

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E
CONTABILIDADE
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

Welfare-improving misreported polls

Ganhos de bem-estar via manipulação de pesquisas eleitorais

Felipe Ricardo Durazzo

Orientador: Prof. Dr. David Turchick

São Paulo - Brasil

2018

Prof. Dr. Marco Antonio Zago
Reitor da Universidade de São Paulo

Prof. Dr. Adalberto Américo Fischmann
Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Hélio Nogueira da Cruz
Chefe do Departamento de Economia

Prof. Dr. Ariaster Baumgratz Chimeli
Coordenador do Programa de Pós-Graduação em Economia

FELIPE RICARDO DURAZZO

Welfare-improving misreported polls

Ganhos de bem-estar via manipulação de pesquisas eleitorais

Dissertação apresentada ao Departamento de Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo como requisito parcial para a obtenção do título de Mestre em Ciências.

Orientador: Prof. Dr. David Turchick

Versão Original

São Paulo - Brasil

2018

FICHA CATALOGRÁFICA

Elaborada por Rafael Mielli Rodrigues – CRB-8/7286
Seção de Processamento Técnico do SBD/FEA/USP

Durazzo, Felipe Ricardo

Welfare-improving misreported polls / Felipe Ricardo Durazzo. -- São Paulo, 2018.

37 p.

Dissertação (Mestrado) – Universidade de São Paulo, 2018.

Orientador: David Turchick.

1. Economia política 2. Teoria dos jogos 3. Eleitorado I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade. II. Título.

CDD – 330

Agradecimentos

Aos meus pais, por todo o apoio.

Ao meu orientador, Prof. David Turchick, por estar sempre disponível para conversar, seja sobre a dissertação ou qualquer outro assunto; por sempre ter depositado confiança em mim, me motivando a prosseguir na carreira acadêmica; por toda a compreensão em diversos momentos e finalmente por ser uma grande referência para mim, seja como pesquisador ou como professor.

Aos professores Marcos Nakaguma e Fernanda Estevan, por terem participado do exame de qualificação e da Avaliação de Progresso, dando sugestões para a melhora deste trabalho, e por sempre estarem dispostos a conversar.

Ao professor Filipe Campante, por ter aceitado me receber em Harvard, pelas ótimas conversas e sugestões.

Aos meus colegas de mestrado, sem os quais esses anos de mestrado teriam sido muito menos divertidos. Em especial, à Grazi, Nayara, Gina, Tiago, Ricardo, Bruce e Rafa.

Ao pessoal da secretaria de pós-graduação: Pinho e Leka, e também ao Luciano, da secretaria de graduação.

Finalmente, à Fundação de Amparo à Pesquisa de São Paulo (FAPESP), por ter me concedido uma bolsa de mestrado (número do processo 2016/18987-4) e a bolsa BEPE (número do processo 2017/12285-0); e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), por ter me concedido uma bolsa durante os primeiros meses de mestrado.

Resumo

Introduz-se um instituto de pesquisa eleitoral em um modelo de voto custoso a fim de estudar quais incentivos os institutos possuem ao divulgarem suas pesquisas. No nosso modelo, o instituto possui informação privada a respeito da distribuição de preferências da sociedade sobre os candidatos, mas pode escolher reportar incorretamente essa informação ao público. Se isso acontece, diremos que o instituto manipulou a pesquisa eleitoral. Uma preocupação comum das pessoas em relação aos institutos de pesquisa é a possibilidade de eles manipularem uma pesquisa eleitoral com o objetivo direto de beneficiar um determinado candidato. Nós mostramos que eles possuem incentivos para manipular a pesquisa mesmo na ausência de motivações partidárias. Ainda, essa manipulação aumenta o bem-estar da sociedade, em comparação com pesquisas verdadeiras.

Palavras-chaves: pesquisa eleitoral, voto custoso, voto pivotal

Abstract

We introduce an electoral pollster in a two-candidate costly voting model to study the incentives that pollsters have regarding the release of poll results. In our model, the pollster has private knowledge about the distribution of citizens' preferences, but it may report false information to the public. If this happens, we say the pollster is misreporting the poll. An often heard criticism about pollsters is that they might manipulate in order to benefit some candidate. We show that they have incentives to misreport even in the absence of ideological motives. Moreover, misreported polls are welfare-improving relative to truthful polls.

Key-words: polls, costly voting, voter turnout, misreporting, pivotal voting

Contents

1	Introduction	13
2	Model Environment	14
3	Electoral equilibrium	15
4	Electoral pollster's objective	18
5	Welfare analysis	21
6	Concluding remarks	22
7	Appendix A	23
8	Appendix B	31
9	Appendix C	36

1 Introduction

In this paper, we seek to understand how an electoral pollster concerned with its reputation would report a poll, given that it knows that the poll itself will alter voters' behavior and consequently the election result, upon which the reputation of the pollster is based.

The canonical pivotal voting model with two candidates, as in Palfrey and Rosenthal (1983) and Ledyard (1984), assumes that citizens participate in elections insofar as there exists a possibility that their vote is pivotal, that is, it changes the result of the election. To compute the probability that his vote is pivotal, a citizen needs to know the distribution of preferences for each candidate among the society. In modern society, this information is provided by electoral pollsters, that are present in virtually every democratic country. However, little is known about the incentives that pollsters have regarding the release of poll results: in other words, can we be certain that electoral pollsters are releasing truthful polls, and does it matter?¹ The opposite question, that is, if citizens have incentives to lie to electoral pollsters, has already been analyzed by Burke and Taylor (2008).

The effects of electoral polls on voter behavior have been studied in Taylor and Yildirim (2010) and Goeree and Grober (2007). Both papers, using a two-candidate costly voting model, have concluded that a poll is harmful to the citizens' expected welfare because it stimulates the 'wrong' group of citizens (that is, the group who most likely is the minority) to vote more, while, at the same time, it stimulates the group who most likely is the majority to vote less. Moreover, the first effect is stronger than the last effect. Therefore, a poll has two negative consequences: it increases the expected voting cost and it decreases the probability that the most preferred candidate wins the election.

In their models, however, the poll results are reported truthfully. We develop a model in which, by formally introducing an electoral pollster, the poll report is endogenously determined and is not necessarily truthful. More precisely, we suppose that the pollster knows the true distribution of the citizens' preferences between the two candidates, but it may report to the public a different distribution. The pollster does not have preferences among the candidates: his objective is to be highly rated, by "getting the election right". An often heard criticism about pollsters is that they might misreport in order to benefit some candidate. We show that they have incentives to misreport even in the absence of partisan motives.

Our main finding is that, for a sufficiently large population, truthful reports do not happen. Moreover, this misreporting of the poll is welfare improving relative to truthfully reporting, for it increases the probability that the 'right' candidate wins the election and it lowers the expected voting cost.

Furthermore, our model provides i) a mechanism to explain why, contrary to the canonical pivotal voting model's unrealistic prediction, elections are not ties in expected value, and ii) another explanation to the emergence of the so called "bandwagon effect".

The paper is organized as follows. Section 2 presents the model. Section 3 characterizes the electoral equilibrium of the citizens, given the information released by the pollster. Section 4 presents the utility function of the pollster and derives the solution of its optimization problem. In Section 5, the welfare comparison between misreported polls and truthfull polls is done. Section 6 concludes.

¹We remark that a pollster can misreport information by faking data, by rigging its sampling procedure, by adjusting its statistical methodology, by framing questions in a specific way or even by signaling to the respondent that it favors a certain candidate, as shown in Bischooping and Shuman (1992).

2 Model Environment

The main elements of the model are described below.

There are $n \geq 2$ citizens, two candidates (B and R) and an electoral pollster. Voting cost, $c > 0$, is homogenous among citizens. A citizen gains 1 unit of utility if his most preferred wins the election, while if his least preferred candidate is the winner, he loses 1 unit of utility. An election tie is broken by the toss of a fair coin. Voting is voluntary.

The probability that a citizen favors candidate B (or R) is q ($1-q$). We assume that $q \in [\bar{q}, 1 - \bar{q}]$, in which $\bar{q} \in (0, 0.5)$.²

Only the electoral pollster knows the probability q . The pollster declares to the citizens that the probability that a citizen favors candidate B is p . Since citizens do not know q , they will use p as an estimate for this parameter in their voting decisions.

We say the pollster is misreporting if $p \neq q$.

The citizens are instrumental voters. A citizen should vote if and only if

$$\Pi b \geq c,$$

in which Π is the probability of being pivotal in the election and b is the benefit associated with being pivotal.

A voter is pivotal if and only if his vote creates or breaks a tie. In both situations, the expected increase in utility for voting is 1, therefore $b = 1$.³

Notice that given the probability p reported by the pollster, the citizens play among themselves a Bayesian Game.

We focus on type-symmetric Bayesian Nash equilibria, in which the strategies played by the citizens are homogenous among types (B-voters or R-voters) but might differ between types.

Given that each B-citizen is voting with probability γ_B and each R-citizen is voting with probability γ_R , the probability that a B-voter is pivotal (through his lens) is given by:

$$\begin{aligned} \Pi_B(n, p, \gamma_B, \gamma_R) = & \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{n-1}{k, k, n-1-2k} (p\gamma_B)^k ((1-p)\gamma_R)^k \times \\ & (1 - p\gamma_B - (1-p)\gamma_R)^{n-1-2k} + \\ & \sum_{k=0}^{\lfloor \frac{n-2}{2} \rfloor} \binom{n-1}{k, k+1, n-2-2k} (p\gamma_B)^k ((1-p)\gamma_R)^{k+1} \times \\ & (1 - p\gamma_B - (1-p)\gamma_R)^{n-2-2k}, \end{aligned} \quad (1)$$

in which the first summation refers to the probability of breaking a tie and the second summation

²The technical reason for this assumption will be clear later on. Still, one can think that \bar{q} is very close to 0 and there would not be much sense in having elections in such harmonious society.

