Campaign Rhetoric and the Hide-&-Seek Game

Sourav Bhattacharya*

Department of Economics, University of Pittsburgh

sourav@pitt.edu

January 7, 2009

Abstract

I examine the choice of political rhetoric when one candidate's willingness to misinform the voter is checked by the other's ability to inform. In the basic model, a debate is a competition between two candidates in which the "good" candidate wants to reveal information about quality and the "bad" candidate wants to avoid such revelation. However, the quality of the rival candidate can only be discovered through costly research. Therefore, debate takes the form of a hide and seek game under incomplete information. The main result is that although negative advertising and slander go down as the prior probability of a candidate being good increases; voter welfare (probability of correct selection) is not monotonic in the prior or the cost of search. All results extend to a model where a continuum of real numbers represents candidate quality, with the additional feature that the definition of "good" and "bad" are determined endogenously.

1 Introduction

Electoral Campaigns are an extremely important source of information for the voter. Candidates choose their campaign messages strategically in order to influence voter perception, and each candidate prefers revelation of information that is favourable to him and/or harmful to the rival. Information provision in electoral campaigns is thus a competitve process involving many aspects of strategic choice (e.g. choice of the issues to be highlighted, the target audience to be focussed

*I thank David Austen-Smith, Allen Brierly, Steve Callander, Joyee Deb, Alexandre Debs, Tim Feddersen, Yuk-Fai Fong, Sean Gailmard, Bard Harstad, Jaehoon Kim, Tapas Kundu and Marciano Siniscalchi for comments and suggestions. Any error that remains in the paper is my responsibility.

on, the extent and allocation of advertisement expenditiures etc). In this paper, I focus on the candidates' choice between positive or negative advertising: whether to highlight one's own quality or to focus on why the rival is unsuitable for office.

Electoral campaigns involve a wide variety of messages: arguments and counter-arguments both tenable and non-tenable, attacks both truthful and slanderous. Existing microeconomic theories of campaign advertising cannot explain the richness of truth, half-truths and lies that are observed in campaigns: while one class of theories considers statements as verifiable and therefore true by definition (see Tirole 1988, chapter 2 for a survey), another class deems advertisements as an exercise in money-burning (Milgrom-Roberts 1986), and thus renders the content of the messages irrelevant[1]. The current paper develops a framework that explains the effect of candidate quality on campaign rhetoric while allowing for both truth and lies. The model postulates that the willingness of a candidate to misinform the voter is checked by his rival's ability to inform, but the power of such a check is limited by how much each candidate knows about the other. Thus, the content of the electoral rhetoric is determined by the interaction of two levels of incomplete information: one between the candidates and the voter, and the other between the two candidates. I study this interaction by modeling electoral competition as a public debate between two candidates with private information about their own quality.

The analysis in the paper has close links with the debate in political science on negative advertising and its role in the democratic process. While some scholars like Ansolabehere and Iyengar (1995) have equated negative advertising with slander and claimed that it reduces the quality of democratic debate, others have pointed out its information-providing role (Polborn and Yi 2006). The contribution of this paper is to show that negative advertising can perform both roles (providing and concealing information) and to demonstrate how these roles interact in equilibrium.

In this model, candidates are initially uninformed about each other's quality, and voters are uninformed about the candidates. Candidates have two actions: first they have an option to privately invest in acquiring information about their rival, and then they use their information to launch a political campaign. The campaign themes are chosen simultaneously and are modelled as a public debate. I model a positive campaign as one in which a candidate provides arguments

[1] See Prat (2004) for a detailed survey of each of these two strands of literature as applied to the special case of political advertising.

in support of his own candidature and a negative campaign is one where a candidate provides reasons why the rival is not a good choice[2]. Given the campaign messages, the actual process of voter opinion formation is blackboxed in what I call the *information revelation protocol* of the debate. I assume that if one candidate employs a positive message and the other a negative one, then there is a *fruitful debate* about the former candidate's quality, and on the other hand, if both candidates use positive messages or both use negative messages, then they are not engaging in arguments and counter-arguments about any one issue: such a situation is termed *cross talk*. The protocol embeds two postulates: first, a fruitful debate reveals more information to the voters about the "focal" candidate whose quality is being debated (as compared to the other candidate), and second, cross talk is less informative than fruitful debate in that there is no focal candidate (or focal issue). Taking the protocol as given, I analyse the incentives for candidates to choose a positive or a negative campaign.

If a good candidate is facing a bad candidate (and if they know each other's types), the former would prefer revelation of information about either candidate through fruitful debate, while the latter would prefer cross talk which would avoid such revelation. Thus, a good candidate wants to counter the rival's argument while the bad candidate wants to mimic the rival's statement, leading to a matching pennies game between the good and bad types. Since the bad candidate wants to hide (conceal information) and the good candidate wants to chase (reveal information), I title the competitive process of information provision as a *Hide and Seek game*.

In the game theoretic literature, the term Hide and Seek game refers to the matching pennies game where the payoffs to the players may be asymmetric. This game comes up not just in the context of political campaigns, but also in commercial advertising, military strategy, market entry, product design and in various other important social, political and economic situations. What makes the game studied in this paper different is the assumption of asymmetric information between candidates and the option of costly search.

The assumption of costly information is crucial from the substantive standpoint too. While information in a positive campaign is free for a candidate, truth in a negative campaign is costly to him. This creates an endogenous distinction between the incentives for positive and negative

[2] Although in reality candidates always use both kinds of messages, this paper concerns with choice of the broader campaign theme which, for our purposes, is either positive or negative.

advertising. If candidates know each other's types, as has been assumed in Polborn and Yi (2006), the incentives to send a positive message or a negative message are exactly symmetric, depending only on the realised types. In the absence of incomplete information between candidates, there is virtually no substantive difference between positive and negative advertising.

The innovation of two-layered incomplete information and search as an option also allows us to explicitly distinguish between truthful negative campaigns which are beneficial for the voter and false attacks or slander which are not. *Slander* in this framework is modeled as a negative message without prior research, and slander against a good candidate is always a lie. Also, due to incomplete information between candidates, the prior beliefs about types plays a major role in determining the nature of the debate. Taking the prior belief as a proxy for average candidate quality, comparative static conclusions can be drawn about the effect of the average (expected) quality in the candidate pool on the information content and nature of the rhetoric.

1.1 Main results

The paper focusses on understanding both the strategic behavior of candidates as a function of the average (expected) candidate quality and the welfare implications of such behavior. The model confirms the conventional wisdom that the incidence of negative advertising and slander increase as the average quality worsens. With lower expected quality, the bad candidates use more negative advertising to deflect attention, especially because the rival is expected to be bad too. However, negative messages have a beneficial role too in that they facilitate information transmission to voters through fruitful debate. In particular, a good candidate can most effectively reveal his type through positive advertising when the average candidate quality is low and there is a lot of negative campaigning. At the other extreme, when the average candidate quality is very high, all candidates engage is more positive campaigning: but this allows the bad candidates to effectively "hide" by avoiding debate with the good types. Thus, while an electorate with low expected candidate quality is competitive but efficient for the purpose of candidate selection, one with a high average quality is conservative and inefficient. When average candidate quality is moderate, the good candidate sometimes uses negative campaigns to expose the bad type, but debate is not as effective. Consequently, voter welfare in terms of the ex-post average quality of the winning candidate may actually *go down* with an increase in the ex-ante expected candidate quality.

The information content of the negative campaigns depends on the propensity to engage in costly search about the rival's type. Search is probabilistic, undertaken only when neither type is too rare, and the bad type searches with a strictly higher frequency than the good type. Since the good type has better information about herself than the bad, the good type cannot commit to searching more frequently than the bad type. This leads to what I term an *inefficiency due to search*. It is worth noting that although the search cost can be seen as an index of the extent of asymmetry in information between the candidates, the welfare loss to the voter is not monotonic in this cost. Similarly, the voter welfare is not monotonic in the payoff from winning office.

