A Note on Forward Induction in a Model of Representative Democracy.*

F. De Sinopoli

CORE
Voie du Roman Pays 1348 Louvain-la-Neuve, Belgium

November, 1999

Abstract

The citizen-candidate approach, proposed to study the performance of representative democracies, builds on a multi-stage game where the same agents are asked whether or not to become a candidate and, successively, to vote. Consistently, the solution concept adopted in Besley and Coate (1997) conforms to backward induction rationality. In this note we remark that it does not conform to forward induction rationality.

JEL Classification: C72, D72

Keywords: Voting games, refinements of Nash equilibrium, citizen-candidate, forward induction, Mertens' stable sets.

*I would like to thank Alessandro Turrini for many discussions which originated this paper. I am deeply indebted to my advisor, Jean-François Mertens, whose suggestions have improved this work. Moreover he pointed out an error in a previous version of this paper. I also thank Michel Le Breton for helpful comments. Financial support from Université Catholique de Louvain is gratefully acknowledged. The usual disclaimer applies.

1 Introduction

The "citizen-candidate" approach to study the outcome and the performance of representative democracies (Osborne and Slivinski (1996), Besley and Coate (1997)) is among the most known in formal voting analysis. In a previous paper (De Sinopoli and Turrini (1999)) it is remarked that the solution concept adopted in Besley and Coate (1997) may fail to capture the behavior of rational agents in that it does not conform to the iterated dominance principle. In a subsequent paper, Dhillon and Lockwood (1999) proposed a different solution concept to solve the game presented in Besley and Coate (1997), by adding the requirement that the equilibrium at the voting stage is to be iteratively weakly undominated.

In this note we want to stress that imposing restrictions on voting behavior is not sufficient to eliminate "irrational" outcomes arising from the solution concept adopted in Besley and Coate (1997).

The solution concept proposed by Besley and Coate (1997) to solve citizen-candidate models makes use of sequential rationality, hence it conforms to the backward induction principle. When deciding upon their own candidacy, citizens are required to make a selection of one particular voting equilibrium among those that may be realized for all possible set of candidates. Each citizen decides to run for election only if this is judged to be profitable, given the expected voting equilibria. However, when the same agents are called to move at different stages of the game, backward induction is not sufficient to induce rationality through the game tree. Exactly as agents are allowed to read the game tree backward, they must also be able to read it "forward", by asking themselves, at each subgame, what the other players should have rationally done in the previous stages of the game. This is an issue that typically arises in citizen-candidate models, since all agents are asked to move both at the candidacy and at the voting stage.

The requirement of forward induction rationality has been first illustrated in Kohlberg (1981), then formalized in Kohlberg and Mertens (1986), and widely used in refinements of signaling games (see, e.g., Cho and Kreps (1987)).[1] Since its introduction, however, the term "forward induction" has been used with slightly different meanings.[2] In this paper we refer to forward induction to mean the process through a player, deviating from a prescribed equilibrium, can signal to the opponents his subsequent behavior, changing the equilibrium induced in some subgame.

Forward induction arguments are of primary relevance when the equilibrium at some non-initial node of the game tree is not unique. At that stage, reading forward what should have happened along the game tree, players might be able to exclude some rivals' strategies from the set of those that might be played with positive probability. As a consequence, forward induction may eliminate

[1] Among relevant contributions concerning the concept of forward induction see also van Damme (1989), Osborne (1990) and Govindan (1995).

[2] Kohlberg (1990) provides an overall view on the subject.

equilibria in later stages of the game. This is of course relevant for the equilibrium outcome. Which move is chosen by some player in early stages of the game may indeed reveal to others how he will move in later stages. Players may thus find it convenient to deviate from some equilibria if by doing that they are able to signal their own intentions to rivals, thereby affecting equilibrium (and payoffs) in later stages.

Forward induction arguments cannot of course be of any help in eliminating the usual multiplicity of equilibria in traditional, one-stage voting games. However, due to their multi-stage nature, forward induction can be relevant in refining citizen-candidate models. The reason is associated to the cost of candidacy. Loosely speaking, since being a candidate is costly, citizens would rationally run for election only if they can affect the final policy outcome. This mere fact may have major consequences in shaping the set of possible equilibria.