³An implicit assumption is that citizens are risk neutral. Grillo (2017) allows for risk-averse citizens.

refers to the probability of creating a tie. An analogous expression holds for an R-voter:

$$\begin{aligned} \Pi_R(n, p, \gamma_B, \gamma_R) &= \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{n-1}{k, k, n-1-2k} ((1-p)\gamma_R)^k (p\gamma_B)^k \times \\ &\quad (1 - (1-p)\gamma_R - p\gamma_B)^{n-1-2k} + \\ &\quad \sum_{k=0}^{\lfloor \frac{n-2}{2} \rfloor} \binom{n-1}{k, k+1, n-2-2k} ((1-p)\gamma_R)^k (p\gamma_B)^{k+1} \times \\ &\quad (1 - (1-p)\gamma_R - p\gamma_B)^{n-2-2k}, \end{aligned}$$

Notice that the following holds:

$$\Pi_R(n, p, \gamma_B, \gamma_R) = \Pi_B(n, 1-p, \gamma_R, \gamma_B).$$

The pollster is not an ideological one: by assumption, his objective is to "get the election right". That is, the electoral result must be coherent with the probability p . Bear in mind that this is not trivial because the voters' equilibrium is a function of p . The exact objective function we assume will be presented later on.

In summary, the timing of the model is as follows:

- Nature chooses the type of each citizen
- The pollster discovers the probability q (probability that a given citizen favors candidate B)
- The pollster reports p
- Each voter chooses to vote or to abstain
- The election happens
- Each citizen and the pollster receive their payoffs

3 Electoral equilibrium

In this section, we characterize the Bayesian-Nash equilibrium among the citizens, given the poll report.

Define

$$\begin{aligned} c_l(n, p) &= \begin{cases} \Pi_B(n, p, 1, 1) & \text{if } p \geq 0.5 \\ \Pi_R(n, p, 1, 1) & \text{if } p < 0.5 \end{cases}, \\ c_u(n, p) &= \begin{cases} \Pi_B\left(n, p, \frac{1-p}{p}, 1\right) & \text{if } p \geq 0.5 \\ \Pi_R\left(n, p, 1, \frac{p}{1-p}\right) & \text{if } p < 0.5 \end{cases}. \end{aligned}$$

Notice that, by Claim 4 (stated and proved in Appendix A), $c_l(n, p) < c_u(n, p)$.

The complete characterization of the electoral equilibrium is given by the following proposition.

Proposition 1 Given n, c and $p \in (0, 1)$ there is one and only one type-symmetric electoral equilibrium (γ_B^*, γ_R^*) . If $c \leq c_l$, then $(\gamma_B^*, \gamma_R^*) = (1, 1)$. If $c \geq 1$, then $(\gamma_B^*, \gamma_R^*) = (0, 0)$. If $c \in (c_u, 1)$ then $(\gamma_B^*, \gamma_R^*) \in (0, 1)^2$ is such that $\Pi_B\left(n, p, \gamma_B^*, \frac{1-p}{p}\gamma_B^*\right) = c$ and $\gamma_R^* = \frac{p}{1-p}\gamma_B^*$. For $p \geq 0.5$, if $c \in (c_l, c_u]$ then $\gamma_R^* = 1$ and $\gamma_B^* \in [\frac{1-p}{p}, 1)$ solves $\Pi_B(n, p, \gamma_B^*, 1) = c$ ($\gamma_B^* = \frac{1-p}{p}$ if and only if $c = c_u$). For $p < 0.5$, if $c \in (c_l, c_u]$ then $\gamma_B^* = 1$ and $\gamma_R^* \in \left(\frac{p}{1-p}, 1\right)$ solves $\Pi_R(n, p, 1, \gamma_R^*) = c$ ($\gamma_R^* = \frac{p}{1-p}$ if and only if $c = c_u$).

Proof. See Appendix A. ■

The comparative statics of the electoral equilibrium with respect to the parameters c and p are described in the following lemmas.

Lemma 1 Given n , in an interior electoral equilibrium, γ_B^* and γ_R^* are strictly decreasing in c . For $p \geq 0.5$ and $c \in (c_l, c_u)$, γ_B^* is strictly decreasing in c . For $p < 0.5$ and $c \in (c_l, c_u)$, γ_R^* is strictly decreasing in c .

Proof. See Appendix A. ■

Lemma 2 Given n and c , in an interior electoral equilibrium, γ_B^* is strictly decreasing in p and γ_R^* is strictly increasing in p .

Proof. See Appendix A. ■

The following two figures illustrate the electoral equilibrium and the comparative statics with respect to c and p .

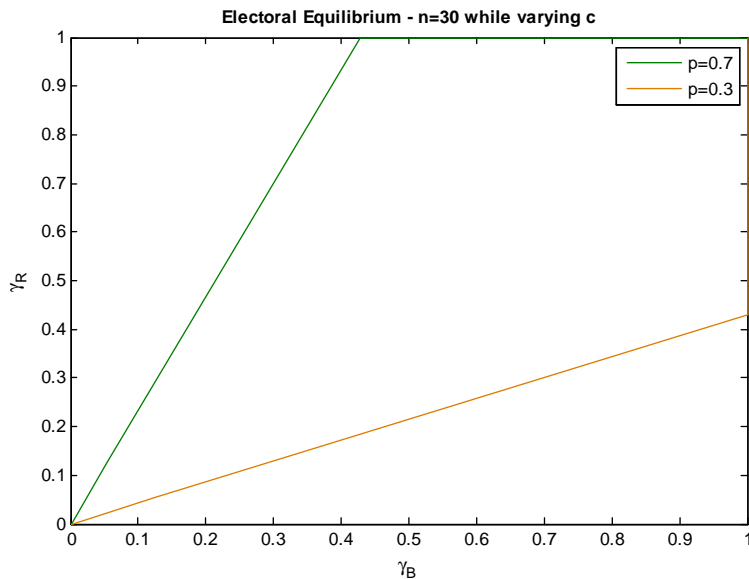


Figure 1

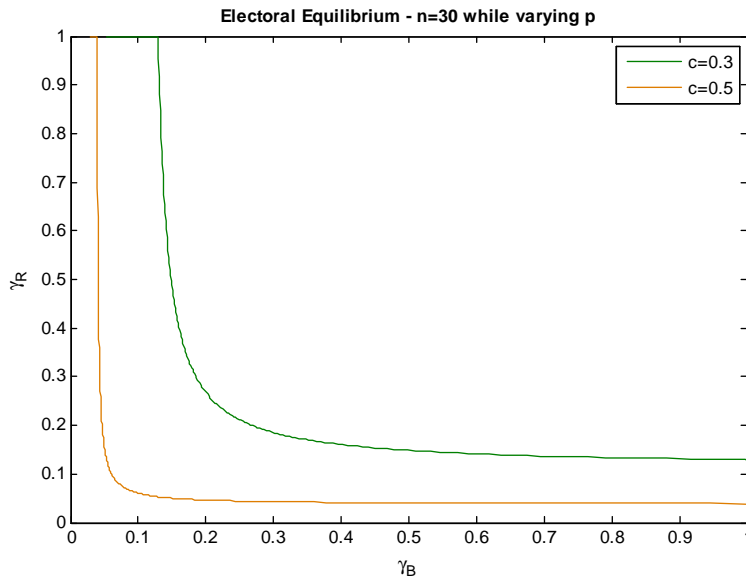


Figure 2

The following lemmas about the behavior of the function $c_u(n, p)$ will be useful.

Lemma 3 Given $p \in (0, 1)$, $c_u(n, p) \searrow 0$. Moreover,

$$\lim_{n \rightarrow \infty} \gamma_B(n, c, p) = \lim_{n \rightarrow \infty} \gamma_R(n, c, p) = 0.$$

and the limits as n grows, of $n\gamma_B(n, c, p)$ and $n\gamma_R(n, c, p)$, exist.

Proof. This is proved by Taylor and Yildirim (2010), in their Lemma 3 and Lemma 7. ■

Lemma 4 For $p \geq 0.5$, $c_u(n, p)$ is strictly increasing in p . For $p < 0.5$, $c_u(n, p)$ is strictly decreasing in p .

Proof. See Appendix A. ■

Notice that Lemma 3 implies that given $c \in (0, 1)$ and p , there exists a critical population size, which depends on p , such that if n is bigger than this critical value, the electoral equilibrium is interior. The following lemma guarantees the existence of a critical population size that does not depend on p .⁴

Lemma 5 Given $c \in (0, 1)$, there exists $n_0(c) \in \mathbb{N}$ such that

$$n \geq n_0(c) \Rightarrow (\gamma_B(n, c, p), \gamma_R(n, c, p)) \in (0, 1)^2, \forall p \in [\bar{q}, 1 - \bar{q}].$$

⁴This is where the assumption that $q \in [\bar{q}, 1 - \bar{q}]$ plays an important role. As we will see in the next section, with this assumption, the pollster would never report $p < \bar{q}$ or $p > \bar{q}$. Thus, its effective choice set is the interval $[\bar{q}, 1 - \bar{q}]$. If we had let $q \in (0, 1)$, there would not exist a uniform critical population size that would guarantee an interior electoral equilibrium regardless of the $p \in (0, 1)$ chosen by the pollster.

Proof. See Appendix A. ■

As an illustration of the above lemma, for $c = 0.5$ and $c = 0.7$, $n_0 = 113$ and $n_0 = 45$, respectively, would suffice.