Polborn and Yi (2006, Proposition 4) derive the empirical prediction that there is a positive correlation between the *incidence* of negative advertising and the probability of victory conditional on going negative. This is due to their assumption that negative advertising is always informative. Taking into account the harmful role of negative advertising, I obtain an overall negative relationship: negative advertising is only sometimes used by the good type to expose the bad candidate, but mostly used by the bad type to "muddle" the debate. However, the point of this paper is that both the incidence and efficacy of negative advertising depends on the many other factors like voter expectations of candidate quality, extent of information asymmetry and the entire profile of messages employed in a debate.

Several experimental and empirical studies (see Lau, Siegelman, Heldman and Babbitt (1999) for a detailed survey) seem to suggest an empirical regularity: while positive advertising increases support for a candidate and negative advertising reduces support for the opponent, negative advertising might hurt the sponsor himself. Earlier formal work on choice of message (Skaperdas and Grofman 1995, Harrington and Hess 1996) assumes an influence function with these effects. In this paper, these effects arise as features of the equilibrium. Further, this paper points out the possibility that the effect of a candidate's message on the voter may depend on the rival's message, and this has implications for experiments to be undertaken in this field.

The remainder of the paper is organized as follows. Section 2 sets up the benchmark model in which the candidates are either "good" or "bad" and section 3 characterises the equilibrium. Section 4 discusses how candidate behavior in the benchmark case changes as the prior belief and the search cost change, and analyses welfare implications. Section 5 presents a few extensions and Section 6 concludes the main body of the paper. Appendix A discusses an extended version of the model with voter inference and a more general type space. Most proofs are in Appendix B.

2 Basic Model: Binary Types

There are two players: candidates 1 and 2 with a private quality $\theta_i, (i = 1, 2)$ which can be either Good (G) and Bad (B)[3]. For both candidates, quality follows a commonly known Bernoulli distribution with the prior probability of a good type $\alpha \in (0, 1)$[4]. The candidates run campaigns (positive or negative) which reveal information to a voter about quality based on which she votes for one of the two. While the voter is not modelled explicitly as a player, her actions based on information revealed through messages are taken into account in the payoffs arising from candidate actions.

2.1 Actions

There are two actions chosen by player i: the debate or message action $M_i \in \{P, N\}$ and the search action $X_i \in \{S, NS\}$. P denotes a positive campaign message and N denotes a negative message. If player i undertakes action S (search), he gets to know the type of his rival $-i$ with certainty. If action NS (no search) is taken, the rival's type is not known. One can think of the search action as being taken before the message action, so that the message can be conditioned on the information obtained through search. But since search itself is private, the search stage and the debate stage can be considered simultaneous. Speeches are also strategically simultaneous: a candidate cannot condition his speech on the message of the other candidate[5]. The message profile is denoted by $\mathbf{M} = \{M_1, M_2\}$, the search profile by $\mathbf{X} = \{X_1, X_2\}$ and the type profile by $\boldsymbol{\theta} = \{\theta_1, \theta_2\}$. The payoff to player i from an action and type profile $\{\mathbf{M}, \mathbf{X}, \boldsymbol{\theta}\}$ is assumed to be the payoff $u_i(\mathbf{M}, \boldsymbol{\theta})$ from debate less the cost of search, which is $c \in (0, \frac{1}{2})$ if $X_i = S$ and zero otherwise. Note that search affects the payoff from debate only by affecting the private information available to a player.

[3] The "quality" of a candidate can be thought of in two ways. In a common values framework where all voters have the same preference, quality captures all characteristics that voters care about. In a private values framework where the two candidates can be assigned two locations on the left-right ideological continuum, quality can simply be thought of as the distance of a candidate from the median voter's ideal point.

[4] In section 5.4, I relax the common prior assumption.

[5] This assumption reflects the fact that most of the time, the campaign strategy of each candidate is decided before campaigning actually starts.

2.2 Information Revelation Protocol

The voter (judge) selects the winner in the debate after listening to the messages. The actual process of debate is assumed as a protocol which determines how information about candidate types is revealed to the judge depending on the message profile. A debate is called *fruitful* if the messages are different, i.e. if $M_1 \neq M_2$. I assume that a fruitful debate fully reveals the type of the candidate whose quality is subject to debate, i.e. the candidate using P. If there is *cross talk*, i.e. if $M_1 = M_2$, no new information is revealed, presumably because the candidates do not coordinate on discussing an issue. The rather extreme assumption that a fruitful debate reveals the true type of the candidate under focus while cross talk reveals nothing is not necessary for the qualitative results of the model to go through. It has been made for technical convenience so as to be able to drive the basic point home without using unnecessary parameters. In section 5.1 I discuss the results in a generalised model where this extreme assumption is relaxed.

In a fruitful debate only one candidate's type is revealed. A candidate found to be good wins and a candidate revealed to be bad loses. If there is cross talk, the winner is chosen randomly with equal probability. Although simplistic, such a passive voter response to debate works for a two-type case because if one candidate is revealed to be good (bad), the other candidate cannot be better (worse). In Appendix A, I show that if we inlcude the voter as a rational player in the game, such "passive" voting strategies considered here arise as equilibrium behavior, and the equilibrium remains unchanged.

Assume that the winner of the debate gets a payoff normalised to 2. The utility $u_i(\mathbf{M}, \boldsymbol{\theta})$ from debate is thus given by:

$$u_i(P, P, \boldsymbol{\theta}) = u_i(N, N, \boldsymbol{\theta}) = 1,$$

$$u_i(P, N, G, \theta_{-i}) = u_i(N, P, \theta_i, B) = 2,$$

$$u_i(P, N, B, \theta_{-i}) = u_i(N, P, \theta_i, G) = 0,$$

Obviously, the utility from debate satisfies $u_1(\mathbf{M}, \boldsymbol{\theta}) + u_2(\mathbf{M}, \boldsymbol{\theta}) = 2$.

2.3 Incentives

To understand the incentives that such a payoff structure creates, it is useful to describe them in a normal form. There are three possible situations, depending on whether a good type faces a good type, a bad type faces another bad type, or whether the two candidates are of different types. Each case is described in a separate 2×2 matrix, with the row player's payoff shown in the matrix.

G-type \ B-type	P	N
P	1	2
N	2	1

G-type \ G-type	P	N
P	1	2
N	0	1

B-type \ B-type	P	N
P	1	0
N	2	1

Suppose for now the candidates knew each other's types. Then they would know which of the above three situations they were in. If a good candidate were facing another good candidate, then it is strictly dominant for both to employ a positive message: both players would want to discuss their own qualities. If a bad candidate were facing another bad one, then again, it is strictly dominant to use a negative message: neither player wants the focus on himself. When a good candidate faces a bad candidate, there is a matching pennies game: the good candidate wants a fruitful debate either about his own quality (P, N) or about the rival's (N, P). The bad candidate, on the other hand, wants to induce cross talk by mimicking the good type's message. What makes this game interesting is that there is incomplete information about the rival's type, and each candidate has an incentive to invest in research about the rival's quality.

In what follows, some of the assumptions underlying the structure of the model are discussed.

2.4 Conditional verifiability of messages

The debate protocol embodies the assumption that information contained in campaign messages is neither hard (completely verifiable) nor soft (cheap talk), but *conditionally verifiable*: the veracity of a statement can be better ascertained if it is actualy debated, i.e. a counterpoint is offered. The model recognises the fact that voters often make judgements about the quality of candidates by comparing the arguments and counter-arguments made during the campaign process, in effect making inferences about candidate types from the *profile* of messages rather than from the individual messages independently, as is common in the economics and politics literature that treats individual messages as signals of candidate types.

