t - b_0}{a_t + b_0}\right) - a_v C_A(a_t)\right] \tag{91}$$

The term $a_v C_A(a_t)$ tends to $a_v C_A(b_0)$, a strictly positive number. Letting $H^* = \liminf_{t\to\infty} h_t\left(\frac{a_t - b_0}{a_t + b_0}\right)$, this number is such that $-1 \leq H^* \leq 1$ and the above equation writes.

$$a_v - \rho b_v \geq \frac{a_v}{1 + H^*} C_A(b_0). \tag{92}$$

If $H^* = -1$ this inequality cannot be satisfied. So $1 + H^* > 0$, and (92) implies that $a_v - \rho b_v$ cannot be too small. However this equation should be interpreted with caution because, unlike equation 87, equation 92 cannot be written in general as a bound on ρ, as could be done in 88. Still we can conclude that if 92 is not satisfied, $\tilde{a} = b_0$ is ruled out and the only limit equilibrium is $\tilde{a} = a_0$.

Comparing the conditions (87) and (92) one can see that they do not match exactly, so that we cannot exclude the possibility that two different equilibria co-exist in the limit: one in which A-voters pay to match the B hardliners and one in which they pay nothing.

Case 2. Next, suppose that $a_v < \rho b_v$. This case is symmetric. In view of Proposition 2, for all t, $a_t = a_0$ and $b_t > b_0$, and b_t satisfies the first-order condition:

$$0 = (\rho b_v - a_v) h_t'\left(\frac{b_t - a_0}{b_t + a_0}\right) \frac{2\kappa a_0}{b_t + a_0} - b_v C_B'(b_t). \tag{93}$$

With the same arguments as previously, one obtains:

- If $b_0 \geq a_0$.

$$\lim_{t\to\infty} b_t = b_0.$$

- If $b_0 < a_0$. The sequence b_t can only have a_0 and b_0 as accumulation points. As previously, conditions may be obtained that imply that one limit or the other is only possible.

□

D Other proofs

D.1 Proof of Lemma 3

Proof. Claim: for any b, there exists a unique A-consistent strategy. To prove this, let us define $\phi : [a_0, \bar{a}] \to [a_0, \bar{a}]$ by

$$\phi(a) = \arg\max_{a^i} EU_A^\kappa\left(a, b, a^i\right). \tag{94}$$

This is indeed well-defined, since the argmax exists and is unique due to the single-peakedness assumption on EU_A^κ. By Berge's maximum theorem, ϕ is continuous. Therefore, Brouwer's fixed point theorem applies and there exists at least one fixed point, i.e. at least one A-consistent strategy. By the assumption that for all a, b, the unique maximum of $a^i \mapsto EU_A^\kappa(a, b, a^i)$ lies in $(a_0, \bar{a})$, we can conclude that any such A-consistent strategy lies in $(a_0, \bar{a})$ (is interior). Hence, it satisfies the first-order condition

$$0 = \frac{\partial}{\partial a^i} EU_A^\kappa\left(a, b, a^i\right)\bigg|_{a^i = a}. \tag{95}$$

Since

$$\frac{\partial}{\partial a}\Phi_A(a, b) = \frac{2\kappa b h'(\alpha(a, b))}{(a + b)^2} - C_A'(a), \tag{96}$$

and recalling (33), we conclude that any fixed point of (94) satisfies

$$0 = \frac{\partial}{\partial a^i} EU_A^\kappa\left(a, b, a^i\right)\bigg|_{a^i = a} = \frac{\partial}{\partial a}\Phi_A(a, b). \tag{97}$$

Since the single-peakedness assumption on Φ_A implies that there is a unique a that maximizes $\Phi_A(a, b)$, this completes the proof of the claim.

By repeating the same argument for B-consistent strategies, we conclude that (a^*, b^*) is an equilibrium of the population game if and only if

1. a^* maximizes $a \mapsto \Phi_A(a, b^*)$, and
2. b^* maximizes $b \mapsto \Phi_B(a^*, b)$.