Finally, the following lemma holds:

Lemma 6 *Given $n, c \in (0, 1)$ and p , if the type-symmetric electoral equilibrium is interior, then $\gamma_B^*(n, c, p) = \gamma^*/(2p)$ and $\gamma_R^*(n, c, p) = \gamma^*/(2(1-p))$, in which $\gamma^* := \gamma_B^*(n, c, 0.5) (= \gamma_R^*(n, c, 0.5))$ – that is, γ^* is such that*

$$\Pi_B(n, 0.5, \gamma^*, \gamma^*) = c.$$

Proof. This is proved by Goeree and Grober (2007), in their Proposition 3. ■

4 Electoral pollster's objective

The main concern of the pollster is to be considered of good quality, by "getting the election right". Notice that the information reported by the pollster and the electoral equilibrium implies a distribution for the result of the election. More precisely, this result is a random vector $x = (i, j, k)$, in which i, j, k denotes the number of votes for B , the number of votes for R and abstentions, respectively. The true probability distribution (known only by the pollster) of x is

$$(i, j, k) \sim \text{Multinomial} \left(\begin{array}{l} n, q\gamma_B(n, c, p), (1-q)\gamma_R(n, c, p), \\ 1 - q\gamma_B(n, c, p) - (1-q)\gamma_R(n, c, p) \end{array} \right)$$

The citizens, however, believe that x is distributed as

$$(i, j, k) \sim \text{Multinomial} \left(\begin{array}{l} n, p\gamma_B(n, c, p), (1-p)\gamma_R(n, c, p), \\ 1 - p\gamma_B(n, c, p) - (1-p)\gamma_R(n, c, p) \end{array} \right)$$

In particular, the citizens believe that, on average, candidate B will have the same number of votes than candidate R, as, by Lemma 3, in an interior electoral equilibrium, we have

$$p\gamma_B^*(n, c, p) = (1-p)\gamma_R^*(n, c, p).$$

The rating (reputation) of the pollster depends on how the actual election result compares with the result implied by the report. We assume that given an election result $x = (i, j, k)$, the rating of the pollster is given by:

$$- \left[(i - np\gamma_B(n, c, p))^2 + (j - n(1-p)\gamma_R(n, c, p))^2 \right],$$

which is a measure of how far the election result was from the result expected by the citizens.⁵

As the actual election result is random, the utility function of the pollster is the expected value of its rating. That is,

$$U(n, c, p, q) = -E \left[(i - np\gamma_B(n, c, p))^2 + (j - n(1-p)\gamma_R(n, c, p))^2 \right],$$

⁵The website FiveThirtyEight.com, for instance, publishes a ranking of electoral pollsters in which the rating of a pollster is computed in a similar way. The difference is that usually pollsters report a direct prediction of the election result (e.g. "candidate B will win by 10 points"), whereas in our model, the pollster reports the probability that someone favors candidate B.

in which E is taken with respect to the true probability distribution (that depends on q).⁶

Note that, as $q \in [\bar{q}, 1 - \bar{q}]$, the pollster would never report $p < \bar{q}$ or $p > 1 - \bar{q}$, for it would become clear to the citizens that the pollster is lying.

The pollster's optimization problem is:

$$\max_{p \in [\bar{q}, 1 - \bar{q}]} U(n, c, p, q).$$

We denote a solution of this problem as $p^*(n, c, q)$.

This function can be rewritten as:

$$\begin{aligned} U(n, c, p, q) &= - \left[\mathbb{E}(i - np\gamma_B(n, c, p))^2 + \mathbb{E}(j - n(1-p)\gamma_R(n, c, p))^2 \right] \\ &= \left[\begin{array}{c} \text{Var}(i - np\gamma_B(n, c, p)) + \\ (np\gamma_B(n, c, p) - n(1-p)\gamma_R(n, c, p))^2 + \\ \text{Var}(j - n(1-p)\gamma_R(n, c, p)) + \\ (n(1-q)\gamma_R(n, c, p) - n((1-p)\gamma_R(n, c, p)))^2 \end{array} \right] \\ &= \left[\begin{array}{c} \text{Var}(i) + n^2(p-q)^2\gamma_B^2(n, c, p) + \\ \text{Var}(j) + n^2(p-q)^2\gamma_R^2(n, c, p) \end{array} \right] \end{aligned}$$

in which $\text{Var}(i) = n(q\gamma_B(n, c, p)(1 - q\gamma_B(n, c, p)))$ and $\text{Var}(j) = n((1-q)\gamma_R(n, c, p)(1 - (1-q)\gamma_R(n, c, p)))$.

The following remark shows that if voting was not costly, then there would not be any incentive to misreport.

Remark 1 For any n and q , if $c = 0$, then $p^*(n, c, q) = q$ is the unique solution of the pollster's problem.

Proof. If $c = 0$, then, by Proposition 1, the electoral equilibrium is $(\gamma_B^*, \gamma_R^*) = (1, 1)$. Plugging these values in the objective function, it becomes:

$$U(n, c, p, q) = -2nq(1-q) - 2n^2(p-q)^2.$$

It is clear that the maximum value of U occurs at $p = q$. ■

Note that for large n ,

$$\text{Var}(i) + \text{Var}(j) = \text{Var}(i - j) + 2\text{Cov}(i, j) \approx \text{Var}(i - j).⁷$$

⁶It may be noted that our utility function resembles, in absolute value,

$$\begin{aligned} &\frac{E(i - np\gamma_B)^2}{np\gamma_B} + \frac{E(j - n(1-p)\gamma_R)^2}{n(1-p)\gamma_R} + \\ &\frac{E((n-i-j) - n(1-p\gamma_B - (1-p)\gamma_R))^2}{n(1-p\gamma_B - (1-p)\gamma_R)}. \end{aligned}$$

This happens to be the expected value of the statistic of the Pearson Chi-Squared Test of goodness of fit for a multinomial distribution, with a null hypothesis that the true probability distribution of x has parameters $(p\gamma_B, (1-p)\gamma_R, 1-p\gamma_B - (1-p)\gamma_R)$. Furthermore, it can be shown that, as $n \rightarrow \infty$, the problem of maximizing U and the problem of minimizing this more complicated expression have the same solution (their FOCs even coincide). In this sense, by maximizing our utility function, the null hypothesis that the pollster is truthful, with a high probability, would not be rejected.

⁷Note that $\text{Cov}(i, j) = -nq\gamma_B(n, c, p)(1-q)\gamma_R(n, c, p)$, the limit of which is 0 as n grows, by Lemma 3.

To gain some intuition on the incentives of the electoral pollster, notice that, assuming an interior equilibrium, we know that the citizens expect the election to be a tie. Only by reporting $p = q$, the number of votes for B will equal the number of votes for R on average. However, the rating of the pollster will be based not on averages of elections, but on this particular election. In this sense, the pollster might be willing to report a p that does not imply an expected election tie, but one such that the difference between the votes for B and R is still small and has less dispersion.

In an interior equilibrium, after some manipulation, the utility function can be rewritten as:

$$U(n, c, p, q) = - \left[\frac{n\gamma}{2} \left(\frac{q}{p} + \frac{1-q}{1-p} \right) - \frac{n\gamma^2}{4} \left(\frac{q^2}{p^2} + \frac{(1-q)^2}{(1-p)^2} \right) + \frac{n^2\gamma^2}{4} (p-q)^2 \left(\frac{1}{p^2} + \frac{1}{(1-p)^2} \right) \right],$$

in which $\gamma = \gamma^*(n, c, 0.5)$.

The following proposition shows that misreporting is the norm, rather than the exception.

Proposition 2 *Given $c \in (0, 1)$, for a sufficiently large n , the solution of the pollster's optimization problem, $p^*(n, c, q)$, is unique and such that:*

- i. *If $q = 0.5$, then $p^*(n, c, q) = q$.*
- ii. *If $q > 0.5$, then $p^*(n, c, q) \in (0.5, q)$.*
- iii. *If $q < 0.5$, then $p^*(n, c, q) \in (q, 0.5)$.*

Proof. See Appendix B. ■

To illustrate this proposition, below we plot the function $p^*(n, c, q)$ for $n = 50$ and $c = 0.7$, together with the 45-degree line.

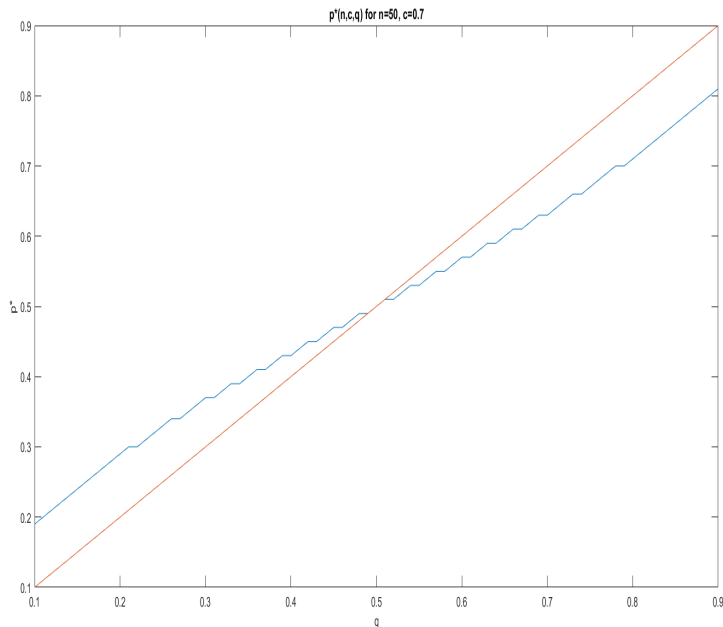


Figure 3

Thus, if there exists an expected majority in the society ($q \neq 0.5$) and n is sufficiently large so that the electoral equilibrium is interior, then a rational pollster driven purely by reputational motives will always misreport information. Moreover, it will do so in such a way that the citizens believe that the majority is not as large as it actually is. As we will see in the next section, contrary to common belief, misreporting actually increases the expected welfare of the citizens, relative to a truthful report.

Some comments about the implications of our model should be made.

Firstly, with misreporting, the election will not be a tie, in expected terms. Rather, if $q > 0.5$, then the expected result of the election is a win for candidate B , as the expected number of votes for B ($nq\gamma/2p$) is greater than the expected number of votes for R ($n(1-q)\gamma/2(1-p)$). Thus, our model provides a mechanism to explain why, in practice, the expected result of an election is not a tie, as implied by the canonical pivotal voting model.⁸

Secondly, we provide another mechanism to the emergence of the so called “bandwagon effect” – the phenomenon in which supporters of a specific candidate are more likely to cast their votes if their candidate ranks first in the polls. Our model implies that part of this effect could actually be an illusion. For example, if $q = 0.80$ and $n = 50$, the pollster would report $p^* = 0.71$. The citizens expect that, on average, the election will be a tie, or, in another way, that the ratio of the expected number of votes for each candidate should be

$$\frac{np\gamma_B(n, c, p)}{n(1-p)\gamma_R(n, c, p)} = 1,$$

from Lemma 6. However, in actuality, this ratio would be

$$\frac{nq\gamma_B(n, c, p)}{n(1-q)\gamma_R(n, c, p)} = \frac{q/p}{(1-q)/(1-p)} \approx 1.62.$$

Thus, candidate B will receive, on average, 62% more votes than candidate R . A political analyst who believes that poll results are truthful, by observing such a discrepancy in a given election, could erroneously be led into thinking that the bandwagon effect was in place – that is, that the electoral equilibrium (γ_B and γ_R) as generated by the pivotal voting model was incorrect – without realizing that the p and the $1-p$ terms in the ratio of the expected number of votes were wrong in the first place. In this way, our model implies that part of what is usually attributed to the bandwagon effect is actually a direct consequence of misreporting.