In the following we show, via an example, how the logic of forward induction can eliminate equilibria, and also outcomes for the whole game, that would be admitted by the solution concept proposed both by Besley and Coate (1997) and by Dhillon and Lockwood (1999) to solve citizen-candidate models. In the main text we give an intuitive argumentation, where, following Besley and Coate (1997) and Dhillon and Lockwood (1999), we limit the analysis to pure strategies in the voting games. In the appendix, allowing players to use mixed strategies in the voting stage too, we show that this intuitive argumentation is corroborated by the fact that Mertens stability eliminates the same policy outcome.

2 The model of representative democracy

Besley and Coate (1997) consider a community consisting of a set of citizens, $i \in \Im = \{1, 2, ..., I\}$, that, in order to implement a policy, must elect a representative among themselves. Citizens' utility may depend both on the chosen policy and on the identity of the policy-maker.

The selection of the community representatives requires an election. Each citizen is allowed to run for election, acting as a candidate. All citizens choosing to be a candidate face a utility cost δ. Each citizen disposes of one vote. The candidate who collects the most votes is elected (in case of ties, all tying candidates have the same probability of election). In other words, the plurality rule applies. Policy alternatives and preferences relative to each citizen are known to each other.

The political process consists of a three-stage game. In the first stage, each citizen $i \in \Im$ decides whether to become a candidate or not. In the second stage, the election occurs. In the third stage, the elected candidate chooses and implements the policy. Since candidates cannot credibly commit to any policy, each elected representative chooses his preferred policy, assumed to be unique. In order to simplify the analysis we eliminate this third stage, considering that the preferred policy of the elected candidate is automatically selected.[3] If no citizen decides to run for office then a "default" policy is selected.

Given the set of candidates $C \subset \Im$, each citizen decides whether to vote for a candidate or to abstain. A vector of voting decision is denoted by $\alpha = (\alpha_1, ..., \alpha_I)$, where $\alpha_j = i$ if citizen-voter j votes for citizen-candidate i, $i \in C$, and $\alpha_j = 0$ if citizen j abstains. Besley and Coate (1997) solve the voting stage with the concept of undominated pure Nash equilibrium. Hence, a voting equilibrium is a vector of voting decisions α^* such that, for each $j \in \Im$, α_j^* is a best response to α_{-j}^* and each α_j^* is an undominated action. Given the structure of the game and the above requirements on α^*, equilibria in the voting game are not unique if there are three or more candidates.

In the first stage, citizens decide non-cooperatively upon their own candidacy. Each citizen decides whether to become a candidate, incurring a cost δ, or not. Denoting by γ the strategy profile of the entry stage[4], a *political equilibrium* is defined in Besley and Coate (1997) by a couple $\{\gamma^*, \alpha(.)\}$ such that *i)* γ^* is an equilibrium in the entry stage associated with the function $\alpha(.)$, and *ii)* $\alpha(C)$ is a voting equilibrium for any $C \subset \Im$. In other words, a *political equilibrium* is a subgame perfect equilibrium such that an undominated pure Nash equilibrium is induced in every proper subgame.

We present in the following section an example in which an equilibrium that is admitted by the Besley and Coate (1997) solution does not conform to forward induction logic.

3 Applying forward induction: an example

There is a community consisting of four citizens, $\Im = \{1, ..., 4\}$. They must decide about three possible policies: a, b and c. Citizens are not directly affected by the identity of policy makers, and their preference order over possible policy alternatives is as follows:

$$\begin{array}{ll} 1: & a \succ b \succ c \\ 2: & a \succ c \succ b \\ 3: & b \succ a \succ c \\ 4: & c \succ a \succ b \end{array}$$

It is immediate that, if citizens vote directly on policies, a is the Condorcet winner.

We show that, for sufficiently small cost of candidacy (δ), there exists a political equilibrium as defined by Besley and Coate (1997) which induces policy

[3] Or, equivalently, we analyse the reduced game obtained by eliminating (not iteratively) dominated strategies.

[4] γ_i denotes the probability that citizen i becomes a candidate.

$\left(\frac{1}{2}b + \frac{1}{2}c\right)$. We also show that this equilibrium is a IWUPE as defined by Dhillon and Lockwood (1999), namely a political equilibrium whose voting equilibria are iteratively weakly undominated. We show, however, that this equilibrium is eliminated by a forward induction argument.