□

D.2 Proof of Proposition 5

Proof. Proving first the first claim, let us define a function $\psi : [a_0, \bar{a}] \times [b_0, \bar{b}] \to [a_0, \bar{a}] \times [b_0, \bar{b}]$ by

$$\psi(a, b) = \begin{pmatrix} \arg\max_{\tilde{a}} \Phi_A(\tilde{a}, b) \\ \arg\max_{\tilde{b}} \Phi_B\left(a, \tilde{b}\right) \end{pmatrix}, \tag{98}$$

which is well-defined by the assumptions on Φ_A and Φ_B. Applying Berge's maximum theorem, we conclude that ψ is continuous, allowing us to apply Brouwer's theorem. We deduce that ψ has at least one fixed point, proving equilibrium existence.

In order to prove uniqueness, we again rely on the auxiliary two-player game. Assume, for a contradiction, that (a_1, b_1) and (a_2, b_2) are both equilibria of that game. Then,

$$\Phi_A(a_1, b_1) \geq \Phi_A(a_2, b_1), \tag{99}$$
$$\Phi_A(a_2, b_2) \geq \Phi_A(a_1, b_2), \tag{100}$$
$$\Phi_B(a_1, b_1) \geq \Phi_B(a_1, b_2), \tag{101}$$
$$\Phi_B(a_2, b_2) \geq \Phi_B(a_2, b_1). \tag{102}$$

Writing these expressions out,

$$\kappa h(\alpha(a_1, b_1)) - C_A(a_1) \geq \kappa h(\alpha(a_2, b_1)) - C_A(a_2), \tag{103}$$
$$\kappa h(\alpha(a_2, b_2)) - C_A(a_2) \geq \kappa h(\alpha(a_1, b_2)) - C_A(a_1), \tag{104}$$
$$-\rho\kappa h(\alpha(a_1, b_1)) - C_B(b_1) \geq -\rho\kappa h(\alpha(a_1, b_2)) - C_B(b_2), \tag{105}$$
$$-\rho\kappa h(\alpha(a_2, b_2)) - C_B(b_2) \geq -\rho\kappa h(\alpha(a_2, b_1)) - C_B(b_1). \tag{106}$$

Rewriting,

$$\kappa h(\alpha(a_1, b_1)) - \kappa h(\alpha(a_2, b_1)) \geq C_A(a_1) - C_A(a_2), \tag{107}$$
$$C_A(a_1) - C_A(a_2) \geq \kappa h(\alpha(a_1, b_2)) - \kappa h(\alpha(a_2, b_2)), \tag{108}$$
$$\rho\kappa h(\alpha(a_1, b_2)) - \rho\kappa h(\alpha(a_1, b_1)) \geq C_B(b_1) - C_B(b_2), \tag{109}$$
$$C_B(b_1) - C_B(b_2) \geq \rho\kappa h(\alpha(a_2, b_2)) - \rho\kappa h(\alpha(a_2, b_1)). \tag{110}$$

Combining and eliminating constant positive factors,

$$h(\alpha(a_1, b_1)) - h(\alpha(a_2, b_1)) \geq h(\alpha(a_1, b_2)) - h(\alpha(a_2, b_2)), \tag{111}$$
$$h(\alpha(a_1, b_2)) - h(\alpha(a_1, b_1)) \geq h(\alpha(a_2, b_2)) - h(\alpha(a_2, b_1)). \tag{112}$$

Rewriting once more,

$$h\left(\alpha\left(a_{1},b_{1}\right)\right)+h\left(\alpha\left(a_{2},b_{2}\right)\right) \geq h\left(\alpha\left(a_{1},b_{2}\right)\right)+h\left(\alpha\left(a_{2},b_{1}\right)\right), \tag{113}$$

$$h\left(\alpha\left(a_{1},b_{2}\right)\right)+h\left(\alpha\left(a_{2},b_{1}\right)\right) \geq h\left(\alpha\left(a_{2},b_{2}\right)\right)+h\left(\alpha\left(a_{1},b_{1}\right)\right). \tag{114}$$

Combining the two inequalities, we have equality throughout:

$$h\left(\alpha\left(a_{1},b_{1}\right)\right)+h\left(\alpha\left(a_{2},b_{2}\right)\right)=h\left(\alpha\left(a_{1},b_{2}\right)\right)+h\left(\alpha\left(a_{2},b_{1}\right)\right). \tag{115}$$