5 Welfare analysis

Now that we have characterized the electoral equilibrium and the solution of the pollster’s problem, we proceed by analyzing the welfare implications of misreporting for the citizens. If we denote N_B

⁸Another mechanism would be through different voting costs among the types of citizens. However, this approach has the disadvantage of imposing, in an ad hoc way, an asymmetry in the model.

as the number of citizens that prefer B , the expected gross benefit of the election, given n, c, p, q is^{9,10}:

$$I(n, c, p, q) = \Pr(B \text{ wins} | n, c, p, q) (E(N_B) - E(n - N_B)) + \Pr(R \text{ wins} | n, c, p, q) (E(n - N_B) - E(N_B))$$

By using that $E(N_B) = nq$, we get

$$I(n, c, p, q) = n(2q - 1) (\Pr(B \text{ wins} | n, c, p, q) - \Pr(R \text{ wins} | n, c, p, q))$$

The expected cost of the election is given by:

$$\begin{aligned} C(n, c, p, q) &= nq\gamma_B(n, c, p) + n(1 - q)\gamma_R(n, c, p) \\ &= n(q\gamma_B(n, c, p) + (1 - q)\gamma_R(n, c, p)) \end{aligned}$$

Thus, the welfare function is given by:

$$W(n, c, p, q) = I(n, c, p, q) - C(n, c, p, q)$$

Proposition 3 *Given $c \in (0, 1)$, $q \in [\bar{q}, 1 - \bar{q}]$, for a sufficiently large election, the misreported poll is welfare-improving relative to a truthful poll. That is,*

$$W(n, c, p^*(n, c, q), q) > W(n, c, q, q).$$

Proof. See Appendix B. ■

The intuition for this proposition is as follows. Suppose that $q > 0.5$. In this case, it would be good for society if candidate B was the expected winner of the election, as, more often than not, the number of B -citizens will be larger than the number of R -citizens. Recall that, with a truthful report ($p = q$), the expected result of the election is a tie, so that the probability that candidate B wins is $1/2$. However, with the misreported poll, $p^* \in (0.5, q)$, by Proposition 2, and by Lemma 2, the B -citizens (R -citizens) are going to vote with a higher (lower) probability relative to the truthful poll, so that the probability that B wins is greater than $1/2$. Besides that, in absolute value, the probability of a B -vote ends up changing less than the probability of an R -vote, in such a way that the expected voting cost decreases. Thus, as the expected benefit increases with misreporting and the expected voting cost decreases, we conclude that a misreported poll is unambiguously welfare-improving relative to a truthful poll.

6 Concluding remarks

In this paper, we formally introduced an electoral pollster in an two-candidate costly voting model, and have concluded that a pollster driven only by reputational reasons will misreport the poll results.

⁹For analytical tractability, we proceed as in Taylor and Yildirim (2010) and disregard the dependence between the number of citizens that prefer B and the probability that candidate B wins the election while computing the gross benefit. To the interested reader, in Appendix C we provide the precise gross benefit function (which can also be seen in Goeree and Grober (2006), page 66) and show that, for large n , the benefit function used in this section is a good approximation of the precise one.

¹⁰We remark that we are only interested in the utility of the citizens, that is, the utility function of the electoral pollster does not enter the welfare function of the model.

Moreover, the misreporting is welfare improving relative to truthful reporting, for it increases the probability that the most preferred wins the election and it also lowers the expected voting cost.

Through misreporting, we can explain why, contrary to what the pivotal voting model would predict, election results are not ties, in expected terms. Our model also implies that part of what is usually attributed to the bandwagon effect could actually be just a consequence of misreporting. In future research, we plan to show that misreported polls can also be welfare improving relative to a world with no polls.

One limitation of our work is that we assume that citizens fully believe the information released by the poll. A more realistic model would let the citizens know the utility function of the pollster, thereby constraining its choice set, as in the literature of information transmission.

Finally, an interesting extension would be to add more electoral pollsters in the model and to analyze how the incentives regarding the release of poll results would change.

References

- [1] K. BISCHOPING AND H. SCHUMAN. Pens and polls in nicaragua: An analysis of the 1990 preelection surveys. *American Journal of Political Science* pp. 331–350 (1992).
- [2] J. BURKE AND C. R. TAYLOR. What’s in a poll? incentives for truthful reporting in pre-election opinion surveys. *Public Choice* **137**(1-2), 221–244 (2008).
- [3] J. K. GOEREE AND J. GROSSER. Welfare reducing polls. *Economic Theory* **31**(1), 51–68 (2007).
- [4] A. GRILLO. Risk aversion and bandwagon effect in the pivotal voter model. *Public Choice* **172**(3-4), 465–482 (2017).
- [5] J. O. LEDYARD. The pure theory of large two-candidate elections. *Public choice* **44**(1), 7–41 (1984).
- [6] T. R. PALFREY AND H. ROSENTHAL. A strategic calculus of voting. *Public choice* **41**(1), 7–53 (1983).
- [7] C. R. TAYLOR AND H. YILDIRIM. Public information and electoral bias. *Games and Economic Behavior* **68**(1), 353–375 (2010).

7 Appendix A

Before proceeding to the proof of Proposition 1, we state and prove five useful claims about the behavior of the functions Π_B and Π_R .

A priori, there are 9 possibilities of type-symmetric electoral equilibria. Each possibility is characterized by the following conditions.

1. $(\gamma_B, \gamma_R) \in (0, 1)^2$; $\Pi_B\left(n, p, \gamma_B, \frac{p}{1-p}\gamma_B\right) = c$ and $\gamma_R = \frac{p}{1-p}\gamma_B$
2. $\gamma_B \in (0, 1), \gamma_R = 0$; $\Pi_B(n, p, \gamma_B, 0) = c$ and $\Pi_R(n, p, \gamma_B, 0) \leq c$
3. $\gamma_B \in (0, 1), \gamma_R = 1$; $\Pi_B(n, p, \gamma_B, 1) = c$ and $\Pi_R(n, p, \gamma_B, 1) \geq c$

4. $\gamma_B = 0, \gamma_R \in (0, 1); \Pi_B(n, p, 0, \gamma_R) \leq c$ and $\Pi_R(n, p, 0, \gamma_R) = c$
5. $\gamma_B = 0, \gamma_R = 0; \Pi_B(n, p, 0, 0) \leq c$ and $\Pi_R(n, p, 0, 0) \leq c$
6. $\gamma_B = 0, \gamma_R = 1; \Pi_B(n, p, 0, 1) \leq c$ and $\Pi_R(n, p, 0, 1) \geq c$
7. $\gamma_B = 1, \gamma_R \in (0, 1); \Pi_B(n, p, 1, \gamma_R) \geq c$ and $\Pi_R(n, p, 1, \gamma_R) = c$.
8. $\gamma_B = 1, \gamma_R = 0; \Pi_B(n, p, 1, 0) \geq c$ and $\Pi_R(n, p, 1, 0) \leq c$
9. $\gamma_B = 1, \gamma_R = 1; \Pi_B(n, p, 1, 1) \geq c$ and $\Pi_R(n, p, 1, 1) \geq c$

Claim 1 *The situations (2), (4), (6) and (8) cannot occur.*

Proof. Notice that, for $\gamma_R > 0$,

$$\begin{aligned} \Pi_B(n, p, 0, \gamma_R) &= (1 - (1-p)\gamma_R)^{n-1} + (n-1)((1-p)\gamma_R)(1 - (1-p)\gamma_R)^{n-2} \\ \Pi_R(n, p, 0, \gamma_R) &= \Pi_B(n, 1-p, \gamma_R, 0) = (1 - (1-p)\gamma_R)^{n-1} \end{aligned}$$

Suppose that (4) or (6) is true. Then,

$$\Pi_B(n, p, 0, \gamma_R) \leq \Pi_R(n, p, 0, \gamma_R)$$

Thus

$$\begin{aligned} &(1 - (1-p)\gamma_R)^{n-1} + (n-1)((1-p)\gamma_R)(1 - (1-p)\gamma_R)^{n-2} \\ &\leq (1 - (1-p)\gamma_R)^{n-1} \end{aligned}$$

which implies that

$$(n-1)((1-p)\gamma_R)(1 - (1-p)\gamma_R)^{n-2} \leq 0$$

Assuming $p \in (0, 1)$, this is a contradiction as all terms of the product are strictly positive.