2.5 Deterministic information revelation

In the benchmark model I assume that if a debate is fruitful in the sense defined above, the type of the "focal" candidate is fully revealed, and no new information about candidates is revealed otherwise. Thus, the information revelation is deterministic. This evidently is an extreme assumption, but all that matters for the results is the incentive structure created by the protocol: the games between similar types have strictly dominant actions, and there is a matching pennies game between dissimilar types. Deterministic information revelation is not necessary for this structure of incentives. Consider the following "probabilistic" information revelation protocol: if the messages are unmatched (i.e. debate is fruitful), the type of the focal candidate is revealed with probability π_P and that of the non-focal candidate with probability π_N. Also, when both candidates employ positive messages, each player's type is revealed with probability ϕ_P and when both employ negative messages, each player's type is revealed with a probability ϕ_N. The following two assumptions on these parameters are both necessary and sufficient for inducing the same incentive structure as in a deterministic information revelation protocol: $(A1)$ $\pi_P > \pi_N$ and $(A2)$ $(1-\pi_P)(1-\pi_N) < \min\left\{(1-\phi_P)^2, (1-\phi_N)^2\right\}$. The first assumption says that if there is a fruitful debate, the type of the focal candidate is revealed with a higher probability than that of the other candidate. The content of the second assumption is that the probability that the voter learns neither candidate's type is lower in a fruitful debate than in cross-talk of either kind[6]. Notice that the deterministic model is a special case of the probabilistic model with $\pi_P = 1$ and $\pi_N = \phi_P = \phi_N = 0$. While the results for this probabilistic model are provided in section 5.1, for most part of the paper I deal with the deterministic framework for parsimony of parameters.

Thus, what matters for the incentives at work in this paper is that more information is revealed about a focal candidate, and there is a focal candidate only if debate is fruitful. One way to see the protocol is to think of the role of campaign themes as "defining the debate": the candidates highlight issues which the media then pursues the issue, checking the facts and so forth. The assumptions seem natural if one takes into account the limitations of media space and public attention span. Obviously, the exact process of opinion formation in the electorate involves many factors, and this model avoids making any assumption on that process.

[6] Notice that we do not need to take a position on which kind of cross talk reveals more, i.e. which of ϕ_P or ϕ_N is larger.

2.6 Costly search

The search cost is an index of how easy it is for a candidate to find out detailed information about the other candidate: in this sense it measures the extent of asymmetry of information between candidates. Since the payoff from winning the debate has been normalised, in effect, c captures the ratio of the actual cost of search to the payoff from winning office. Thus, an increase in the importance of the contested office (with the difficulty of finding information about the rival remaning the same) would imply a drop in search cost. Among other things, I am interested in showing how almost-free information is qualitatively different from completely free information: therefore I consider low but positive values of c.

2.7 Alternative interpretation: issue choice

The basic framework discussed in this paper can be applied to debates that are not limited to candidates discussing each others' qualities. For instance, in the context of American politics, one can think of a competition between a Democratic and a Republican candidate where each candidate chooses to argue on one of two issues - say homeland security and redistribution. The former is thought to be a Republican issue while the latter is thought to be a Democratic one. The way it relates to this framework is that the Republican candidate knows more about both the costs and benefits of homeland security measures, but talks only about the benefits while discussing the issue. However, the Democratic candidate can invest in learning about the costs of homeland security measures and highlight them if he decides to discuss the issue. The framework in the paper would be equally applicable to this situation, and would predict, based on actual costs and benefits of each policy, which issue would be discussed by each candidate in equilbrium. In particular, a candidate discussing the benefits of the issue espoused by his own party would be equivalent to what is positive advertising in the rest of the paper, and likewise, a candidate revealing the costs of the issues espoused by the party of his rival would be equivalent to negative advertising.

3 Equilibrium and its properties

Each candidate has to decide whether to acquire information and to choose a message conditional on the information acquired (if any). Since a strategy for a player is a map from types to probabilties of each action, the strategy set for a candidate of type θ should include the following three elements:

1. $p(\theta)$: probability of search.

2. $q(\theta)$: probability of using message P conditional on not searching

3. $r(\theta, \theta')$: probability of using message P conditional on searching, and discovering the rival candidate to be of type θ'

The equilibrium concept considered here is symmetric Bayesian Nash equilibrium. Since only symmetric strategy equilibria are considered, the strategies are not indexed by the identity i of the player.

From the previous discussion, we must have $r(G, G) = 1$ and $r(B, B) = 0$ (strictly dominant strategies). The remaining elements in the strategy set of a player is the set of probabilities $p(G)$, $p(B)$, $q(G)$, $q(B)$, $r(G)$ and $r(B)$ where, with a slight abuse of notation, $r(G)$ denotes $r(G, B)$ and $r(B)$ denotes $r(B, G)$.

It is useful to define $\theta P\theta'$ as the probability of the event that when the two candidates facing each other are of types θ and θ' respectively, the candidate with type θ uses message P. Therefore:

$$\theta P\theta' = p(\theta)r(\theta, \theta') + (1 - p(\theta))q(\theta), \text{ where } \theta, \theta' = G, B \tag{1}$$

Expanding on equation (1), we define:

$$\begin{aligned} BPG &= p_B r_B + (1 - p_B)q_B \\ GPB &= p_G r_G + (1 - p_G)q_G \end{aligned} \tag{2}$$

Lemma 1, which deals with message choice for each type of candidate conditional on available information, brings out the hide and seek nature of the game.

Lemma 1 *For $\theta \in \{G, B\}$, the message choice $q(\theta)$ conditional on not searching, and the message choice $r(\theta)$ conditional on finding the rival type to be different from own type, is given by:*

(a) On searching and finding the rival to be a bad type, the good type uses a positive message ($r(G) = 1$) if $BPG < \frac{1}{2}$ and a negative message ($r(G) = 0$) if $BPG > \frac{1}{2}$. Conditional on not

searching, the good type employs a positive message ($q(G) = 1$) if $\Pr(B) \cdot BPG < \frac{1}{2}$ *and a negative message ($q(G) = 0$) if* $\Pr(B) \cdot BPG > \frac{1}{2}$*, where* $\Pr(B) = 1 - \alpha$.

(b) *On searching and finding the rival to be a good type, the bad type uses a positive message ($r(B) = 1$) if $GPB > \frac{1}{2}$ and a negative message ($r(B) = 0$) if $GPB < \frac{1}{2}$. Conditional on not searching, the bad type employs a positive message ($q(B) = 1$) if* $\Pr(G) \cdot GPB > \frac{1}{2}$ *and a negative message ($q(B) = 0$) if* $\Pr(G) \cdot GPB < \frac{1}{2}$*, where* $\Pr(G) = \alpha$.

Proof. See Appendix B ■

According to Lemma 1, the message choice of a good type depends only on BPG, the probability with which she expects a bad type of the rival to employ the positive message against her. If the bad type goes positive with a high probability, the good type prefers to go negative and expose the bad type, and if the bad type goes negative with a high probability, the good type prefers to go positive and reveal her true type. In the same way, the message of the bad type depends only on GPB, the probability that the good type of the rival uses a positive message against him. The bad type always prefers to avert a fruitful debate with the good type, and therefore tries to mimic her message. As an implication of Lemma 1(a), conditional on not searching, the good type has a strictly dominant message of P if it is the more common type. Similarly, part (b) implies that conditional on not searching, the bad type finds it strictly dominant to play N if the bad type is more common. In other words,

$$\left.\begin{array}{l} \alpha > \frac{1}{2} \Rightarrow q(G) = 1 \\ \alpha < \frac{1}{2} \Rightarrow q(B) = 0 \end{array}\right\} \tag{3}$$

Proposition 1 demonstrates the equilibrium behavior of candidates.