Multiplying with κ and subtracting $C_A\left(a_1\right)$ as well as $C_A\left(a_2\right)$ on both sides,

$$\Phi_A\left(a_1,b_1\right)+\Phi_A\left(a_2,b_2\right)=\Phi_A\left(a_1,b_2\right)+\Phi_A\left(a_2,b_1\right). \tag{116}$$

Since $\Phi_A\left(a_1,b_1\right)\geq\Phi_A\left(a_2,b_1\right)$ and $\Phi_A\left(a_2,b_2\right)\geq\Phi_A\left(a_1,b_2\right)$, we have

$$\Phi_A\left(a_1,b_1\right)=\Phi_A\left(a_2,b_1\right) \text{ and } \Phi_A\left(a_2,b_2\right)=\Phi_A\left(a_1,b_2\right). \tag{117}$$

Since we assumed that $a\mapsto\Phi_A\left(a,b\right)$ is single-peaked for any b, we deduce $a_1=a_2$. By repeating the same argument on Φ_B, we deduce $b_1=b_2$, thus proving equilibrium uniqueness of the auxiliary game. Lemma 3 then implies that this is also the unique equilibrium of the population game. This completes the proof of the first claim of the proposition.

We turn now to the second claim of the proposition. To begin, for any turnout levels (a,b) let us write the expected utility of B-supporter i as a function of the cutoff strategy s_B^i:

$$EU_B^\kappa\left(a,b,s_B^i\right)=\rho h\left(\beta^\kappa\left(a,b,s_B^i\right)\right)-\int_0^{s_B^i}cf\left(c\right)dc, \tag{118}$$

where

$$\beta^\kappa\left(a,b,s_B^i\right)=\frac{\left(1-\kappa\right)b+\kappa\left(b_vF_B\left(s_B^i\right)+b_0\right)-a}{\left(1-\kappa\right)b+\kappa\left(b_vF_B\left(s_B^i\right)+b_0\right)+a}. \tag{119}$$

Likewise, write the associated auxiliary function as a function of s_B:

$$\Phi_B\left(a,s_B\right)=\rho\kappa h\left(\beta\left(a,s_B\right)\right)-\int_0^{s_B}cf\left(c\right)dc. \tag{120}$$

Clearly, since F_B is strictly increasing, single-peakedness of $EU_B^\kappa\left(a,b,s_B^i\right)$ in s_B^i holds if and only if single-peakedness of $EU_B^\kappa\left(a,b,b^i\right)$ in b^i holds. Likewise, single-peakedness of $\Phi\left(a,s_B\right)$ in s_B holds if and only if single-peakedness of $\Phi_B\left(a,b\right)$ in b holds.

We will show that for all (a,b), $s_B^i \mapsto EU_B^\kappa\left(a,b,s_B^i\right)$ and $s_B \mapsto \Phi_B\left(a,s_B\right)$ are single-peaked, as the proof goes analogously for EU_A^κ and Φ_A (with $\rho=1$).

In order to ease notation, let

$$\begin{aligned} M^\kappa\left(s_B^i\right) &= (1-\kappa)\,b+\kappa\left(b_v F_B\left(s_B^i\right)+b_0\right)-a, \text{ and} \\ T^\kappa\left(s_B^i\right) &= (1-\kappa)\,b+\kappa\left(b_v F_B\left(s_B^i\right)+b_0\right)+a. \end{aligned} \tag{121}$$

We then have

$$\frac{\partial}{\partial s_B^i}\beta^\kappa\left(a,b,s_B^i\right)=\frac{2\kappa a b_v f_B\left(s_B^i\right)}{\left(T^\kappa\left(s_B^i\right)\right)^2} \tag{122}$$

so that

$$\frac{\partial}{\partial s_B^i}EU_B^\kappa\left(a,b,s_B^i\right)=\rho h'\left(\beta^\kappa\left(a,b,s_B^i\right)\right)\frac{2\kappa a b_v f_B\left(s_B^i\right)}{\left(T^\kappa\left(s_B^i\right)\right)^2}-s_B^i f_B\left(s_B^i\right). \tag{123}$$