The claim for the situations (2) and (8) follows from the above and by noticing that

$$\Pi_R(n, p, \gamma_B, \gamma_R) = \Pi_B(n, 1-p, \gamma_R, \gamma_B)$$

That is, supposing that (2) or (8) is true would implies that

$$(n-1)(p\gamma_B)(1 - p\gamma_B)^{n-2} \leq 0$$

another contradiction. ■

Claim 2 *Situation (5), that is, $(\gamma_B, \gamma_R) = (0, 0)$ occurs if and only if $c \geq 1$.*

Proof. Simply notice that

$$\Pi_B(n, p, 0, 0) = \Pi_R(n, p, 0, 0) = 1$$

Condition 5 becomes $c \geq 1$, as claimed. ■

Claim 3 *Situation (7) occurs only if $\gamma_R \geq \frac{p}{1-p}$ and $p < 0.5$. Situation (3) occurs only if $\gamma_B \geq \frac{1-p}{p}$ and $p \geq 0.5$.*

Proof. Notice that

$$\begin{aligned}\Pi_B(n, p, 1, \gamma_R) &= \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{n-1}{k, k, n-1-2k} p^k ((1-p)\gamma_R)^k \times \\ &\quad (1-p - (1-p)\gamma_R)^{n-1-2k} + \\ &\quad \sum_{k=0}^{\lfloor \frac{n-2}{2} \rfloor} \binom{n-1}{k, k+1, n-2-2k} p^k ((1-p)\gamma_R)^{k+1} \times \\ &\quad (1-p - (1-p)\gamma_R)^{n-2-2k}\end{aligned}$$

and

$$\begin{aligned}\Pi_B(n, p, 1, \gamma_R) &= \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{n-1}{k, k, n-1-2k} p^k ((1-p)\gamma_R)^k \times \\ &\quad (1-p - (1-p)\gamma_R)^{n-1-2k} + \\ &\quad \sum_{k=0}^{\lfloor \frac{n-2}{2} \rfloor} \binom{n-1}{k, k+1, n-2-2k} p^{k+1} ((1-p)\gamma_R)^k \times \\ &\quad (1-p - (1-p)\gamma_R)^{n-2-2k}\end{aligned}$$

Suppose that (7) holds. Thus

$$\Pi_B(n, p, 1, \gamma_R) \geq \Pi_R(n, p, 1, \gamma_R)$$

which implies that

$$\begin{aligned}&\sum_{k=0}^{\lfloor \frac{n-2}{2} \rfloor} \binom{n-1}{k, k+1, n-2-2k} (1-p - (1-p)\gamma_R)^{n-2-2k} \times \\ &p^k ((1-p)\gamma_R)^k ((1-p)\gamma_R - p) \geq 0\end{aligned}$$

Notice that the terms $(1-p - (1-p)\gamma_R)^{n-2-2k}$ and $((1-p)\gamma_R)^k$ are strictly positive. Thus the inequality is true only if

$$(1-p)\gamma_R - p \geq 0$$

This holds only if $\gamma_R \geq \frac{p}{1-p}$, which, in turn, is true only if $\frac{p}{1-p} < 1$. This implies that we must have $p \leq 0.5$.

The proof for the situation (3) is completely analogous. By the same reasoning above, we conclude that (3) holds only if $\gamma_B \geq \frac{1-p}{p}$, which implies that $p \geq 0.5$. ■

Claim 4 *The function $\Pi_B\left(n, p, \gamma_B, \frac{p}{1-p}\gamma_B\right)$ is strictly decreasing in γ_B for $\gamma_B \in \left[0, \frac{1-p}{p}\right]$. The function $\Pi_B(n, p, \gamma_B, 1)$ is strictly decreasing in γ_B for $\gamma_B \in \left[\frac{1-p}{p}, 1\right]$ and the function $\Pi_R(n, p, 1, \gamma_R)$ is strictly decreasing in γ_R for $\gamma_R \in \left[\frac{p}{1-p}, 1\right]$.*

Proof. The first assertion is proved in Taylor and Yildirim (2010). We prove here the second one by computing $\frac{\partial \Pi_B(n, p, \gamma_B, 1)}{\partial \gamma_B}$ and checking its sign.

$$\begin{aligned}
& \frac{\partial \Pi_B(n, p, \gamma_B, 1)}{\partial \gamma_B} \\
= & \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \frac{(n-1)! p^k k \gamma^{k-1}}{k! k! (n-1-2k)!} (1-p)^k (p(1-\gamma))^{n-1-2k} - \\
& \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \frac{(n-1)!}{k! k! (n-1-2k)!} (p\gamma)^k (1-p)^k p^{n-1-2k} (n-1-2k) (1-\gamma)^{n-2-2k} + \\
& \sum_{k=0}^{\lfloor \frac{n-2}{2} \rfloor} \frac{(n-1)! p^k k \gamma^{k-1} (1-p)^{k+1}}{k! (k+1)! (n-2-2k)!} (p(1-\gamma))^{n-2-2k} - \\
& \sum_{k=0}^{\lfloor \frac{n-2}{2} \rfloor} \frac{(n-1)! (p\gamma)^k (1-p)^{k+1} p^{n-2-2k}}{k! (k+1)! (n-2-2k)!} (n-2-2k) (1-\gamma)^{n-3-2k} \\
\leq & (n-1)! \left[\frac{\sum_{k=1}^{\lfloor \frac{n-1}{2} \rfloor} \frac{p^k k \gamma^{k-1} (1-p)^k}{k! k! (n-1-2k)!} (p(1-\gamma))^{n-1-2k} -}{\sum_{k=0}^{\lfloor \frac{n-2}{2} \rfloor} \frac{(p\gamma)^k (1-p)^{k+1} p^{n-2-2k}}{k! (k+1)! (n-2-2k)!} (n-2-2k) (1-\gamma)^{n-3-2k}} \right] \\
& + \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} (k(1-p) - (k+1)p\gamma) \frac{(n-1)!}{k! (k+1)! (n-2-2k)!} p^{n-2-k} (1-p)^k \gamma^{k-1} (1-\gamma)^{n-2-2k} \\
= & (n-1)! \left[\frac{\sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor - 1} \frac{p^{k+1} \gamma^k (1-p)^{k+1}}{k! (k+1)! (n-3-2k)!} (p(1-\gamma))^{n-3-2k}}{\sum_{k=0}^{\lfloor \frac{n-2}{2} \rfloor} \frac{(p\gamma)^k (1-p)^{k+1} p^{n-2-2k}}{k! (k+1)! (n-2-2k)!} (n-2-2k) (1-\gamma)^{n-3-2k}} \right] + \\
& \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} (k(1-p) - (k+1)p\gamma) \frac{(n-1)! p^{n-2-k} (1-p)^k}{k! (k+1)! (n-2-2k)!} \gamma^{k-1} (1-\gamma)^{n-2-2k} \\
\leq & \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor - 1} \left((1-\gamma)^{n-3-2k} - (1-\gamma)^{n-3-2k} \right) \frac{(n-1)! p^{n-2-k}}{k! (k+1)! (n-3-2k)!} \gamma^k (1-p)^{k+1} + \\
& \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} (k(1-p) - (k+1)p\gamma) \frac{(n-1)! p^{n-2-k} (1-p)^k}{k! (k+1)! (n-2-2k)!} \gamma^{k-1} (1-\gamma)^{n-2-2k} \\
& + \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} (k(1-p) - (k+1)p\gamma) \frac{(n-1)! p^{n-2-k} (1-p)^k}{k! (k+1)! (n-2-2k)!} \gamma^{k-1} (1-\gamma)^{n-2-2k}
\end{aligned}$$

where the first inequality combines the fourth sum with the first one (whose first summand vanishes when $k=0$) and the second sum with the third one (possibly with an extra final – and positive – term, in case n is odd), and the second inequality possibly (if n is even) discards the final – also positive – term of the sum to which a negative sign is attached.

Now notice that for each k ,

$$k(1-p) - (k+1)p\gamma \leq k(1-p) - (k+1)p\frac{1-p}{p} = p-1 < 0.$$

Thus

$$\frac{\partial \Pi_B(n, p, \gamma_B, 1)}{\partial \gamma_B} < 0.$$

Because

$$\Pi_R(n, p, 1, \gamma_R) = \Pi_B(n, (1-p), \gamma_R, 1),$$

we know that

$$\begin{aligned} \frac{\partial \Pi_R(n, p, 1, \gamma_R)}{\partial \gamma_R} &= \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} (kp - (k+1)(1-p)\gamma_R) \frac{(n-1)!}{k!(k+1)!(n-2-2k)!} \times \\ &\quad (1-p)^{n-2-k} p^k \gamma_R^{k-1} (1-\gamma_R)^{n-2-2k} \end{aligned}$$

Remark that

$$kp - (k+1)(1-p)\gamma_R \leq kp - (k+1)(1-p)\frac{p}{(1-p)} = -p < 0.$$

Thus

$$\frac{\partial \Pi_R(n, p, 1, \gamma_R)}{\partial \gamma_R} < 0.$$

■

Claim 5 For $p \geq 0.5$ and $\gamma_B \geq \frac{1-p}{p}$, $\Pi_R(n, p, \gamma_B, 1) \geq \Pi_B(n, p, \gamma_B, 1)$. For $p < 0.5$ and $\gamma_R \geq \frac{p}{1-p}$, $\Pi_B(n, p, 1, \gamma_R) \geq \Pi_R(n, p, 1, \gamma_R)$.

Proof. Recall that

$$\begin{aligned} \Pi_B(n, p, \gamma_B, 1) &= \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{n-1}{k, k, n-1-2k} (p\gamma_B)^k \times \\ &\quad (1-p)^k (1-p\gamma_B - (1-p))^{n-1-2k} \\ &\quad + \sum_{k=0}^{\lfloor \frac{n-2}{2} \rfloor} \binom{n-1}{k, k+1, n-2-2k} (p\gamma_B)^k \times \\ &\quad (1-p)^{k+1} (1-p\gamma_B - (1-p))^{n-2-k} \end{aligned}$$

and

$$\begin{aligned} \Pi_R(n, p, \gamma_B, 1) &= \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{n-1}{k, k, n-1-2k} (1-p)^k \times \\ &\quad (p\gamma_B)^k (1-p\gamma_B - (1-p))^{n-1-2k} \\ &\quad + \sum_{k=0}^{\lfloor \frac{n-2}{2} \rfloor} \binom{n-1}{k, k+1, n-2-2k} (1-p)^k \times \\ &\quad (p\gamma_B)^{k+1} (1-p\gamma_B - (1-p))^{n-2-2k} \end{aligned}$$

Thus $\Pi_R(n, p, \gamma_B, 1) - \Pi_B(n, p, \gamma_B, 1)$ is given by

$$\sum_{k=0}^{\lfloor \frac{n-2}{2} \rfloor} \binom{n-1}{k, k+1, n-2-2k} (1-p\gamma_B - (1-p))^{n-2-2k} ((1-p)p\gamma_B)^k (p\gamma_B - (1-p))$$

If $\gamma_B \geq \frac{1-p}{p}$ then

$$p\gamma_B - (1-p) \geq p \frac{(1-p)}{p} - (1-p) = 0$$

and

$$1 - p\gamma_B - (1-p) = p - p\gamma_B \geq 0$$

which implies that

$$\Pi_R(n, p, \gamma_B, 1) - \Pi_B(n, p, \gamma_B, 1) \geq 0$$

thus proving the first assertion.