Proposition 1 *The unique equilibrium for different values of the prior*[7] *is as follows:*

(i) *If $\alpha < c$, there is a fully separating equilibrium where no type searches, the good type uses the positive message and the bad type attacks (slander) i.e.* $p(G) = p(B) = 0$, $q(G) = 1$, $q(B) = 0$, *and (off equilibrium),* $r(G) = r(B) = 1$.

[7] We do not consider the case $\alpha = c$ since it is non-generic. In this case, we can have a continuum of equilibria. However, the equilibria discussed in the proposition extended to $c \to \alpha$ still exist in the limit $c = \alpha$. The case of $\alpha = 1 - c$ is not considered due to the same reasons.

(ii) If $\alpha \in (c, 1-c)$, there is a partially separating equilibrium where both types search with positive probability, i.e. $p(G) = \frac{1}{2}(1-\frac{c}{\alpha})$, $p(B) = \frac{1}{2}(1+\frac{c}{1-\alpha})$. *The good type employs a positive message when she does not search and a negative message when she searches and finds the rival to be a bad type, i.e.* $q(G) = 1$, *and* $r(G) = 0$. *The bad type employs a negative message when he does not search and a positive message when he searches and finds the rival to be a good type, i.e.* $q(B) = 0$ *and* $r(B) = 1$.

(iii) If $\alpha > 1 - c$ there is a fully pooling equilibrium where both types send the positive message and neither type searches, i.e. $p(G) = p(B) = 0$, $q(G) = q(B) = 1$, *and (off equilibrium),* $r(G) = 0$, $r(B) = 1$.

Proof. See appendix B. ■

If either type is too rare (when $\alpha < c$ or $\alpha > 1 - c$), then there is no incentive for search, and each candidate assumes the rival to be of the common type. When $\alpha < c$, the bad type finds it strictly dominant to use N (by equation 3).Since is no search, the action of the bad type is predictable, i.e. $BPG = 0$, and by Lemma 1(a), the good type employs message P. Thus, when expected candidate quality is very low, we have a "competitive" electorate, there is a lot of negative advertising, but the good type can always separate itself from the bad type by ensuring a fruitful debate. At the other extreme, when $\alpha > 1 - c$, the good type always uses P (equation 3), and Lemma 1(b) dictates that then the bad type will use P too, and successfully ensure cross talk. Therefore, when the expected candidate quality is very high, we have a "conservative" electorate, where there is only positive advertising, and it is impossible to distinguish the good type from the bad. Notice that the no-search case demonstrates that the rarer type has the advantage in the hide-and-seek game.

Search is undertaken only when neither type is very rare, i.e. the expected candidate quality is moderate ($c < \alpha < 1 - c$). The good candidate always provides arguments supporting himself (positive message) unless he is sure that the rival is a bad type, in which case he tries to expose the rival by going negative. The bad candidate on the other hand has a default message which is negative, but when he is sure that the rival is a good type, he tries to ensure cross talk by defending himself (positive message), hoping that the rival has not searched and is going to employ a positive message too. This equilibrium is supported by the fact that the good candidate searches less frequently than the bad type.

To see the technical intuition for the equilibrium with search, notice that if a candidate were to invest in search with any positive probability, he must play different actions with different types of the rival. Since the good type plays P if the rival is also good, he must play N when search reveals the rival to be bad, i.e. $r(G) = 0$. In the hide-and-seek game with the bad type, the good type does not want to play the same action against the bad type all the time. Hence, the good type plays P conditional on not searching and N conditional on searching and discovering the rival to be bad. Similarly, the bad type has $r(B) = 1$ and $q(B) = 0$. The search probabilities are chosen by each type so as to keep the "other" type indifferent between searching and not searching: search thus performs the role of mixing between the two different actions in the debate.

As demonstrated by the above proposition, irrespective of whether search occurs or not, when two good candidates are in competition, there is cross talk with both candidates arguing in support of themselves (positive message). When two bad candidates face each other, we again have cross talk, but with negative messages if $\alpha < 1 - c$ and with positive messages otherwise. A fruitful debate can occur only between a good and a bad type. However, when two different types face each other, we might also have cross talk with both types either going positive or both going negative.

Notice that although we have considered "mechanistic" or passive voting, including the voter as a rational player does not alter this equilibrium. First, note there is nothing more to learn from cross-talk because both candidates take the same action. As long as there is a common prior over both candidates, it is rational for the voter to randomly choose the winner. In equilibrium, fruitful debate occurs only between two different types of candidates - thus the passive action is again rational. I make the additional assumption that off the equilibrium, if the voter observes a candidate to be of a type that is not supposed to be observed in equilibrium (e.g. type B in the case $\alpha < c$), then the voter assumes that each type of the rival candidate has a small positive probability of having played N, which implies that there remains an uncertainty about the type of the rival. Thus, off the equilibrium path too, the voter strictly prefers to vote for the candidate revealed to be good and against the candidate revealed to be bad. Therefore, the "naive" voting action hardwired in the payoffs does not change if we include a rational voter in the model. Voter inference is formally discussed in Appendix A.

3.1 Properties of Search

A few properties of search are worth mentioning here:

1. Given the equilibrium message strategies, search has the property of *strategic substitutability*[8]: certain search by one type takes away the incentive of the other type to search. Moreover, the type that does not search can mix messages in such a way as to nullify the informational advantage of the type that has searched. Therefore, no candidate searches with certainty, even if search cost is very low. The result that search must be probabilistic casts doubt over the assumption of full information between candidates which is common in the related literature.

2. Search is *reciprocal*, i.e. if one type searches with a positive probability, the other type does so too.

3. If search occurs, the bad type searches with a higher frequency than the good type. In the equilibrium with search, we must have $p(B) > \frac{1}{2} > p(G)$. Since the good type has better information about herself than the bad type has about himself, the marginal value of positive advertising is higher to the good type than the bad. It is this advantage that depresses the good type's incentive to search, and raises the bad type's motivation for the same. This inefficiency due to different incentives for search is present only when search is worthwhile, and I call it the *inefficiency due to search.* For a given prior α, the size of this inefficiency (measured as the difference between $p(G)$ and $p(B)$) increases with the cost of search $\alpha \in (c, 1-c)$. Thus, the cost of search drives a wedge between the incentives to search of the two types. To drive the point home further, note that as $c \to 0$, the inefficiency due to search vanishes in the limit.

4. As the cost of search goes to zero, the outcome of a debate between a good and a bad type approaches that of a matching pennies game, which is a game of complete information. As $c \to 0$, for any α, both $p(B)$ and $p(G)$ approach $\frac{1}{2}$. Thus, both types mix the positive and negative message almost equally when they face each other. It is as if the good type ignores

[8] See Aghion and Tirole (1997) for another model where the principal and agent can both invest in information acquisition about a project, and such investment is a strategic substitute.

the uncertainty and takes into consideration only the bad type of the rival and conversely. Even though in equilibrium there is residual incompleteness of information between the two candidates (since search probabilities are close to half), we have the same outcome that would have come about if the only information asymmetry was between the voter and the candidates.

3.2 Messages and candidate quality

To analyse what messages mean about candidate quality, one really has to look at the voter as a rational player. I have informally discussed before and formally demonstrated in appendix that the equilibrium strategies derived earlier in this section can be used to analyse voter inference of candidate types from messages spoken.