Hence, for $h(x)=\frac{\arctan(mx)}{\arctan(m)}$,

$$\begin{aligned} \frac{\partial}{\partial s_B^i}EU_B^\kappa\left(a,b,s_B^i\right) &= \rho\frac{2m\kappa a b_v f_B\left(s_B^i\right)}{\arctan(m)}\frac{1}{\left(T^\kappa\left(s_B^i\right)\right)^2}\frac{1}{1+m^2\frac{\left(M^\kappa\left(s_B^i\right)\right)^2}{\left(T^\kappa\left(s_B^i\right)\right)^2}}-s_B^i f_B\left(s_B^i\right) \\ &= f_B\left(s_B^i\right)\left(\frac{2m\kappa\rho a b_v}{\arctan(m)}\frac{1}{\left(T^\kappa\left(s_B^i\right)\right)^2+m^2\left(M^\kappa\left(s_B^i\right)\right)^2}-s_B^i\right) \\ &= \underbrace{\frac{f_B\left(s_B^i\right)}{\left(T^\kappa\left(s_B^i\right)\right)^2+m^2\left(M^\kappa\left(s_B^i\right)\right)^2}}_{>0}\left(\frac{2m\kappa\rho a b_v}{\arctan(m)}-\left(\left(T^\kappa\left(s_B^i\right)\right)^2+m^2\left(M^\kappa\left(s_B^i\right)\right)^2\right)s_B^i\right). \end{aligned} \tag{124}$$

Consider now the auxiliary function in (120). Since

$$\frac{\partial}{\partial s_B}\beta\left(a,s_B\right)=\frac{2ab_v f_B\left(s_B\right)}{\left(T^1\left(s_B\right)\right)^2}, \tag{125}$$

for $h(x)=\frac{\arctan(mx)}{\arctan(m)}$ we obtain

$$\begin{aligned} \frac{\partial}{\partial s_B}\Phi_B\left(a,s_B\right) &= \rho\kappa h'\left(\beta\left(a,s_B\right)\right)\frac{2\kappa a b_v f_B\left(s_B\right)}{\left(T^1\left(s_B\right)\right)^2}-s_B f_B\left(s_B\right) \\ &= \underbrace{\frac{f_B\left(s_B\right)}{\left(T^1\left(s_B\right)\right)^2+m^2\left(M^1\left(s_B\right)\right)^2}}_{>0}\left(\frac{2m\kappa\rho a b_v}{\arctan(m)}-\left(\left(T^1\left(s_B\right)\right)^2+m^2\left(M^1\left(s_B\right)\right)^2\right)s_B\right). \end{aligned} \tag{126}$$

Therefore, to show the single-peakedness of $EU_B^\kappa(a, b, s_A^i)$ in s_B^i and of $\Phi_B(a, s_B)$ in s_B, it is sufficient that $\phi_B^{\tilde{\kappa}}(s_B) \overset{def}{=} \left(\left(T^{\tilde{\kappa}}\right)^2 + m^2\left(M^{\tilde{\kappa}}\right)^2\right) s_B$ is, for both $\tilde{\kappa} = \kappa$ and $\tilde{\kappa} = 1$,

(a) strictly increasing, and

(b) eventually greater than $\frac{2m\kappa\rho a b_v}{\arctan(m)}$.

To prove (a) it is sufficient to prove that prove that $\left(\phi_B^{\tilde{\kappa}}\right)'(s_B) > 0$, where

$$\left(\phi_B^{\tilde{\kappa}}\right)'(s_B) = \left(T^{\tilde{\kappa}}\right)^2 + m^2\left(M^{\tilde{\kappa}}\right)^2 + 2s_B\tilde{\kappa}b_v f_B(s_B)\left(T^{\tilde{\kappa}} + m^2 M^{\tilde{\kappa}}\right). \tag{127}$$

It is straightforward to see that $\left(\phi_B^{\tilde{\kappa}}\right)'(s_B) > 0$ holds for $m \leq 1$. For $m > 1$, writing $\tilde{b} = b - b_0$, and minimizing the expression over $a, \tilde{b}$ and b_0 using SymPy (the code is included below), we show that $\left(\phi_B^{\tilde{\kappa}}\right)'(s_B) > 0$ if