The second assertion follows naturally from the proof above and recalling that $\Pi_R(n, p, 1, \gamma_R) = \Pi_B(n, 1-p, \gamma_R, 1)$. ■

Now we are ready to prove Proposition 1.

Proof of Proposition 1. Suppose that $p \geq 0.5$. By Claims 1 and 3, an equilibrium can occur only in situations (1), (3), (5) or (9).

If $c \leq c_l$, that is, $\Pi_B(n, p, 1, 1) \geq c$, then, by Claim 5, $\Pi_R(n, p, 1, 1) \geq c$ and by the condition in (9) we know that $(\gamma_B, \gamma_R) = (1, 1)$ is an equilibrium.

If $c \geq 1$, then by Claim 2, we know that $(\gamma_B, \gamma_R) = (0, 0)$ is an equilibrium.

If $c \in (c_l, c_u)$ then we know, by the definitions of c_l and c_u that

$$\Pi_B\left(n, p, \frac{1-p}{p}, 1\right) > c$$

and

$$\Pi_B(n, p, 1, 1) < c$$

The function $\Pi_B(n, p, \gamma_B, 1)$ is polynomial in γ_B , thus continuous in γ_B . By the Intermediate Value Theorem (IVT), there exists $\gamma'_B \in \left(\frac{1-p}{p}, 1\right)$ such that

$$\Pi_B(n, p, \gamma'_B, 1) = c$$

By Claim 5,

$$\Pi_R(n, p, \gamma'_B, 1) \geq c$$

thus $(\gamma'_B, 1)$ is an equilibrium (situation 3).

If $c = c_u$ then

$$\Pi_B\left(n, p, \frac{1-p}{p}, 1\right) = c$$

and by Claim 5,

$$\Pi_R\left(n, p, \frac{1-p}{p}, 1\right) \geq c$$

Thus $\left(\frac{1-p}{p}, 1\right)$ is an equilibrium (situation 3).

Define

$$\Pi(n, p, \gamma_B) := \Pi_B\left(n, p, \gamma_B, \frac{p}{1-p}\gamma_B\right)$$

Notice that

$$\Pi_B\left(n, p, \frac{1-p}{p}, 1\right) = \Pi\left(n, p, \frac{1-p}{p}\right)$$

and that

$$\Pi_B(n, p, 0, 0) = \Pi(n, p, 0)$$

If $c \in (c_u, 1)$ then

$$\Pi_B\left(n, p, \frac{1-p}{p}, 1\right) < c$$

and

$$1 = \Pi_B(n, p, 0, 0) > c$$

which is equivalent to

$$\Pi(n, p, 0) > c \text{ and } \Pi\left(n, p, \frac{1-p}{p}\right) < c$$

By the IVT, there exists $\gamma_B'' \in \left(0, \frac{1-p}{p}\right)$ such that $\Pi(n, p, \gamma_B'') = c$. By the definition of the function Π , $\left(\gamma_B'', \frac{p}{1-p}\gamma_B''\right)$ is an equilibrium (situation 1).

By Claim 4, there can be at most one $\gamma_B^* \in [0, \frac{1-p}{p})$ that solves

$$\Pi_B\left(n, p, \gamma_B^*, \frac{p}{1-p}\gamma_B^*\right) = c$$

and at most one $\gamma_B^{**} \in \left[\frac{1-p}{p}, 1\right]$ that solves

$$\Pi_B(n, p, \gamma_B^{**}, 1) = c$$

Furthermore, also by Claim 4, existence of such γ_B^* and γ_B^{**} at the same time is not possible (Π_B is a polynomial, and hence continuous, in γ_B - particularly, at $\gamma_B = \frac{1-p}{p}$). Thus the type-symmetric equilibrium is unique.

The proof for the case $p < 0.5$ is entirely similar and is omitted here. ■

Proof of Lemma 1. In an interior electoral equilibrium, we know that (γ_B^*, γ_R^*) is such that $\gamma_B^* \in \left(0, \frac{1-p}{p}\right)$,

$$\Pi_B\left(n, p, \gamma_B^*, \frac{p}{1-p}\gamma_B^*\right) = c \text{ and } \gamma_R^* = \frac{p}{1-p}\gamma_B^*$$

By the first assertion in Claim 4, we know that $\Pi_B\left(n, p, \gamma_B, \frac{p}{1-p}\gamma_B\right)$ is strictly decreasing in γ_B .

Thus if c increases (decreases) then γ_B^* must decrease (increase) otherwise the equality $\Pi_B\left(n, p, \gamma_B^*, \frac{p}{1-p}\gamma_B^*\right) = c$ would not hold. Furthermore, γ_R^* must change in the same direction as γ_B^* , since $\gamma_R^* = \frac{p}{1-p}\gamma_B^*$.

Now suppose that $p \geq 0.5$ and $c \in (c_l, c_u)$. By Proposition 1, we know that γ_B^* is such that it solves

$$\Pi_B(n, p, \gamma_B, 1) = c$$

By the second assertion in Claim 4, we know that $\Pi_B(n, p, \gamma_B, 1)$ is strictly decreasing in γ_B . Thus if c increases (decreases) then γ_B^* must decrease (increase).

The proof for $p < 0.5$ is similar and is omitted here. ■

Proof of Lemma 2. Recall that, by Lemma 3, in an interior electoral equilibrium,

$$\gamma_B^*(n, c, p) = \frac{\gamma(n, c, 0.5)}{2p} \text{ and } \gamma_R^*(n, c, p) = \frac{\gamma(n, c, 0.5)}{2(1-p)}.$$

Thus

$$\frac{\partial \gamma_B(n, c, p)}{\partial p} < 0 \text{ and } \frac{\partial \gamma_R(n, c, p)}{\partial p} > 0.$$

■

Proof of Lemma 4. Suppose that $p \geq 0.5$. Recall that, in this case,

$$\begin{aligned} c_u(n, p) &= \Pi_B\left(n, p, \frac{1-p}{p}, 1\right) \\ &= \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{n-1}{k, k, n-1-2k} (1-p)^k ((1-p))^k (1-(1-p)-(1-p))^{n-1-2k} + \\ &\quad \sum_{k=0}^{\lfloor \frac{n-2}{2} \rfloor} \binom{n-1}{k, k+1, n-2-2k} (1-p)^k ((1-p))^{k+1} (1-(1-p)-(1-p))^{n-2-2k} \end{aligned}$$

Notice that this is the expression of function $P(\alpha, n)$, defined in Taylor and Yildirim (2010) at page 357, evaluated at $\alpha = 1-p$.

In their Lemma 2, they proved that $P(\alpha, n)$ is strictly decreasing in $\alpha \in [0, 0.5]$.

Thus, if p increases, then a decreases, which implies that c_u increases.

The proof for $p < 0.5$ follows directly from the same reasoning above, by noticing that, in this case, $c_u(n, p)$ is equal to $P(\alpha, n)$ evaluated at $\alpha = p$. ■

Proof of Lemma 5. Consider the function $\bar{n}(c, p)$ such that, for $(c, p) \in (0, 1) \times (\bar{q}, 1 - \bar{q})$, $\bar{n}(c, p)$ is the smallest number k such that $c_u(k, p) < c$.

Notice that, as $c_u(n, p) \searrow 0$ (Lemma 1), $\bar{n}(c, p)$ is well-defined and if $n' \geq \bar{n}(c, p)$ then $c_u(n', p) < c$.

We will argue that for $p \in [0.5, 1 - \bar{q}]$, $\bar{n}(c, p)$ is increasing in p , whereas for $p \in [\bar{q}, 0.5]$, $\bar{n}(c, p)$ is decreasing in p .

Suppose that $p \in [0.5, 1 - \bar{q}]$ and let $n_1 := \bar{n}(c, 1 - \bar{q})$. By Lemma 2,

$$c_u(n_1, p) \leq c_u(n_1, 1 - \bar{q}) < c$$

This implies that

$$\bar{n}(c, p) \leq n_1 = \bar{n}(c, 1 - \bar{q})$$

by the definition of the function $\bar{n}(c, p)$.

Suppose now that $p \in [\bar{q}, 0.5)$ and let $n_2 := \bar{n}(c, \bar{q})$. By Lemma 2,

$$c_u(n_2, p) \leq c_u(n_2, \bar{q}) < c$$

whence, by the same reasoning as before,

$$\bar{n}(c, p) \leq n_2 = \bar{n}(c, \bar{q})$$

Define $n_0(c) := \max\{\bar{n}(c, \bar{q}), \bar{n}(c, 1 - \bar{q})\}$

Since $c_u(n, p)$ decreases with n , for all $p \in [\bar{q}, 1 - \bar{q}]$,

$$n \geq n_0(c) \Rightarrow c \in (c_u(n, p), 1)$$

which, by Proposition 1, is equivalent to saying that the electoral equilibrium is interior. ■

8 Appendix B

Before proving Proposition 2, we state a lemma about the utility function.