The most important issue highlighted in this paper is that information is revealed by the profile of messages rather than an individual message. Therefore, the same message may induce very different inferences about candidate quality depending on the message spoken by the other candidate and on the expected candidate quality. It has already been pointed out a fruitful debate has either a good candidate "exposing" a bad one, or a good candidate successfully defending herself against "slander" by a bad one. Thus, if in a debate a candidate is revealed to be good (bad), the other type must be bad (good). In case of cross talk, the candidates mimic each other's message. Since the type of both candidates is drawn from a common prior distribution, the voter cannot distinguish between the two, and assigns to each candidate a common, updated posterior distribution of types. When there is cross talk with positive messages, either or both of the candidates must be good. This implies that when there is cross talk with a positive message, the voters would adjust their quality assesement of candidates upwards from the prior. Similarly, when there is cross talk with negative messages, at least one of the two candidates is bad, and voters adjust their assessment downwards. Suppose both candidates cross talk with message M,and the inferred probability of the candidates being good be $\widehat{\alpha}(M)$. A comparison of the posterior $\widehat{\alpha}(M)$ with the prior α allows us to draw several conclusions about the "meaning" of messages to the electorate.

1. If $\alpha \in (c, 1-c)$, then $\widehat{\alpha}(N) < \alpha$, i.e. the assessment of candidate quality goes down when the voters observe both candidates attacking each other in the debate. In fact, the voters believe that the candidates are more likely to be bad than good, i.e. $\widehat{\alpha}(N) < \frac{1}{2}$.

2. If $\alpha \in (c, 1-c)$, then $\widehat{\alpha}(P) > \alpha$, i.e. the assessment of candidate quality goes up when the voters observe positive messages from both candidates. In fact, the voters believe that the candidates are more likely to be good than bad, i.e. $\widehat{\alpha}(P) > \frac{1}{2}$.

3. If $\alpha > 1-c$ there is full pooling on the positive message. The message becomes completely uninformative about quality, and thus $\widehat{\alpha}(P) = \alpha$.

4. If $\alpha < c$, we have full separation. Then, anyone speaking a positive message distinguishes herself as a good type while anyone attacking the rival in debate reveals himself to be a bad type. Formally, in this case, $\widehat{\alpha}(N) = 0$.

4 Comparative Statics

In this section, I examine the comparative static properties of the equilibrium for different levels of the search cost and average candidate quality, and discuss the implications of such properties.

4.1 Welfare Analysis: Candidate Selection

Looking at political campaigns as debates between candidates with partial information about each other helps us understand a few important issues about the efficiency of the campaign process especially in terms of its ability to select the better candidate. The major finding of the model is that as the prior probability of a candidate being good increases, the probability of a good candidate being selected through the electoral process may actually *go down*. Figure 1 plots the total probability of selection of the good candidate against the prior α for some given search cost c[9]. In the graph, there exist downward jumps in otherwise piecewise continuous and monotonically increasing graphs. There are three regimes based on ranges of α : full separation, partial separation and no separation, and the downwards jumps occur when we move from a more efficient regime to a less efficient one as the average candidate quality increases. Note also that since an increase in the search cost favours the bad type, the downward jumps are larger as the cost of information increases.

[9] Denote the total probability of selection of the good type given c and α as $f_c(\alpha)$. It can be deduced from Propositions 1 that:

$$f_c(\alpha) = \begin{cases} 2\alpha - \alpha^2 \text{ if } \alpha < c \\ \frac{3}{2}\alpha - \frac{1}{2}\alpha^2 - \frac{1}{2}c^2 \text{ if } c < \alpha < 1-c \\ 2\alpha - \alpha^2 \text{ if } \alpha > 1-c \end{cases}$$

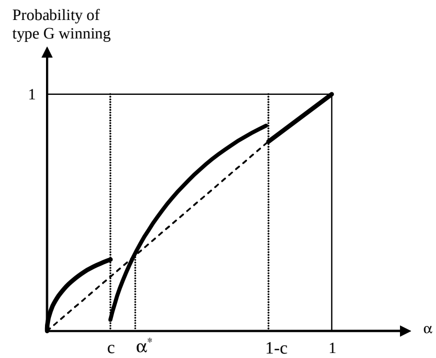

Figure 1: Total probability of selection of a good candidate

The total probability of selection may not fully reflect the efficiency of debate as a selection mechanism. Perhaps a better indicator of the efficiency of debates in selecting the right candidate would be a measure of how often the good candidate wins when competing against a bad candidate. Hence I look at the equilibrium probability of a fruitful debate conditional on candidates being of different types. Denote this probability $\beta(\alpha, c)$. Proposition 2 shows how β changes with the parameters of the model.

Proposition 2 *Suppose one of the two competing candidates is a good type and the other is a bad type. Debate is alway fruitful if $\alpha < c$, never fruitful if $\alpha > 1 - c$, and if $\alpha \in (c, 1 - c)$ debate is fruitful with a probability $\beta(\alpha, c) = \frac{1}{2} - \frac{c^2}{2\alpha(1-\alpha)}$.*

Proof. Follows from proposition 1. ■

When there is no search, the rarer type has full advantage in the hide and seek game. For very low priors, debate is fully efficient and for very high priors, debate is fully inefficient. When the prior is moderate, both types search in equilibrium. In the case when the search cost is very low, i.e. $c \to 0$, the debate between the good and bad types reduces to the matching pennies game in which there is no advantage to either type: hence each message profile occurs with equal probability. Then $\beta(\alpha, c) \to \frac{1}{2}$ for almost all values of the prior. As the search cost increases, there is a further inefficiency due to differential incentives for search that the two types have. This inefficiency due to search leads to a further welfare loss of $\frac{c^2}{2\alpha(1-\alpha)}$. Figure 2 shows how $\beta(\alpha, c)$ changes for different values of α. The dashed line shows β when $c \to 0$.

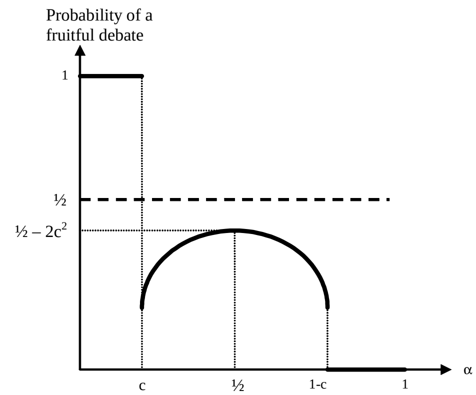

Figure 2: Probability of correct selection when two candidates are of different types

Given an expected candidate quality, if a social planner could choose the search cost c (or more realistically, the extent of office perks), what would the best choice be?

Note that if there is no search, the rarer type has an advantage in the hide and seek game. Thus, if the good type is more common, we unambiguously improve welfare by reducing c (increasing the prize from winning office) and the best choice of c when $\alpha > \frac{1}{2}$ would be as close to zero as possible.

If the bad type is more common, there are two opposing effects. When $c < \alpha$, a marginal reduction in c reduces the inefficiency due to search and inceases β. But in an equilibrium with search, β is always bounded below $\frac{1}{2}$. On the other hand, for any given α less than $\frac{1}{2}$, if we set $c > \alpha$, we get full separation of types. Hence, if the bad type is more common, a social planner can achieve full separation by sufficiently reducing the prize from office.

4.1.1 Commitment

While the voter always prefers that the good type win the debate, she does not internalise the candidate's cost of search. The good type trades off the cost of search with the possible gain from search in terms of the probability of winning the debate. This leads to a welfare loss for the voter, and the problem is worsened by the fact that the good type economises on search cost more than the bad type does. If the good type could somehow commit to search with certainty, the voter would be strictly better off for all α.

To illustrate the commitment problem, suppose the good type is forced to search in the game studied in the previous section. In other words, consider the new game $\mathcal{C}(\alpha, c)$ which is the same game as in the previous section except that the good type searches with certainty. The strategy space of the good type consists only of the functions $q(G)$ and $r(G, \theta)$. In this game, the inefficiency due to search is eliminated.