We make our point clear by first solving (with the concept of undominated pure Nash equilibrium) all the possible voting subgames.[5]

Candidates	Eq. policy	Eq. strategy profiles
1	a	$(1,1,1,1)$
2	a	$(2,2,2,2)$
3	b	$(3,3,3,3)$
4	c	$(4,4,4,4)$
$\{1,2\}$	a	*Every undominated strategy*
$\{1,3\}$	a	$(1,1,3,1)$
$\{1,4\}$	a	$(1,1,1,4)$
$\{2,3\}$	a	$(2,2,3,2)$
$\{2,4\}$	a	$(2,2,2,4)$
$\{3,4\}$	$\frac{1}{2}b + \frac{1}{2}c$	$(3,4,3,4)$
$\{1,2,3\}$	a	*Every undominated equilibrium*
$\{1,2,4\}$	a	*Every undominated equilibrium*
$\{1,3,4\}$	$\frac{1}{2}b + \frac{1}{2}c$	$(3,4,3,4)$
	a	$(1,1,3,1)\,,(1,1,3,4)\,,(1,1,1,4)\,,(1,1,1,1)$
$\{2,3,4\}$	$\frac{1}{2}b + \frac{1}{2}c$	$(3,4,3,4)$
	a	$(2,2,2,2)\,,(2,2,3,4)\,,(2,2,2,4)\,,(2,2,4,2)$
$\{1,2,3,4\}$	$\frac{1}{2}b + \frac{1}{2}c$	$(3,4,3,4)$
	a	*Every other undominated equilibrium*

It is to remark that the voting equilibrium $(3,4,3,4)$ that arises with $\{1,3,4\}$, $\{2,3,4\}$ and $\{1,2,3,4\}$ as set of candidates is always a strict equilibrium. Hence, it cannot be eliminated by any usual refinement *if applied to the voting subgame*.

At this point, it is straightforward to show that, provided the utility cost of candidacy δ is not too high, there exists a Besley and Coate political equilibrium with the set of candidates $C = \{3,4\}$ that induces $\left(\frac{1}{2}b + \frac{1}{2}c\right)$ as policy outcome. This equilibrium is supported only by the following function $\alpha'(.)$:

[5] In order to save space, we omit the description of the voting equilibria that induces policy outcome a in the subgames $\{1,2,3\}$, $\{1,2,4\}$ and $\{1,2,3,4\}$. The reader can easily compute them.

	voting strategies	*outcome*
$\alpha'(C)$	$(3,4,3,4)$	$\frac{1}{2}b+\frac{1}{2}c$
$\alpha'(C\cup 1)$	$(3,4,3,4)$	$\frac{1}{2}b+\frac{1}{2}c$
$\alpha'(C\cup 2)$	$(3,4,3,4)$	$\frac{1}{2}b+\frac{1}{2}c$
$\alpha'(C\setminus 3)$	$(4,4,4,4)$	c
$\alpha'(C\setminus 4)$	$(3,3,3,3)$	b

This equilibrium, however, is not compatible with the logic of forward induction. The reason, as will be clear, is that citizens 1 and 2 have an incentive to deviate in the first stage. By doing that, they are able to signal to the other citizens that they are going to vote for themselves.

Assume that the actual equilibrium is $(\gamma' = (0,0,1,1),\ \alpha'(.))$. There, if citizen 1 deviates by running for election, then citizen 3 should, in the voting stage, rationally think that 1 has done so in order to gain a payoff that is at least equal to the one he gets by staying out (i.e., his payoff associated to $\frac{1}{2}b+\frac{1}{2}c$). May citizen 1 hope to achieve a payoff strictly greater than the one he gets in $\frac{1}{2}b+\frac{1}{2}c$ by running for election? Of course yes, because citizen 1 gains the payoffs associated to a (minus δ) in all the other equilibria of the voting subgame with the $\{1,3,4\}$ as set of candidates. In such cases, however, citizen 1 must vote for himself, and not for 3. It follows that citizen 3 after having observed that 1 is candidate, has to deduce that citizen 1 votes for himself. But if this is the case, then citizen 3 must also recognize that he has no chance of being elected (if the opponents do not use dominated strategies), and that his own best voting strategy is not anymore 3, but 1. The same occurs for citizen 2: he should also vote 1 after a deviation on the part of citizen 1. It follows, by the logic of forward induction, that it is not rational for citizen 1 not to run for election. He anticipates that by doing so he will be a sure winner. The same argument applies to a possible deviation of citizen 2 in the first stage. Again, 4 conjectures that 2 is going to vote for himself and this induces a change in his own voting behavior: he would prefer to vote for 2. Therefore, policy $\frac{1}{2}b+\frac{1}{2}c$ is not compatible with forward induction.