Lemma 7 *Given c , if $n \geq n_0(c)$ (so that the electoral equilibrium is interior, regardless of the p chosen) and $n\gamma(n, c, 0.5) \leq 1$, then $U(n, c, p, q)$ is strictly concave in p .*

Proof. Rearranging terms, we can rewrite U as:

$$\begin{aligned} U &= - \left[\frac{n\gamma}{2} \left(\frac{q}{p} + \frac{1-q}{1-p} \right) - \frac{n\gamma^2}{4} \left(\frac{q^2}{p^2} + \frac{(1-q)^2}{(1-p)^2} \right) + \frac{n^2\gamma^2}{4} (p-q)^2 \left(\frac{1}{p^2} + \frac{1}{(1-p)^2} \right) \right] \\ &= \left[\frac{n\gamma^2}{4} \frac{q^2}{p^2} - \frac{n\gamma}{2} \frac{q}{p} - \frac{n^2\gamma^2}{4} (p-q)^2 \frac{1}{p^2} \right] \\ &\quad + \left[\frac{n\gamma^2}{4} \frac{(1-q)^2}{(1-p)^2} - \frac{n\gamma}{2} \frac{(1-q)}{(1-p)} - \frac{n^2\gamma^2}{4} (p-q)^2 \frac{1}{(1-p)^2} \right] \end{aligned}$$

We now check that each term in square brackets is strictly concave by computing its second derivative.

$$\frac{\partial^2}{\partial p^2} \left(\frac{n\gamma^2}{4} \frac{q^2}{p^2} - \frac{n\gamma}{2} \frac{q}{p} - \frac{n^2\gamma^2}{4} (p-q)^2 \frac{1}{p^2} \right) = \frac{1}{2} nq\gamma \frac{(n\gamma - 1)2p + 3q\gamma(1-n)}{p^4}$$

The sign of this derivative is the same as the sign of

$$(n\gamma - 1)2p + 3q\gamma(1-n)$$

which is negative for $n\gamma \leq 1$. Therefore, the function in the first square bracket is strictly concave in p .

$$\begin{aligned} &\frac{\partial^2}{\partial p^2} \left(\frac{n\gamma^2}{4} \frac{(1-q)^2}{(1-p)^2} - \frac{n\gamma}{2} \frac{(1-q)}{(1-p)} - \frac{n^2\gamma^2}{4} (p-q)^2 \frac{1}{(1-p)^2} \right) \\ &= \frac{1}{2} n\gamma(1-q) \frac{(1-n\gamma)2p - (n\gamma - 3\gamma + 3q\gamma - 3nq\gamma + 2)}{(1-p)^4} \end{aligned}$$

The sign of this derivative is the as the same the sign of

$$(1 - n\gamma) 2p - (n\gamma - 3\gamma + 3q\gamma - 3nq\gamma + 2)$$

As $n\gamma \leq 1$, this last expression is strictly increasing in p . Plugging $p = 1$, it becomes

$$2(1 - n\gamma) - (n\gamma - 3\gamma + 3q\gamma - 3nq\gamma + 2) = 3\gamma(q - 1)(n - 1)$$

which is negative. Therefore, as the maximum value of this expression is negative, it follows that it is negative for all $p \in (0, 1)$.

If $n\gamma = 1$, then derivative becomes

$$-(1 - 3\gamma + 3q\gamma - 3q + 2) = 3(1 - \gamma)(q - 1) < 0$$

Therefore, U , being the sum of strictly concave functions, is also strictly concave. ■

Proof of Proposition 2. First, we prove ii).

Suppose that $q > 0.5$. By Lemma 5, there exists $n_0(c)$ such that, for all $p \in [\bar{q}, 1 - \bar{q}]$, the electoral equilibrium is interior. We assume that $n \geq n_0$.

The partial derivative of $U(n, c, p, q)$, in an interior equilibrium, with respect to p is

$$\begin{aligned} & \frac{\partial}{\partial p} U(n, c, q, p) \\ = & -\frac{1}{2} \frac{n}{p^3} \frac{\gamma}{(1-p)^3} \left[\begin{array}{l} (2q + n\gamma - 2nq\gamma - 1)p^4 + \\ (2q\gamma - \gamma - 4q - 2q^2\gamma + 2nq\gamma + 2nq^2\gamma + 1)p^3 \\ + (3q + 3q^2\gamma - 3nq\gamma - 3nq^2\gamma)p^2 + \\ (nq\gamma - 3q^2\gamma - q + 3nq^2\gamma)p + (q^2\gamma - nq^2\gamma) \end{array} \right] \end{aligned}$$

For notational simplicity, denote the term in square brackets as the function $f(n, \gamma, q, p)$. The first-order condition of the optimization problem implies that if $p^* \in (\bar{q}, 1 - \bar{q})$ is a solution to the problem, then

$$f(n, \gamma, q, p^*) = 0$$

We now show that there exists a root of f between 0.5 and q . Notice that

$$f(n, \gamma, q, q) = q^2(2q - 1)(1 - q)^2(1 - \gamma) > 0$$

and that

$$f(n, \gamma, q, 0.5) = -\frac{1}{16}(-2\gamma + n\gamma + 1)(2q - 1) \sim 2\gamma - n\gamma - 1 = (2 - n)\gamma - 1$$

which is a strictly decreasing of γ . At the minimum value ($\gamma = 0$) it is equal to $-1 < 0$. Therefore, $f(n, \gamma, q, 0.5)$ is negative for all $\gamma \in [0, 1]$.

Because f is polynomial in p (and hence, continuous), by the Intermediate Value Theorem, there exists $p^* \in (0.5, q)$ such that $f(n, \gamma, q, p^*) = 0$. That is, p^* is a critical point of U in the interval $(\bar{q}, 1 - \bar{q})$. We now claim that p^* is the only critical point in this interval.

For $n\gamma \leq 1$, by Lemma 7, $U(n, c, p, q)$ is a strictly concave function, thus there can be at most one critical point, which also is the solution of optimization problem.

For $n\gamma > 1$, the proof is less straightforward. We proceed by showing that $f(n, \gamma, q, p)$ is strictly monotone in p . Notice that

$$\begin{aligned} \frac{\partial}{\partial p} f(n, \gamma, q, p) &= \left(\begin{aligned} &4np^3 + 3nq^2 + 6pq^2 + 6p^2q - 6p^2q^2 + nq - \\ &3p^2 - 3q^2 - 6npq^2 + 6np^2q - 8np^3q + 6np^2q^2 - 6npq \end{aligned} \right)^\gamma \\ &\quad + (8p^3q - 12p^2q - q + 6pq + 3p^2 - 4p^3) \end{aligned}$$

which is affine in γ . At $\gamma = 1/n$, the derivative is equal to

$$\begin{aligned} \frac{\partial}{\partial p} f(n, \gamma, q, p) &= 3(n-1) \frac{-2pq^2 - 2p^2q + 2p^2q^2 + p^2 + q^2}{n} \sim \\ &\quad -2pq^2 - 2p^2q + 2p^2q^2 + p^2 + q^2 \\ &= p^2(1-q)^2 + q^2(1-p)^2 > 0 \end{aligned}$$

At $\gamma = 1$, the derivative is equal to

$$\begin{aligned} \frac{\partial}{\partial p} f(n, 1, q, p) &= (n-1) \left((6p^2 - 6p + 3)q^2 + (6p^2 - 8p^3 - 6p + 1)q + 4p^3 \right) \sim \\ g(p, q) &: = (6p^2 - 6p + 3)q^2 + (6p^2 - 8p^3 - 6p + 1)q + 4p^3 \end{aligned}$$

We now argue that $g(p, q)$ is positive. Notice that $g(p, q)$ is a convex parabola, as its leading coefficient is

$$6p^2 - 6p + 3 = 3(1-p)^2 + 3p^2 > 0$$

The vertex of the parabola, given by $-(6p^2 - 8p^3 - 6p + 1) / (2(6p^2 - 6p + 3))$ is plotted in the following figure:

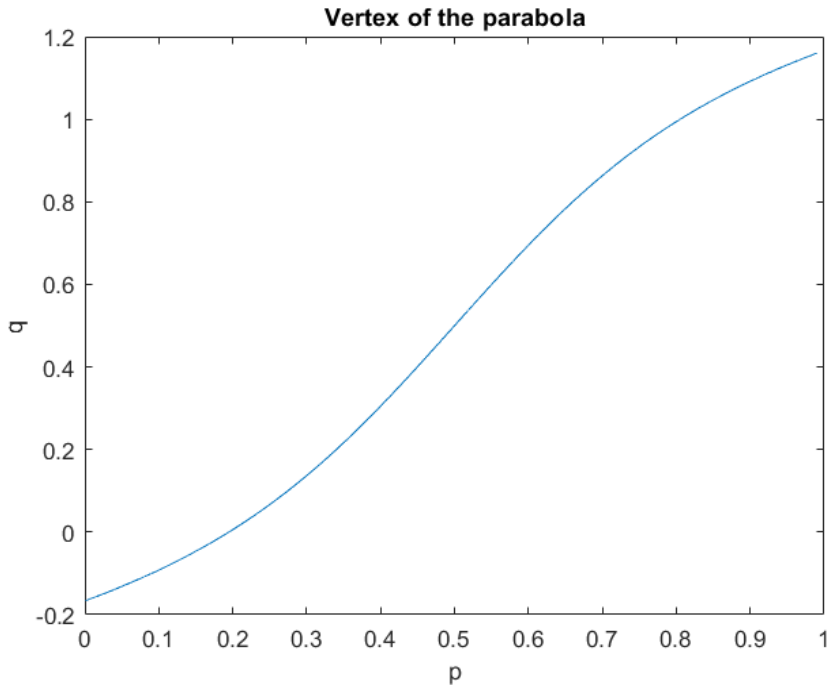


Figure 4

Notice that for $p > 0.85$, the vertex occurs at $q > 1$, implying that the parabola is decreasing for $q \in [0, 1]$. That is, the minimum value, in this case, occurs at $q = 1$. Hence,

$$g(p, 1) = 4(1 - p)^3 > 0$$

For $p < 0.19$, the vertex occurs at $q < 0$, implying that the parabola is increasing for q . The minimum value of g occurs at $q = 0$.

$$g(p, 0) = 4p^3 > 0$$

The minimum value that g attains is given by

$$-\frac{\left((6p^2 - 8p^3 - 6p + 1)^2 - 4(6p^2 - 6p + 3)4p^3\right)}{4(6p^2 - 6p + 3)}$$

and is plotted in the following figure:

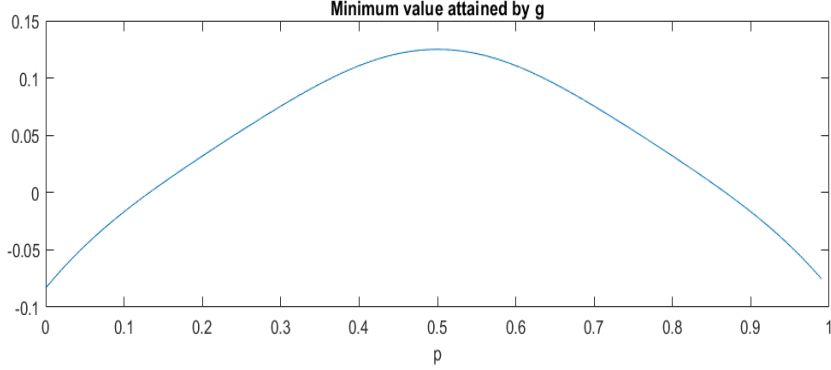


Figure 5

As it can be seen, for $p \in [0.19, 0.85]$ the minimum value of g is also positive, from which we conclude that g is positive for all $p \in [\bar{q}, 1 - \bar{q}]$.