$$s_B < \frac{2F_B(s_B)\, m}{\rho f_B(s_B)(m-1)^2} \tag{128}$$

or, equivalently,

$$\frac{f_B(s_B)\, s_B}{F_B(s_B)} < \frac{2m}{\rho(m-1)^2}. \tag{129}$$

Since $f_B(c)c/F_B(c)$ is decreasing (by assumption 2), it is sufficient to have

$$\lim_{c\to 0} \frac{f_B(c)\, c}{F_B(c)} < \frac{2m}{\rho(m-1)^2}. \tag{130}$$

For this, in turn, it is sufficient to have, for some $r < \frac{2m}{\rho(m-1)^2}$,

$$\lim_{c\to 0} \frac{F_B(c)}{c^r} > 0. \tag{131}$$

Let us now turn to (b). It is sufficient to give a condition on $\bar{s}_B$ such that

$$\bar{s}_B\left(\left(T^{\tilde{\kappa}}\right)^2 + m^2\left(M^{\tilde{\kappa}}\right)^2\right) > \frac{2m\tilde{\kappa}\rho a b_v}{\arctan(m)}, \tag{132}$$

where M^κ and T^κ are evaluated at $s_B^i = \bar{s}_B$, for both $\tilde{\kappa} = \kappa$ and $\tilde{\kappa} = 1$.

Using SymPy, we minimize $\psi = \left(\left(T^{\tilde{\kappa}}\right)^2 + m^2\left(M^{\tilde{\kappa}}\right)^2\right)$ over a, $\tilde{b}$ and b_0; we also maximize the right-hand side by plugging in $a = \bar{a}$. We find that (132) holds if

$$\bar{s}_B \frac{4\tilde{\kappa}^2 b_v^2 m^2}{m^2+1} \geq \frac{2m\tilde{\kappa}\rho\bar{a}b_v}{\arctan m}, \tag{133}$$

or equivalently,

$$\bar{s}_B \geq \frac{\rho\bar{a}\left(m^2+1\right)}{2\tilde{\kappa}b_v m \arctan\left(m\right)}. \tag{134}$$

We also observe that this condition holds for both $\tilde{\kappa} \in \{\kappa, 1\}$ if

$$\bar{s}_B \geq \frac{\rho\bar{a}\left(m^2+1\right)}{2\kappa b_v m \arctan\left(m\right)}. \tag{135}$$

□

SymPy calculations

[1]:
```
from sympy import *
```

[2]:
```
m,rho, sbi,bars,bv, kap = symbols(r'm \rho s_B \bar{s}_B b_v␣
↪\tilde{\kappa}',positive=True)
b = Symbol(r"\tilde{b}")
a = Symbol("a")
abar = Symbol(r"\bar{a}",positive=True)
b0 = Symbol("b_0")
FB = Function("F_B")(sbi)
fB = Function("f_B")(sbi)
FBsymb = Symbol("F_B")
fBsymb = Symbol("f_B")
```

Analyzing ϕ_B

[9]:
```
Tkap = kap*bv*FB + (1-kap)*b + b0 + a
Mkap = kap*bv*FB + (1-kap) * b + b0 - a


phi = Tkap**2 + m**2 * Mkap**2 + 2*sbi*kap*rho *bv*fB *(Tkap + m**2*Mkap)
display(collect(expand(phi),[a**2,a]))
```