The above example shows that the solution concepts proposed by Besley and Coate (1997) and by Dhillon and Lockwood (1999) fail to capture forward induction rationality. Imposing restrictions on voting behavior (such as undominance, iterated undominance or strategic stability) does not suffice for overcoming this lack. In the appendix we show how the intuitive argumentation given above is corroborated by the fact that Mertens' stability eliminates policy $\frac{1}{2}b+\frac{1}{2}c$ as a solution outcome.

4 Conclusions

In this note we have shown how the solution concept adopted in Besley and Coate (1997) model of representative democracy can be refined through the logic of forward induction. The emergence of forward induction arguments is deeply rooted in the very structure of citizen-candidate models: the same agents are called to act in different stages of the game and the choice of running as a candidate involves an utility cost. Under such conditions players may have an incentive to deviate from a prescribed equilibrium with the aim of signalling to others their subsequent behavior, thus changing the equilibrium induced in some subgame.

Refining the political equilibrium proposed by Besley and Coate in the voting stage only is not sufficient to guarantee solution conform to forward induction logic. We have shown how such a criterium may eliminate equilibria and outcomes in the citizen-candidate model also when voting equilibria are assumed to be strict, and, hence, iteratively undominated (as proposed by Dhillon and Lockwood (1999)).

In a previous note (De Sinopoli and Turrini (1999)), it is shown that the solution concept to be adopted to solve this citizen-candidate model must be invariant after the iterated elimination of weakly dominated strategies in the voting stage. The results about the Besley and Coate model of representative democracy presented in that paper and in the present one, together with the results on plurality games presented in De Sinopoli (1999) lead us to the conclusion that, in order to capture strategic behavior in this model, we need the full strength of the Mertens' stability concept.

References

[1] Besley, T. and S. Coate (1997), "An economic model of representative democracy", *Quarterly Journal of Economics*, 112, 85-114.

[2] Cho, I. and D. Kreps, (1987), "Signaling games and stable equilibria", *Quarterly Journal of Economics*, 102, 2, 179-221.

[3] De Sinopoli, F. (1999), "Sophisticated voting and equilibrium refinements under plurality rule" *forthcoming Social Choice and Welfare*.

[4] De Sinopoli, F. and A. Turrini, (1999), "A remark on voters' rationality in Besley and Coate model of representative democracy", CORE Discussion Paper n. 9927

[5] Dhillon, A. and B. Lockwood, (1999), "When are plurality rule voting games dominance solvable?", mimeo, University of Warwick.

[6] Govindan, S. (1995), "Stability and the chain store paradox", *Journal of Economic Theory*, 66, 536-547.

[7] Kohlberg, E. (1981), "Some problems with the concept of perfect equilibrium", in NBER Conference on Theory of General Equilibrium, University of California, Berkeley.

[8] Kohlberg, E. (1990), "Refinements of Nash equilibrium: the main ideas", in T. Ichiishi, A. Neyman and Y. Tauman Eds. *Game Theory and Applications*. Academic Press, NY.

[9] Kohlberg, E., and J. F. Mertens, (1986), "On the strategic stability of equilibria", *Econometrica*, 50, 863-894.

[10] Mertens, J.F. (1989), "Stable Equilibria-A reformulation. Part I: Definition and basic properties", *Mathematics and Operations Research*, 4, 575-625.

[11] Osborne, M. (1990), "Signalling, forward induction, and stability in finitely repeated games", *Journal of Economic Theory*, 50, 22-36.

[12] Osborne, M. and A. Slivinski, (1996), "A model of political competition with citizen candidates", *Quarterly Journal of Economics*, 111, 65-96.

[13] van Damme, E. (1989), "Stable equilibria and forward induction", *Journal of Economic Theory*, 49, 476-496.

Appendix

In this appendix we show how strategic stability (Mertens (1989)) eliminates the policy outcome $\left(\frac{1}{2}b+\frac{1}{2}c\right)$ in the example discussed in section 3. In order to apply strategic stability, we must allow the players to use mixed strategies in the voting stage too. This obviously implies that the intuitive analysis developed in section 3 cannot be applied to all games with the same preference order, since other voting equilibria may appear.