Therefore as both $\frac{\partial}{\partial p} f(n, 1, q, p)$ and $\frac{\partial}{\partial p} f(n, 1, q, p)$ are positive for all $(p, q) \in (0, 1) \times (\bar{q}, 1 - \bar{q})$ and $\frac{\partial}{\partial p} f(n, \gamma, q, p)$ is affine in γ , it follows that $f(n, \gamma, q, p)$ is strictly increasing for all $p \in [0, 1]$. Thus there can be at most one root of $f(n, \gamma, q, p)$ for $p \in [0, 1]$, which implies that there can be at most one critical point of $U(n, c, q, p)$.

Finally, note that $\frac{\partial}{\partial p} U(n, c, q, p)$ has the opposite sign of $f(n, \gamma, q, p)$. As we showed that $f(n, \gamma, q, 0.5) < 0$, $\frac{\partial}{\partial p} U(n, c, q, p)$ evaluated at $p = 0.5$ is positive, hence $p = 0.5$ cannot be a solution. Besides, as $f(n, \gamma, q, p^*) = 0$ and $p^* < 1 - \bar{q}$ it follows that $f(n, \gamma, q, 1 - \bar{q}) > 0$, which implies that $\frac{\partial}{\partial p} U(n, c, q, p)$ at $p = 1 - \bar{q}$ is negative, hence $1 - \bar{q}$ is not a solution.

Therefore, $p^* \in (0.5, q)$ is the unique solution of the optimization problem, as claimed.

The proof for iii) is entirely analogous and is omitted here.

To prove i), suppose that $q = 0.5$. In this case,

$$\frac{\partial}{\partial p} U(n, c, 0.5, p) = -\frac{1}{8} \frac{n}{p^3} \gamma \frac{(2p-1)}{(p-1)^3} (\gamma - 2p - n\gamma - p\gamma + p^2\gamma + 2p^2 + 3np\gamma - 3np^2\gamma)$$

which equals 0 at $p = 0.5$, as $(2p - 1) = 0$. So $p = 0.5$ is a critical point of U , which is also the only one, as shown in part ii). ■

Proof of Proposition 3. Suppose that $q > 0.5$.

First, we argue that $I(n, c, p^*(n, c, q), q) > I(n, c, q, q)$.

Denote as X_B and X_R the number of votes for B and R , respectively. Then,

$$\Pr(B \text{ wins} | n, c, p, q) = \Pr(X_B > X_R | n, c, p, q) + \frac{1}{2} \Pr(X_B = X_R | n, c, p, q)$$

Denoting by N_B the number of citizens that favors candidate B , we have that

$$I(n, c, p, q) = E_{N_B} [I(n, c, p, q) | N_B]$$

Conditional on N_B , X_B and X_R are binomial random variables with probabilities of success $\theta_B(p, q) := \left(q \frac{\gamma}{2p}\right)$ and $\theta_R(p, q) := \left((1-q) \frac{\gamma}{2(1-p)}\right)$, respectively. Thus

$$\begin{aligned} & I(n, c, p^*(n, c, q), q | N_B) - I(n, c, q, q | N_B) \\ &= n(2q-1) \left[\frac{\Pr(X_B > X_R | N_B, c, p^*(n, c, q), q) - \Pr(X_B > X_R | N_B, c, q, q)}{\Pr(X_B < X_R | N_B, c, q, q) - \Pr(X_B < X_R | N_B, c, p^*(n, c, q), q)} + \right] \end{aligned}$$

As $\theta_B(p^*(n, c, q), q) > \theta_B(q, q)$ and $\theta_R(p^*(n, c, q), q) < \theta_R(q, q)$, from Proposition 2, it follows that

$$\begin{aligned} & (\Pr(X_B > X_R | N_B, c, p^*(n, c, q), q) - \Pr(X_B > X_R | N_B, c, q, q)) > 0 \\ & (\Pr(X_B < X_R | N_B, c, q, q) - \Pr(X_B < X_R | N_B, c, p^*(n, c, q), q)) > 0 \end{aligned}$$

Thus, for each possible realization of N_B ,

$$I(n, c, p^*(n, c, q), q | N_B) - I(n, c, q, q | N_B) > 0$$

which, taking the expected value with respect to N_B , implies that

$$I(n, c, p^*(n, c, q), q) > I(n, c, q, q)$$

Now we argue that $C(n, c, p^*(n, c, q), q) < C(n, c, q, q)$. Suppose, aiming at a contradiction, that this is not true. Then, by Lemma 6,

$$\begin{aligned} & n \left(q \frac{\gamma}{2p^*} + (1-q) \frac{\gamma}{2(1-p^*)} \right) \geq n\gamma \\ \Rightarrow & \left(\frac{q}{p^*} + \frac{1-q}{1-p^*} \right) \geq 2 \\ \Rightarrow & \frac{q(1-p^*) + p^*(1-q)}{p^*(1-p^*)} \geq 2 \\ \Rightarrow & q(1-p^*) + p^*(1-q) \geq 2p^*(1-p^*) \\ \Rightarrow & q + p^* - 2p^*q \geq 2p^*(1-p^*) \\ \Rightarrow & q - 2p^*q \geq 2p^*(1-p) - p^* \\ \Rightarrow & q(1-2p^*) \geq p^*(1-2p^*) \end{aligned}$$

Notice that as $p^* > 1/2$, from Proposition 2, $1-2p^* < 0$, so cancelling out $1-2p^*$ in the last expression, one gets

$$q \leq p^*$$

a contradiction, since $p^* < q$, again by Proposition 2.

Thus $W(n, c, p^*(n, c, q), q) > W(n, c, q, q)$.

The proof for the case $q < 0.5$ is analogous and is omitted here. ■

9 Appendix C

The goal of this Appendix is to state the precise gross benefit function and to show that this function converges in probability to the function used in the text, thereby making the latter a good approximation of the former. We denote the precise gross benefit by \hat{I} .

Suppose that there exists $0 \leq N_B \leq n$ citizens that favor candidate B . Recall that N_B is binomially distributed, with parameters (n, q) . Conditional on N_B , the benefit is:

$$\begin{aligned}
\hat{I}(n, c, p, q, N_B) &= \left[\begin{array}{l} \Pr(B \text{ wins}|n, c, p, N_B) (N_B - (n - N_B)) + \\ \Pr(R \text{ wins}|n, c, p, N_B) (n - N_B - N_B) \end{array} \right] \\
&= (2N_B - n) (\Pr(B \text{ wins}|n, c, p, N_B) - \Pr(R \text{ wins}|n, c, p, N_B)) \\
&= (2N_B - n) (2\Pr(B \text{ wins}|n, c, p, N_B) - 1) \\
&= 4N_B \Pr(B \text{ wins}|n, c, p, N_B) - 2N_B - 2n \Pr(B \text{ wins}|n, c, p, N_B) + n
\end{aligned}$$

Taking the expected value of $\hat{I}(n, c, p, q, N_B)$ with respect to N_B yields

$$E \left[\hat{I}(n, c, p, q, N_B) \right] = 4E [N_B \Pr(B \text{ wins}|n, c, p, N_B)] - 2nq - 2n \Pr(B \text{ wins}|n, c, p, q) + n$$

in which it was used that $E [\Pr(B \text{ wins}|n, c, p, N_B)] = \Pr(B \text{ wins}|n, c, p, q)$.

Dividing the expression by n , one gets:

$$\begin{aligned}
E \left[\frac{1}{n} \hat{I}(n, c, p, q, N_B) \right] &= 4E \left[\frac{N_B}{n} \Pr(B \text{ wins}|n, c, p, q, N_B) \right] - 2q - \\
&\quad 2\Pr(B \text{ wins}|n, c, p, q) + 1 \\
&= \left[4 \left(\frac{\text{Cov} \left(\frac{N_B}{n}, \Pr(B \text{ wins}|n, c, p, q, N_B) \right) +}{q \Pr(B \text{ wins}|n, c, p, q)} \right) - \right] \\
&\quad \left[2q - 2\Pr(B \text{ wins}|n, c, p, q) + 1 \right] \\
&= \left[4\text{Cov} \left(\frac{N_B}{n}, \Pr(B \text{ wins}|n, c, p, q, N_B) \right) \right] \\
&\quad \left[+ 2\Pr(B \text{ wins}|n, c, p, q) (2q - 1) - (2q - 1) \right] \\
&= \left[4\text{Cov} \left(\frac{N_B}{n}, \Pr(B \text{ wins}|n, c, p, q, N_B) \right) + \right] \\
&\quad \left[(2q - 1) (2\Pr(B \text{ wins}|n, c, p, q) - 1) \right]
\end{aligned}$$

Notice that, by the Weak Law of Large Numbers,

$$\text{plim}_{n \rightarrow \infty} \frac{N_B}{n} = q$$

and that $\Pr(B \text{ wins}|n, c, p, q, N_B) - \Pr(B \text{ wins}|n, c, p, q)$ is bounded. Hence, as the term in square brackets converges (in probability) to 0, we have

$$E \left[\frac{1}{n} \hat{I}(n, c, p, q, N_B) \right] \approx (2q - 1) (2\Pr(B \text{ wins}|n, c, p, q) - 1) = \frac{1}{n} I(n, c, p, q).$$