$-2\rho\tilde{\kappa}^2\tilde{b}b_v m^2 s_B^i f_B\left(s_B^i\right) - 2\rho\tilde{\kappa}^2\tilde{b}b_v s_B^i f_B\left(s_B^i\right) + 2\rho\tilde{\kappa}^2 b_v^2 m^2 s_B^i F_B\left(s_B^i\right) f_B\left(s_B^i\right) + 2\rho\tilde{\kappa}^2 b_v^2 s_B^i F_B\left(s_B^i\right) f_B\left(s_B^i\right) + 2\rho\tilde{\kappa}\tilde{b}b_v m^2 s_B^i f_B\left(s_B^i\right) + 2\rho\tilde{\kappa}\tilde{b}b_v s_B^i f_B\left(s_B^i\right) + 2\rho\tilde{\kappa}b_0 b_v m^2 s_B^i f_B\left(s_B^i\right) + 2\rho\tilde{\kappa}b_0 b_v s_B^i f_B\left(s_B^i\right) + \tilde{\kappa}^2\tilde{b}^2 m^2 + \tilde{\kappa}^2\tilde{b}^2 - 2\tilde{\kappa}^2\tilde{b}b_v m^2 F_B\left(s_B^i\right) - 2\tilde{\kappa}^2\tilde{b}b_v F_B\left(s_B^i\right) + \tilde{\kappa}^2 b_v^2 m^2 F_B^2\left(s_B^i\right) + \tilde{\kappa}^2 b_v^2 F_B^2\left(s_B^i\right) - 2\tilde{\kappa}\tilde{b}^2 m^2 - 2\tilde{\kappa}\tilde{b}^2 - 2\tilde{\kappa}\tilde{b}b_0 m^2 - 2\tilde{\kappa}\tilde{b}b_0 + 2\tilde{\kappa}\tilde{b}b_v m^2 F_B\left(s_B^i\right) + 2\tilde{\kappa}\tilde{b}b_v F_B\left(s_B^i\right) + 2\tilde{\kappa}b_0 b_v m^2 F_B\left(s_B^i\right) + 2\tilde{\kappa}b_0 b_v F_B\left(s_B^i\right) + \tilde{b}^2 m^2 + \tilde{b}^2 + 2\tilde{b}b_0 m^2 + 2\tilde{b}b_0 + a^2\left(m^2+1\right) + a\left(-2\rho\tilde{\kappa}b_v m^2 s_B^i f_B\left(s_B^i\right) + 2\rho\tilde{\kappa}b_v s_B^i f_B\left(s_B^i\right) + 2\tilde{\kappa}\tilde{b}m^2 - 2\tilde{\kappa}\tilde{b} - 2\tilde{\kappa}b_v m^2 F_B\left(s_B^i\right) + 2\tilde{\kappa}b_v F_B\left(s_B^i\right) -\right.$

$-2\tilde{b}m^2 + 2\tilde{b} - 2b_0m^2 + 2b_0\Big) + b_0^2m^2 + b_0^2$

Indeed, the leading coefficient is positive.

[10]:
```
asol = simplify(solve(diff(phi,a),a)[0])
phi = simplify(phi.subs(a,asol))*(m**2 +1)
display(collect(expand(phi),[b**2,b]))
```

$$-\rho^2\tilde{\kappa}^2b_v^2m^4\left(s_B^i\right)^2 f_B^2\left(s_B^i\right) + 2\rho^2\tilde{\kappa}^2b_v^2m^2\left(s_B^i\right)^2 f_B^2\left(s_B^i\right) - \rho^2\tilde{\kappa}^2b_v^2\left(s_B^i\right)^2 f_B^2\left(s_B^i\right) + 8\rho\tilde{\kappa}^2b_v^2m^2s_B^iF_B\left(s_B^i\right)f_B\left(s_B^i\right) + 8\rho\tilde{\kappa}b_0b_vm^2s_B^if_B\left(s_B^i\right) + 4\tilde{\kappa}^2b_v^2m^2F_B^2\left(s_B^i\right) + 8\tilde{\kappa}b_0b_vm^2F_B\left(s_B^i\right) + \tilde{b}^2\cdot\left(4\tilde{\kappa}^2m^2 - 8\tilde{\kappa}m^2 + 4m^2\right) + \tilde{b}\left(-8\rho\tilde{\kappa}^2b_vm^2s_B^if_B\left(s_B^i\right) + 8\rho\tilde{\kappa}b_vm^2s_B^if_B\left(s_B^i\right) - 8\tilde{\kappa}^2b_vm^2F_B\left(s_B^i\right) - 8\tilde{\kappa}b_0m^2 + 8\tilde{\kappa}b_vm^2F_B\left(s_B^i\right) + 8b_0m^2\right) + 4b_0^2m^2$$

Notice that $\tilde{\kappa}^2 - 2\tilde{\kappa} + 1 = (\tilde{\kappa} - 1)^2 > 0$ for $\tilde{\kappa} \neq 1$, showing that the leading coefficient is positive. Moreover, observe that for $\tilde{\kappa} = 1$, the polynomial is actually constant in $\tilde{b}$, i.e. of degree zero.