As for the analysis of mixed strategies a full specification of payoffs is needed, let us consider the following:[6]

$$\begin{aligned}
u_1(a) &= 2, \quad u_1(b)=1, \quad u_1(c)=0, \\
u_2(a) &= 2, \quad u_2(b)=0, \quad u_2(c)=1, \\
u_3(a) &= 1, \quad u_3(b)=2, \quad u_3(c)=0, \\
u_4(a) &= 1, \quad u_4(b)=0, \quad u_4(c)=2, \\
\frac{3-\sqrt{5}}{4} &< \delta < \frac{1}{2}.
\end{aligned}$$

It is immediate to verify that, for the above specification of payoffs, $(\gamma', \alpha'(.))$ (as described in section 3) is a *political equilibrium* with policy outcome $\left(\frac{1}{2}b+\frac{1}{2}c\right)$. We show how this outcome is excluded by Mertens' stable sets (i.e., no element of a stable set induces policy $\left(\frac{1}{2}b+\frac{1}{2}c\right)$).

The main idea of the proof is to find the set $(\hat{\Sigma})$ of undominated Nash equilibria of the reduced normal form representation that induces the policy outcome $\left(\frac{1}{2}b+\frac{1}{2}c\right)$. We then prove that $\hat{\Sigma}$ is disconnected from its complement in the set of undominated equilibria and we show that the normal form strategies that prescribe to player 1 to run as a candidate in the first stage and to vote for 3 in the subgame $\{1,3,4\}$ are strictly inferior responses to every element of $\hat{\Sigma}$. Finally, we note how after eliminating these strategies (as well as any dominated strategy), no undominated equilibrium belongs to $\hat{\Sigma}$. This proves that $\hat{\Sigma}$ does not contain any element of a stable set.

We prove those points using Kuhn's equivalent behavioral strategies. Notice that every undominated strategy of the reduced normal form has to prescribe an undominated behavior in each subgame (reached with positive probability under some strategy of the opponents). This simple consideration drastically simplifies the analysis, since it implies that we have to study only the candidacy stage and the subgames with three or more candidates.

We start the study of the set $\hat{\Sigma}$, by characterizing the prescribed behavior in the "candidacy" stage.

As in Besley and Coate (1997), we denote by γ_i the probability that player i is a candidate. We obtain the following Lemma.

[6] The utility can be perturbed to make the game generic, without affecting the results, cf footnote 8.

Lemma 1 *If σ is an undominated equilibrium with policy outcome $\left(\frac{1}{2}b+\frac{1}{2}c\right)$ then:*
α) $(\gamma_3 \mid \sigma) = (\gamma_4 \mid \sigma) = 1$
β) $(\gamma_1 \mid \sigma) = (\gamma_2 \mid \sigma) = 0$

Proof. To simplify the notation we omit the dependence of the γ's from σ.

α) By contradiction, we consider two cases:
i) $\gamma_1+\gamma_2 > 0$. If $\gamma_3 \cdot \gamma_4 \neq 1$, then a subgame where policy a is selected by every undominated strategy combination is reached with positive probability.
ii) $\gamma_1+\gamma_2 = 0$. If $\gamma_3, \gamma_4 \neq 1$, then the default outcome (i.e., no candidate) is selected with positive probability. If $\gamma_3 = 1, \gamma_4 < 1$ then, for every undominated strategy combination, the policy outcome is $(1-\gamma_4)\,b + \gamma_4\left(\frac{1}{2}b+\frac{1}{2}c\right) \neq \left(\frac{1}{2}b+\frac{1}{2}c\right)$. Analogously for the case $\gamma_3 < 1, \gamma_4 = 1$.

β) We remark that every undominated equilibrium induces an undominated equilibrium in every subgame reached with positive probability and the preference's orders over policies implies that, in the subgames $\{1,3,4\}, \{2,3,4\}$, $\{1,2,3,4\}$, there is only one undominated equilibrium outcome with policy a selected with probability 0, namely $\left(\frac{1}{2}b+\frac{1}{2}c\right)$. This simple considerations and (α) implies there exists no undominated equilibrium with policy outcome $\left(\frac{1}{2}b+\frac{1}{2}c\right)$ in the following cases:
i) $0 < \gamma_1, \gamma_2 < 1$ (by contradiction, since, if σ is an undominated equilibrium with policy outcome $\left(\frac{1}{2}b+\frac{1}{2}c\right)$, the policy outcome induced in $\{1,3,4\}, \{2,3,4\}$, $\{1,2,3,4\}$ and in $\{3,4\}$ is the same, hence players 1 and 2 would prefer not running, saving δ)
ii) $\gamma_1 > 0, \gamma_2 = 0$ (i.e., since the policy is the same in $\{1,3,4\}$ as in $\{3,4\}$, player 1 would prefer not running)
ii') $\gamma_1 = 0, \gamma_2 > 0$
iii) $0 < \gamma_1 < 1, \gamma_2 = 1$ (since the policy is the same in $\{2,3,4\}$ as in $\{1,2,3,4\}$, player 1 would prefer not running)
iii') $\gamma_1 = 1, 0 < \gamma_2 < 1$.