Proposition 3 *If the prior is moderate* ($c < \alpha < 1 - c$), *in the unique equilibrium of the game* $\mathcal{C}(\alpha, c)$ *the bad type searches with probability* $\frac{1}{2}$, *and the good type mixes both messages with a positive probability against the bad type. For all values of* α *in this range, we have fruitful debate with probability* $\frac{1}{2}$ *when the competing candidates are of different types.*

Proof. In appendix B. ■

Since the good type has full information of her rival's type, the best response of the bad type is to mix the two messages equally against the good type. As a result, we get back the outcome in the matching pennies game, and $\beta = \frac{1}{2}$ for all values of the prior α. Note that here the good type is indifferent between the two messages to play with the bad type. If she were not forced to search, her best response would have been to save on search cost and play P against both types. This is precisely why the good type cannot commit to searching with certainty in the original game.

Thus, if the good type could commit to searching, for all values of $\alpha > c$, the voter would be strictly better off. What is more, if $\alpha > \frac{1}{2}$, the good type would receive a higher payoff in the game $\mathcal{C}(\alpha, c)$ than her Nash payoff and thus would herself like to commit to search if she could.

4.2 Negative Advertising

Next, I look at the type of messages exchanged in the debate as the prior varies. As the average candidate quality improves, the ex-ante probability that a candidate will play a negative message goes down.There are several reasons for this. In general, the bad type's propensity to attack goes down as the probability of the rival being good goes up. For moderate values of the prior, the volume of negative messages decreases with α because the good type's increasing aggressiveness (in searching and attacking the bad type) is more than compensated by the bad type's increasing conservatism (in searching and sending positive messages). This is illustrated in figure 3 and stated in Remark 1, which follows from Proposition 1[10].

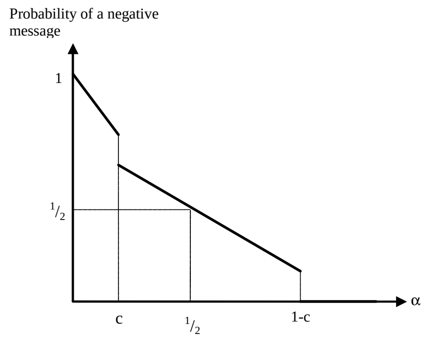

Figure 3: Probability of a negative message as a function of the prior

Remark 1 *The probability of negative advertising is strictly decreasing in the prior for $\alpha < 1 - c$ and is equal to 0 for $\alpha > 1 - c$. Negative advertising is more frequent than positive advertising if the bad type is more common, and less frequent if the good type is more common.*

4.3 Slander

Slander is defined as an attack on the rival in the debate without finding out information about him[11]. Slander may save the cost of searching, but if by mistake one attacks a good type who successfully defends herself, then one loses the election. One major finding of the model is that such a risk is taken only by the bad type. Just like negative advertising, the incidence of slander goes down as the prior improves[12].

If $\alpha < c$, since search is too expensive to undertake, all messages of the bad type are slanderous, and for moderate priors, only $(1 - p(B))$ share of messages of the bad type are slanderous. The probability of slander decreases with improvement in expected candidate quality because the bad type becomes less common and also starts searching more often. For high values of the prior, there is no negative advertising and therefore, no slander. Note that a "clean" electorate in this sense

[10] The probability of a negative message $\gamma(\alpha, c)$ is: $\gamma(\alpha, c) = \begin{cases} 1 - \alpha \; if \; \alpha < c \\ (1 - \alpha) - \frac{c}{2}(1 - 2\alpha) \; if \; \alpha \in (c, 1 - c) \\ 0 \; if \; \alpha > 1 - c \end{cases}$

[11] Slander may also be defined in this model as simply an attack on a good type. With that definition, all qualitative statements about slander made in the remark are true. However, since there is no pre-defined "good" type in a continuous type model, we cannot carry over this definition to the next section.

[12] The probability of slander is $1 - \alpha$ for $\alpha < c$, $\frac{1}{2}(1 - \alpha) - \frac{c}{2}$ if $\alpha \in (c, 1 - c)$ and 0 if $\alpha > 1 - c$.

is very inefficient in terms of candidate selection. As the expected candidate quality improves, a growing share of negative advertising is informed attacks by the good type. Thus, not only the absolute volume of slander, but also the share of slander in all negative messages exhibits a decreasing trend with the prior.

5 Extensions

Although the hide and seek framework discussed in the paper is very simple, it can serve as a legitimate model of electoral competition. This framework can be extended to discuss several features necessary for a richer model of politics.

5.1 Probabilistic Information Revelation

Consider a generalization of the model discussed in section 2 and 3 where the quality of the candidate is revealed to the voter with a probability dependent on the profile of messages. Suppose the revelation of a candidate's quality is a Bernoulli random variable distributed independently conditional on the message profile. If both candidates employ positive messages, each candidate's quality is revealed to the voter with probability ϕ_P, and similarly, with probability ϕ_N if both engage in negative advertising. If, on the other hand, one candidate employs a positive message and the other a negative message, then the former candidate's type is revealed with probability π_P and the latter's type with probability π_Q. I call this structure the "probabilistic information revelation protocol". I make two assumptions on the parameters:

$$(A1) \quad 1 \geq \pi_P > \pi_N \geq 0$$

$$(A2) \quad (1-\pi_P)(1-\pi_N) \quad < \quad \min\left\{(1-\phi_P)^2, (1-\phi_N)^2\right\}$$

$(A1)$ captures the focality of the positive candidate in a fruitful debate, and $(A2)$ says that the probability that the voter learns nothing is lower in a fruitful debate than in cross talk of either kind. First, I demonstrate that under a probabilistic information revelation protocol, these two assumptions are both necessary and sufficient for preserving the incentive structure in the basic framework.

If the types of both candidates are revealed, the voter compares the two and decides. If only one type is revealed, the voter votes for the candidate if he is revealed to be good and against him if he is revealed to be bad. If neither type is revealed, the voter votes randomly. In the notation used in section 2, the payoff to the candidates from each message profile when the two candidates are of the same type are as follows:

G-type \ G-type	P	N
P	*1*	*k*
N	*(2-k)*	*1*

B-type \ B-type	P	N
P	*1*	*m*
N	*(2-m)*	*1*

where $k = \pi_P\pi_N + (1-\pi_P)(1-\pi_N) + 2\pi_P(1-\pi_N) = 1 + (\pi_P - \pi_N)$, and similarly, $m = 1 - (\pi_P - \pi_N)$.

Notice that assumption $(A1)$ implies and is implied by $1 < k \leq 2$ and $0 < m \leq 1$. This ensures that when a good type meets another good type, both find it strictly dominant to use a positive message; and when a bad type meets another bad type, both find it strictly dominant to use negative messages. Next, consider the game between a good type and a bad type, where we can find the payoffs in the individual cells through a little algebra.

G-type \ B-type	P	N
P	a_1	b
N	b	a_2

where

$$a_1 = 2 - (1 - \phi_P)^2$$

$$a_2 = 2 - (1 - \phi_N)^2$$

$$b = 2 - (1 - \pi_P)(1 - \pi_N)$$

Assumption $(A2)$ is equivalent to $b > \max(a_1, a_2)$, which induces the matching pennies structure. Therefore, when a good type faces a bad type, the former prefers fruitful debate (unmatched messages) and the latter prefers cross talk (matched messages).

The following proposition describes the equilibrium in the probabilistic revelation case. Under two additional restrictions, the features of the equilibrium are exactly the same as those in the deterministic revelation case. In fact, proposition 1 is a special case of proposition 4.

Proposition 4 *Assume* $(A1), (A2)$, $\phi_P - \phi_N \geq b - a_1$, *and* $c < \frac{(b-a_1)^2}{(b-a_1)+(\phi_P-\phi_N)}$. *The following is the equilibrium for different values of the prior* α :

(i) If $\alpha < \frac{c}{b-a_1}$, *there is a fully separating equilibrium with no search, i.e.* $p(G) = p(B) = 0$, $q(G) = 1$, $q(B) = 0$, *and (off equilibrium),* $r(G) = r(B) = 1$.