[11]:
```
bsol = simplify(solve(Eq(diff(phi,b),0),b)[0])
display(bsol)
```

$$\frac{\rho\tilde{\kappa}b_vs_B^if_B(s_B^i) + \tilde{\kappa}b_vF_B(s_B^i) + b_0}{\tilde{\kappa} - 1}$$

If $\tilde{\kappa} \neq 1$, after minimizing over $\tilde{b}$, we conclude that the expression is minimal for a negative $\tilde{b}$. For $\tilde{\kappa} = 1$, the expression does not change with $\tilde{b}$. In either case, we can plug in $\tilde{b} = 0$ because it minimizes the expression over $\tilde{b}$ in our admissible range $[0, b_v]$.

[12]:
```
phi = collect(expand(simplify(phi.subs(b,0))),b0)
display(phi)
```

$$-\rho^2\tilde{\kappa}^2b_v^2m^4\left(s_B^i\right)^2 f_B^2\left(s_B^i\right) + 2\rho^2\tilde{\kappa}^2b_v^2m^2\left(s_B^i\right)^2 f_B^2\left(s_B^i\right) - \rho^2\tilde{\kappa}^2b_v^2\left(s_B^i\right)^2 f_B^2\left(s_B^i\right) + 8\rho\tilde{\kappa}^2b_v^2m^2s_B^iF_B\left(s_B^i\right)f_B\left(s_B^i\right) + 4\tilde{\kappa}^2b_v^2m^2F_B^2\left(s_B^i\right) + 4b_0^2m^2 + b_0\cdot\left(8\rho\tilde{\kappa}b_vm^2s_B^if_B\left(s_B^i\right) + 8\tilde{\kappa}b_vm^2F_B\left(s_B^i\right)\right)$$

Observe that the coefficient in front of b_0 is strictly positive, allowing us to minimize easily over b_0.

[13]:
```
b0sol = simplify(solve(Eq(diff(phi,b0),0),b0)[0])
display(b0sol)
```

$$-\tilde{\kappa}b_v\left(\rho s_B^if_B\left(s_B^i\right) + F_B\left(s_B^i\right)\right)$$

This expression is negative. Since we allow only positive values for b_0, our expression is minimal for $b_0 = 0$. The next step is to plug in $b_0 = 0$ and to obtain a polynomial in s_B, regarding F_B and f_B as constants.

[14]:
```
phi = phi.subs(b0,0).subs(fB,fBsymb).subs(FB,FBsymb)
display(collect(expand(phi),sbi))
```

$$4F_B^2\tilde{\kappa}^2b_v^2m^2 + 8F_B\rho\tilde{\kappa}^2b_v^2f_Bm^2s_B^i + \left(s_B^i\right)^2\left(-\rho^2\tilde{\kappa}^2b_v^2f_B^2m^4 + 2\rho^2\tilde{\kappa}^2b_v^2f_B^2m^2 - \rho^2\tilde{\kappa}^2b_v^2f_B^2\right)$$

$$\frac{2F_Bm}{\rho f_B\left(m^2-2m+1\right)}$$

$$-\frac{2F_Bm}{\rho f_B\left(m^2+2m+1\right)}$$

We see here that that the leading coefficient is negative as long as $m^4-2m^2+1 = (m^2-1)^2$ is positive, i.e. for $m \neq 1, -1$. Therefore, the polynomial expression is positive as long as s_B is between the roots of the polynomial expression. For $m = 1$, the expression is always positive for $s_B \geq 0$. We calculate the roots below.

[16]:
```
sols = solve(phi,sbi)
for sol in sols:
    display(simplify(sol))
```

$$\frac{2F_Bm}{\rho f_B\left(m^2-2m+1\right)}$$

$$-\frac{2F_Bm}{\rho f_B\left(m^2+2m+1\right)}$$

The second root being negative whereas $s_B \geq 0$, this allows us to find a positivity criterion by comparing s_B to the first root.

Analyzing ψ_B

[17]:
```
psi = Tkap**2 + m**2 * Mkap **2
display(collect(expand(psi),[a**2,a]))
```