Hence, in order to prove the lemma we only need to exclude the case $\gamma_1 = \gamma_2 = 1$. In this case we can prove that player 1 (resp. 2) would prefer to not running. It is not difficult to see that, in the subgame $\{2,3,4\}$, for every undominated strategy of the opponents, player 1 can guarantee himself a payoff of $\frac{1}{3}$, then strictly greater than $\frac{1}{2}-\delta$.[7] ■

Lemma 1 fully characterizes the behavior prescribed in the "candidacy" stage by every undominated equilibrium with policy outcome $\left(\frac{1}{2}b+\frac{1}{2}c\right)$. It is immediate to verify that such behavior is best reply for player 3 and player 4 to every undominated strategy vector with such a probability of candidacy. Lemma 1 implies that if player 1 becomes a candidate the subgame that is reached is, with probability one, $\{1,3,4\}$. Hence, every undominated equilibrium σ with

[7] In fact, if the opponents play undominated strategy in $\{2,3,4\}$, the payoff of player 1 is decreasing in the probabilities that players 2 and 4 vote for 4. Setting these probabilities equal to one, we find that the minimum player 1 can guarantee himself is when player 3 votes for 2 with probability $\frac{1}{3}$. Hence the result. We remark that this is the only step of the proof where the preference's orders over policies do not suffice.

policy outcome $\left(\frac{1}{2}b+\frac{1}{2}c\right)$ has to satisfy the following constraints:

$$\begin{aligned} i)\ U_1(E1|\sigma) &\leq \frac{1}{2} \\ ii)\ U_1(E3|\sigma) &\leq \frac{1}{2} \end{aligned}$$

where $E1$ denotes any normal form strategy of player 1 that prescribes "being a candidate and voting for himself in the subgame $\{1,3,4\}$", while $E3$ denotes any strategy of player 1 that prescribes "being a candidate and voting for 3 in the subgame $\{1,3,4\}$".

Lemma 2 *The strategy $E3$ is, for player 1, a (strictly) inferior reply to any undominated equilibrium σ with policy outcome $\left(\frac{1}{2}b+\frac{1}{2}c\right)$. The analogous holds for player 2.*

Proof. Let us concentrate on undominated strategies in the subgame $C=\{1,3,4\}$ and let:

$$\begin{aligned} x &= prob\left(\alpha_2\left(C\right)=1\right) \qquad (1-x)=prob\left(\alpha_2\left(C\right)=4\right) \\ y &= prob\left(\alpha_3\left(C\right)=1\right) \qquad (1-y)=prob\left(\alpha_3\left(C\right)=3\right) \\ z &= prob\left(\alpha_4\left(C\right)=1\right) \qquad (1-z)=prob\left(\alpha_4\left(C\right)=4\right) \end{aligned}$$

The first step of the proof consists in showing that if $U_1(E3\mid\sigma)\geq U_1(E1\mid\sigma)$ then:

$$x+z-\frac{4}{3}xz\leq\frac{1}{3}.$$

First of all, we notice that if σ is an undominated equilibrium with policy outcome $\left(\frac{1}{2}b+\frac{1}{2}c\right)$ then $y\neq 1$. This is obvious since, if $y=1$ the payoff of $E1$, under every undominated strategy combinations, is at least equal to $1-\delta$, i.e., strictly greater than $\frac{1}{2}$. The remainder of the Lemma is proved by direct computation. We observe that

$$\begin{aligned} U_1(E1\mid\sigma)-U_1(E3\mid\sigma) &= \tfrac{1}{2}xz(1-y)+(1-y)\left[x(1-z)+z(1-x)\right]+ \\ &\quad +y(1-x)(1-z)-\tfrac{1}{2}(1-x)(1-y)(1-z)\geq \\ &\geq \tfrac{1}{2}(1-y)\left[xz+2x(1-z)+2z(1-x)-(1-x)(1-z)\right]= \\ &= \tfrac{1}{2}(1-y)(3x+3z-4xz-1). \end{aligned}$$

Hence, $U_1(E1\mid\sigma)-U_1(E3\mid\sigma)$ is strictly positive strictly positive if $x+z-\frac{4}{3}xz>\frac{1}{3}$.