(ii) If $\alpha \in (\frac{c}{b-a_1}, 1 - \frac{c}{b-a_1})$, *there is a partially separating equilibrium where both types search with positive probability, and* $q(G) = 1, r(G) = 0$, $q(B) = 0, r(B) = 1, p(G) = \frac{1}{2b-(a_1+a_2)}\left[(b-a_1) - \frac{c}{\alpha}\right]$, *and* $p(B) = \frac{1}{2b-(a_1+a_2)}\left[(b-a_2) + \frac{c}{1-\alpha}\right]$

(iii) If $\alpha > 1 - \frac{c}{b-a_1}$ *there is a fully pooling equilibrium on the positive message and neither type searches, i.e.* $p(G) = p(B) = 0$, $q(G) = q(B) = 1$, *and (off equilibrium),* $r(G) = 0$, $r(B) = 1$.

I skip the proof of this proposition as it is similar to the proof of proposition 1.

In this case too, inclusion of the voter as a rational player does not change the equilibrium outcome. All the results in sections 3 and 4 go through qualitatively, under appropriate restrictions on the parameters.

5.2 Continuous Type Space

To discuss the continuous type extension we need to formally include the voter as a third player in the game and discuss his beliefs given candidate strategies and revealed information. In Appendix A, I discuss voter beliefs and then show that there is an equilibrium in the continuous type case where there is a cut-off quality such that candidates with their type above the cut-off behave like the good type in the baseline (two-type) model, and those with type below the cut-off behave like the bad type. Therefore, all results in the baseline model carry over. Moreover, we get an endogenous classification of the continuum of quality into "good" and "bad", dictated by voter beliefs. The results are presented for the determinisitic information revelation protocol, but I believe that similar results can be derived for the probabilistic revelation too.

The continuous type extension is not of mere technical interest. It provides a comparison with the existing literature on adverse selection with one principal (the median voter) and two competing agents (the candidates). Banks (1990) analyses Downsian competition where candidates may make false announcements about their preferred positions, but lying has an exogenous cost that increases in the distance between their preferred and announced positions. While Banks finds pooling of candidate types over an interval containing the median voter's best point, Callander and Wilkie

(2005) show that if there is a cheap talking type in the model, pooling happens at two disjoint intervals on either side of the median voter's best point. Thus, the "best" types (those that are preferred most by the median voter) pool in the former paper while the "moderate" types pool in the latter, and all the other types separate. The continous type space in this model can be interpreted as a space of possible candidate locations on the Hotelling line, with higher quality implying a location closer to the median voter's ideal point. In the equilibrium in the current model there are two clusters of pooling - one for the "good" types and the other for the "bad" types. The message in this paper is that if competition reveals information about only one agent, it is possible to separate the good set of types from the bad set of types (where good and bad types are defined endogenously), but one cannot separate within the good or bad set of types.

5.3 Long-term Average Quality

In this paper, the probability of a good type has been treated as a fixed parameter. One can think of the proportion of the good type (average quality) in the candidate pool continuously evolving through successive elections. Imagine that each time a good candidate is selected, the public assessment of the average quality goes up, and each time a bad candidate is selected, the assessment goes down. If this adjustment process occurs sufficient number of times, the assessment of α will settle at that level where the probability of selection of the good candidate is equal to her probability of occurrence in the candidate pool. In other words, in the long term assessment of α is a fixed point of the function $f_c(\alpha)$, which is the probability of selection of the good type studied in Section 4.1. In Figure 1, the fixed points are those points where the 45^0 line intersects the function $f_c(\alpha)$. Note that there are multiple fixed points. In particular, all values of $\alpha > 1 - c$ can serve as fixed points. Thus, in the long term a candidate is more likely to be good than bad. However, it is also possible that the system settles at some $\alpha < \frac{1}{2}$, which has been indicated by α^* in Figure 1, which implies that the political system can be stuck in a cycle of low expectations[13]. It can also be shown that α^* is increasing in c, implying that the average candidate quality will worsen in the long term if information becomes cheaper.or the office becomes more lucrative. Note that there is

[13] The existence of α^* in the range $(1-c, c)$ is guaranteed by the fact that in the range, the function starts below the 45^0 line, and ends above it. Examining the functional form, $\alpha^* = \frac{1}{2} - \frac{\sqrt{1-4c^2}}{2} < \frac{1}{2}$. Also, the concave graph shifts upwards everywhere in this range if c goes down, hence α^* goes down.

a fixed point at $\alpha = 0$ too. However, none of the equilibria are stable: a small perturbation can lead to a path moving away from the fixed point.

5.4 Incumbency Advantage

The debates framework can be extended to examine the incumbency advantage. According to this model, the source of the advantage is the ability to alter voters' perceptions. Suppose the incumbent has successfully been able to raise the prior probability of being good to α_I, which is higher than the prior α_C from which the challenger is drawn. If $0 < \alpha_C < \alpha_I < 1$, then in equilibrium, the good type of the incumbent always employs a positive message, the bad type of the challenger always goes negative. The bad type of incumbent mixes through search, imitiating the good type partially. Similarly, the good type of the challenger mixes, unable to separate himself from the bad type fully. The good type of the incumbent always wins. The bad type of the incumbent sometimes wins against the good type of challenger.

Conversely, if the incumbent is assessed to be worse than the average challenger ($\alpha_C > \alpha_I$), then the incumbent always loses to the good type of the challenger. Even the good type of the incumbent may lose to the bad type of the challenger. Thus, the voter rewards good performance and punishes bad performance of the incumbent. Note that this result hinges on observability and not on risk aversion of the voter. If the voter is risk averse, then the incumbent can have an advantage over the challenger even if he performs somewhat badly, since his quality would be known with more certainty than that of the challenger. This analysis indicates that the whole objective of political activity is to alter voters' expectation of candidate quality before the actual campaign begins.

5.5 Other extensions

There are a few more assumptions that limit the scope of the model. First, since the model is static, one cannot look at the dynamic choice of arguments and counter-arguments. It is assumed, with some support from what is observed in real electoral races, that the public campaign strategy is fixed far in advance of the time the messages are actually spoken. Second, the candidate is constrained to use either a positive message or a negative message. This assumption, although standard in the literature, is an abstraction from reality where we observe candidates using a combination of both positive and negative messages. Third, candidate quality is one-dimensional in this model.

If candidate quality were multidimensional, a candidate could dodge a fruitful debate by focussing on a different dimension of the same person rather than on a different person. Thus, there could still be a fruitless debate with one candidate sending a positive message and the other sending a negative message. This assumption of unidimensional quality simplifies the analysis considerably and demonstrates some of the basic tradeoffs involved in competitive information provision at the cost of a certain loss of richness. Each of these assumptions could be relaxed and the model can thereby be extended in promising directions for future research.

6 Conclusion

This paper examines the choice between positive and negative campaigns as a particular case of the more general decision problem of contradicting the opponent's argument (revealing information through fruitful debate) and mimicking his argument (concealing information through cross talk). The fundamental idea of the paper is that information transmitted in a debate is *conditionally verifiable*: the truth of an argument is more transparent when compared with a counter-argument. Using this model, most of the results in the existing literature on positive and negative advertising are confirmed, and additionally, issues like lies and slander are explained. The model points out that a lie or slander, while useless to the voter, might be useful to the candidate in "muddling" the debate and thus distracting the attention of the electorate. Besides, it demonstrates several interesting features of information search by candidates, the most important of which is that the bad type searches more frequently than the good type. This arises from a commitment problem of the good type since the good type has a higher incentive to economise on search cost than the bad type. This divergence of interest between the good type of candidate and the voter is an additional source of welfare loss for the electorate, and this loss can be captured analytically only if one allows for the option of investment in information acquisition. The model also shows that, contrary to popular perception, voter welfare is not monotonic in the average candidate quality because of the hide-and-seek nature of political campaigns.