$$\tilde{\kappa}^2\tilde{b}^2m^2 + \tilde{\kappa}^2\tilde{b}^2 - 2\tilde{\kappa}^2\tilde{b}b_vm^2F_B\left(s_B^i\right) - 2\tilde{\kappa}^2\tilde{b}b_vF_B\left(s_B^i\right) + \tilde{\kappa}^2b_v^2m^2F_B^2\left(s_B^i\right) + \tilde{\kappa}^2b_v^2F_B^2\left(s_B^i\right) - 2\tilde{\kappa}\tilde{b}^2m^2 - 2\tilde{\kappa}\tilde{b}^2 - 2\tilde{\kappa}\tilde{b}b_0m^2 - 2\tilde{\kappa}\tilde{b}b_0 + 2\tilde{\kappa}\tilde{b}b_vm^2F_B\left(s_B^i\right) + 2\tilde{\kappa}\tilde{b}b_vF_B\left(s_B^i\right) + 2\tilde{\kappa}b_0b_vm^2F_B\left(s_B^i\right) + 2\tilde{\kappa}b_0b_vF_B\left(s_B^i\right) + \tilde{b}^2m^2 + \tilde{b}^2 + 2\tilde{b}b_0m^2 + 2\tilde{b}b_0 + a^2\left(m^2+1\right) + a\left(2\tilde{\kappa}\tilde{b}m^2 - 2\tilde{\kappa}\tilde{b} - 2\tilde{\kappa}b_vm^2F_B\left(s_B^i\right) + 2\tilde{\kappa}b_vF_B\left(s_B^i\right) - 2\tilde{b}m^2 + 2\tilde{b} - 2b_0m^2 + 2b_0\right) + b_0^2m^2 + b_0^2$$

The leading coefficient in a being positive, we can minimize easily in a. We plug the result into ψ_B.

[18]:
```
asol = solve(psi.diff(a),a)[0]
psi = simplify(psi.subs(a,asol))
display(collect(expand(psi),[b**2,b]))
```

$$\frac{4\tilde{\kappa}^2 b_v^2 m^2 F_B^2(s_B^i)}{m^2+1} + \frac{8\tilde{\kappa} b_0 b_v m^2 F_B(s_B^i)}{m^2+1} + \tilde{b}^2 \cdot \left(\frac{4\tilde{\kappa}^2 m^2}{m^2+1} - \frac{8\tilde{\kappa} m^2}{m^2+1} + \frac{4m^2}{m^2+1}\right) + \tilde{b}\left(-\frac{8\tilde{\kappa}^2 b_v m^2 F_B(s_B^i)}{m^2+1} - \frac{8\tilde{\kappa} b_0 m^2}{m^2+1} + \frac{8\tilde{\kappa} b_v m^2 F_B(s_B^i)}{m^2+1} + \frac{8 b_0 m^2}{m^2+1}\right) + \frac{4 b_0^2 m^2}{m^2+1}$$

This expression as a polynomial in $\tilde{b}$ has a positive leading coefficient for $\tilde{\kappa} \neq 1$, and is otherwise constant in $\tilde{b}$. For $\tilde{\kappa} \neq 1$, we minimize over $\tilde{b}$.

[19]:
```
bsol = solve(psi.diff(b),b)[0]
display(bsol)
```

$$\frac{\tilde{\kappa} b_v F_B(s_B^i) + b_0}{\tilde{\kappa} - 1}$$

This solution being negative for $\tilde{\kappa} \neq 1$ and the aforementioned expression being constant in $\tilde{b}$ for $\tilde{\kappa} = 1$, we may plug in $\tilde{b} = 0$ as our lowest possible $\tilde{b}$.

[20]:
```
psi = simplify(psi.subs(b,0))
display(collect(expand(psi),[b0**2,b0]))
```

$$\frac{4\tilde{\kappa}^2 b_v^2 m^2 F_B^2(s_B^i)}{m^2+1} + \frac{8\tilde{\kappa} b_0 b_v m^2 F_B(s_B^i)}{m^2+1} + \frac{4 b_0^2 m^2}{m^2+1}$$

We minimize this expression over b_0, which is straightforward due to the positive leading coefficient.

[21]:
```
b0sol= solve(psi.diff(b0),b0)[0]
display(b0sol)
```

$$-\tilde{\kappa} b_v F_B\left(s_B^i\right)$$

This expression being negative, we may substitute $b_0 = 0$. Moreover, recall that we assumed $s_B = \bar{s}_B$, so that $F_B(s_B) = 1$, yielding the minimal ψ_B for $s_B = \bar{s}_B$.

[23]:
```
psi = psi.subs(b0,0)
psi = psi.subs(FB,1)
display(psi)
```

$$\frac{4\tilde{\kappa}^2 b_v^2 m^2}{m^2+1}$$