Now we can prove the claim of the Lemma. In fact, if $E3$ is not a (strictly) inferior reply, then $U_1(E3\mid\sigma)\geq U_1(E1\mid\sigma)$ and, therefore, $x+z-\frac{4}{3}xz\leq\frac{1}{3}$. Direct computations yield

$$\begin{aligned} U_1(E3\mid\sigma) &= 2xyz+2y\left[z(1-x)+x(1-z)\right]+ \\ &\quad +\tfrac{3}{2}xz(1-y)+(1-y)\left[x(1-z)+z(1-x)\right]+\tfrac{1}{2}(1-x)(1-z)(1-y)-\delta= \\ &= \tfrac{1}{2}xz(1-y)+\tfrac{1}{2}(1-y)+(x+z-\tfrac{4}{3}xz)(\tfrac{1}{2}+\tfrac{3}{2}y)+\tfrac{1}{3}xz(\tfrac{1}{2}+\tfrac{3}{2}y)-\delta. \end{aligned}$$

Hence, for $x+z-\frac{4}{3}xz \leq \frac{1}{3}$, we have[8]
$U_1(E3 \mid \sigma) \leq \frac{1}{2}(1-y)(1+xz)+\frac{1}{3}(\frac{1}{2}+\frac{3}{2}y)+\frac{1}{3}xz(\frac{1}{2}+\frac{3}{2}y)-\delta = \frac{2}{3}(xz+1)-\delta$.
It is easy to see that the constraint $x+z-\frac{4}{3}xz \leq \frac{1}{3}$ implies $xz \leq \left(\frac{3-\sqrt{5}}{4}\right)^2$, and, hence, $U_1(E3 \mid \sigma) < \frac{1}{2}$.[9] By symmetry the analogous holds for player 2. ■

Denote by Σ' the space of undominated strategies. We have shown that the set of undominated equilibria with policy outcome $\left(\frac{1}{2}b+\frac{1}{2}c\right)$ is described by

$$\hat{\Sigma} = \left\{ \begin{array}{c} \sigma \in \Sigma' : (\gamma_1 \mid \sigma) = (\gamma_2 \mid \sigma) = 0, (\gamma_3 \mid \sigma) = (\gamma_4 \mid \sigma) = 1, \\ U_1(E1 \mid \sigma) \leq \frac{1}{2}, U_2(E2 \mid \sigma) \leq \frac{1}{2} \end{array} \right\}$$

where $E2$ denotes player 2's strategy "being a candidate and voting for himself in the subgame $\{2,3,4\}$".

It is easy to see that the set $\hat{\Sigma}$ is not connected to any other undominated equilibrium component of the game. In fact players 3 and 4 strictly prefer being a candidate than not, hence this is true for every sufficiently close undominated strategy combination of the opponents. Moreover, $E3$ is not a best reply for player 1 to any sufficiently close undominated strategy combination of the opponents. Furthermore, necessary condition to have an undominated equilibrium nearby $\hat{\Sigma}$, where player 1 "runs for office and votes for himself in $\{1,3,4\}$" with positive probability is that the strategy combination $(1,x,y,z)$, is an undominated equilibrium of the subgame $\{1,3,4\}$. It is easy to see that in all the undominated equilibria of this subgame where player 1 votes for himself he is surely elected (i.e., player 2 can guarantee 1's election by voting for him). Hence player 1 could guarantee himself, by $E1$, a payoff close to $2-\delta$. A symmetric reasoning applies to player 2. Hence nearby $\hat{\Sigma}$, there is no undominated equilibrium with different "candidacy" strategies. Finally it is immediate to realize that, nearby $\hat{\Sigma}$, there is no undominated equilibrium with different strategies in the voting stage since if $U_1(E1 \mid \sigma) > \frac{1}{2}$ (resp. $U_2(E2 \mid \sigma) > \frac{1}{2}$) player 1 (resp.

[8] This analysis does not depend on the exact payoffs. Consider the following "generic" specification for player 1's utility:
$u_1(a) = 2, u_1(b) = \alpha, u_1(c) = 0,\ \frac{\alpha^2}{4+2\alpha} < \delta < \frac{2-\alpha}{2}$ with $0 < \alpha < \sqrt{2}$
If the other players have the same preference's orders over policies we can prove the same results of Lemma 1. In this case, in the subgame $\{2,3,4\}$ player 1 can, at least, guarantee himself a payoff of $\frac{\alpha}{2+\alpha}$. It is easy to see that we can prove that if σ is an undominated equilibrium with policy outcome $\left(\frac{1}{2}b+\frac{1}{2}c\right)$ and $U_1(E3 \mid \sigma) \geq U_1(E1 \mid \sigma)$ then:
$x+z-\frac{6-2\alpha}{4-\alpha}xz \leq \frac{\alpha}{4-\alpha}$.
Moreover:
$U_1(E3 \mid \sigma) = \left[x+z-\frac{6-2\alpha}{4-\alpha}xz\right]\left[2y+\frac{\alpha(1-y)}{2}\right]+\frac{2-\alpha}{2}(1-y)xz+\frac{\alpha}{2}(1-y)+$
$\quad +\frac{2-\alpha}{4-\alpha}xz\left[2y+\frac{\alpha(1-y)}{2}\right]-\delta$
hence for $x+z-\frac{6-2\alpha}{4-\alpha}xz \leq \frac{\alpha}{4-\alpha}$
$U_1(E3 \mid \sigma) \leq \frac{\alpha}{4-\alpha}\left[2y+\frac{\alpha(1-y)}{2}\right]+\frac{2-\alpha}{2}(1-y)xz+\frac{\alpha}{2}(1-y)+\frac{2-\alpha}{4-\alpha}xz\left[2y+\frac{\alpha(1-y)}{2}\right]-\delta =$
$\quad = \frac{4\alpha+(8-4\alpha)xz}{8-2\alpha}-\delta$.
[9] The maximum of xz such that $0 \leq x,z \leq 1$ and $x+z-\frac{4}{3}xz \leq \frac{1}{3}$ is obviously attained in $x=z$ and $2x-\frac{4}{3}x^2=\frac{1}{3}$.

2) would strictly prefer to run for office. This completes the proof of the next Proposition.

Proposition 3 $\hat{\Sigma}$ *is the set of undominated equilibria with policy outcome* $\frac{1}{2}b + \frac{1}{2}c$*. This set is not connected to any other undominated equilibrium component and*

$$U_1(E3 \mid \hat{\sigma}) < U_1(\hat{\sigma}) \quad \forall \hat{\sigma} \in \hat{\Sigma}.$$

To show that no element of a stable set inducing the policy outcome $\left(\frac{1}{2}b + \frac{1}{2}c\right)$ exists, it is enough to consider the following three properties of this solution concept:

i) Stable sets always exist.

ii) Stable sets are connected sets of normal form perfect (hence undominated) equilibria.

iii) A stable set contains a stable set of every game obtained by eliminating a strategy that is used with minimum probability in any $\varepsilon-$*perfect* equilibrium close to it.

By property (*ii*) and the above Proposition we know that if some elements of a stable set induces policy outcome $\left(\frac{1}{2}b + \frac{1}{2}c\right)$, then a stable set contained in $\hat{\Sigma}$ exists. But since $\forall \hat{\sigma} \in \hat{\Sigma}$, $U_1(E3 \mid \hat{\sigma}) < U_1(\hat{\sigma})$ we have that $E3$ is used with minimum probability in any $\varepsilon-$*perfect* equilibrium close to $\hat{\Sigma}$. If we eliminate all the normal form strategies prescribing $E3$, as well as all the other dominated strategies, we have that voting for 1 in the subgame $\{1, 3, 4\}$ is dominant for players 2 and 3. It follows that there is no undominated equilibrium of this reduced game contained in $\hat{\Sigma}$, hence $\hat{\Sigma}$ contains no stable set.

Remark: The above analysis considers undominated equilibria. We have constructed the example fixing the utility associated to policies and finding sufficient conditions on δ such that when player 1 prefers $E3$ to $E1$ he is strictly better off if not running for election. We think that these results can be improved (in the sense that a larger interval for δ can be found) considering perfect equilibria. A preliminary cost-benefit analysis discouraged us from such a project.