This model is interesting due to several reasons independent of the analysis of electoral rhetoric. While the theoretical prediction of a matching pennies game is that both the players should mix actions equally, in laboratory experiments, the matcher and unmatcher have often exhibited different behavior. This divergence of observed outcomes from the equilibrium prediction has troubled economists (Crawford and Iriberri (2005) and Rosenthal, Shachat and Walker (2003) are two recent examples). In the game studied in this paper, the assumption of incomplete information and costly search allows for an equilibrium where each type employs a pure action in the debate conditional on information available, and mixing (if any) occurs through probabilistic search. The innovation in this paper may be thought of a way to break the type-independence and mixed-action outcome in the matching pennies game. However, it is to be noted that the solution, though suited to this particular case, makes several strong assumptions which may not be applicable in a general sense.

Last but not the least, the paper also suggests a new reason why advertisements for experiential goods are believed, at least partially. The current paper points out that since each advertiser's willingness to misinform is potentially checked by the ability of his competitor to inform, advertisements can be treated as partially credible signals of the private information of the sponsor. These contributions, I believe, would make the paper independently interesting to those not actively engaged with the literature on positive and negative advertising.

7 References

1. Aghion, P. and J. Tirole (1997): "Formal and Real Authority in Organizations", Journal of Political Economy 105 (1): 1-29

2. Ansolabehere, S. and S. Iyengar (1995): Going Negative: How Political Advertisements Shrink and Polarize the Electorate, (New York: Free Press)

3. Banks, J. (1990): "A Model of Electoral Competition with Incomplete Information", Journal of Economic Theory 50: 309-25

4. Callander, S. and S. Wilkie (2005): "Lies, Damned Lies and Political Campaigns", Working paper, Northwestern University

5. Crawford, V. and N. Iriberri (2004): "Focality, Naivette and Sophistication in Experimental "Hide and Seek" Games, Working Paper, UCSD

6. Lau, R., L. Spiegelman, C. Heldman and P. Babbitt (1999): The Effects of Negative Political

Advertisements: A Meta-analytic Assessment, The American Political Science Review 93(4), 851-875

7. Harrington, J. and G. Hess (1996): “A Spatial Theory of Positive and Negative Campaigning", Games and Economic behavior 17, 209 – 229

8. MilGrom, P. and J. Roberts (1986): “Price and Advertising as Signals of Product Quality", Journal of Political Economy 94: 796 - 821

9. Prat, A. (2004): “Rational Voters and Political Advertising", Oxford Handbook of Political Economy, eds. Barry Weingast and Donald Wittman (Oxford University Press)

10. Polborn, M. and D. T. Yi (2006): “Informative Positive and Negative Campaigning", Quarterly Journal of Political Science 1(4): 351-71

11. Rosenthal, R., J. Shachat and M. Walker (2003): “Hide and Seek in Arizona", International Journal of Game Theory 32: 273 - 293

12. Skaperdas, S. and B. Grofman (1995): “Modelling Negative Campaigning", The American Political Science Review 89, 49 – 61

13. Tirole, J. (1988): The Theory of Industrial Organization, MIT Press

8 Appendix A : Rational voter model

In the main body of the paper, the voter is treated as "passive" . In this section, we include a rationally updating voter and a more general space of types.

Consider a game $\mathcal{G}(\Theta, F, c)$ with a voter and two candidates $i = 1, 2$. Each candidate has a private quality type θ_i that is drawn from a common type space $\Theta \subset R$ according to a common distribution $F(\cdot)$. The action space of candidates $i = 1, 2$ is exactly the same as specified in Section 2. Formally, this is a three stage game. In stage 1, the candidates observe their own realised type and choose search action X privately. In stage 2, the candidates play debate action $M \in \{P, N\}$. Since action in stage 1 is private, stages 1 and 2 are considered strategically simultaneous. The (deterministic) information revelation protocol is now formally introduced as a function $R : M^2 \times$

$\Theta^2 \to \Theta \cup \phi$, which has the following form

$$R(P,P,\boldsymbol{\theta}) = R(N,N,\boldsymbol{\theta}) = \phi$$
$$R(P,N,\boldsymbol{\theta}) = \theta_1 \text{ and } R(N,P,\boldsymbol{\theta}) = \theta_2$$

In stage 3, the voter observes $R(\mathbf{M},\boldsymbol{\theta}) = \rho \in \Theta \cup \phi$ and votes for either candidate, and the utilities are realised. Assume that if the winning candidate's type is θ, the voter receives a utility of θ. The voter's action is $v_i(\rho)$, which is the probability of voting for candidate i. Assuming that the voter votes for both candidates with the same probability whenever he is indifferent, we have $v_i(\rho) \in \{0, \frac{1}{2}, 1\}$. The utility of candidate i from debate is $u_i(\mathbf{M},\boldsymbol{\theta}) = 2 \cdot v_i(\rho),\ i = 1,2$. The total utility of the candidate, as before, is the utility from winning the debate less the search cost $c \in (0, \frac{1}{2})$. Our equilibrium concept is symmetric Bayesian Nash equilibrium.

8.1 Voter and candidate strategies

In the game $\mathcal{G}$, the voter rationally forms beliefs about candidate types based on information ρ and $\mathbf{M}$. However, the structure of the game is such that the only thing that matters for ranking the candidates is ρ and not $\mathbf{M}$. Suppose the expected type of candidate i is $\mu(\theta_i|\rho)$. Since this is also the expected utility from a candidate when he is elected, the voter compares $\mu(\theta_1|\rho)$ and $\mu(\theta_2|\rho)$ and votes for whichever is higher. If they are equal, the voter randomises with equal probability.

If $\rho = \phi$, both candidates are playing the same action. Since the candidates are playing symmetric strategies, the voter cannot distinguish between the two. Hence, $\mu(\theta_1|\phi) = \mu(\theta_2|\phi)$, and therefore $v_i(\phi) = \frac{1}{2}$. If $\rho = \theta_i$, the type of candidate i is known to be θ_i while $\mu(\theta_{-i}|\theta_i)$ is inferred by the voter either by Bayes Rule from equilibrium strategies or by out-of-equilibrium beliefs. Hence, the called *rational voting strategies* are:

$$v_i(\theta_i) = \begin{cases} 1 \ if \ \theta_i > \mu(\theta_{-i}|\theta_i) \\ 0 \ if \ \theta_i < \mu(\theta_{-i}|\theta_i) \\ \frac{1}{2} \ if \ \theta_i = \mu(\theta_{-i}|\theta_i) \end{cases} \tag{4}$$
$$v_i(\phi) = \frac{1}{2}$$

The strategy space is defined by the functions $p(\theta)$, $q(\theta)$ and $r(\theta, \theta')$ as mentioned in section 3. Note that this set-up can handle any type space of reasonable generality. First, consider a binary type space.

8.2 Binary Type space

Suppose that there are only two types: Good (G) and Bad (B). In formal terms, $\Theta = \{G, B\}$, where G and B are two real numbers with $G > B$. The distribution $F(\cdot)$ now becomes Bernoulli, with the prior $\Pr(\theta_i = G) = \alpha \in (0, 1)$ for $i = 1, 2$. It is shown here that for a natural specification of out-of-equilibrium beliefs, the game $\mathcal{G}(\{G, B\}, \alpha, c)$ has the same unique equilibrium as the game $\mathcal{H}(\alpha, c)$ discussed in the main body of the paper[14]. To demonstrate that, I show that for game $\mathcal{G}(\{G, B\}, \alpha, c)$, rational voting strategy is the same as *passive voting strategy* considered in $\mathcal{H}$, and therefore leads to the same outcome. Define passive voting strategy as

$$v_i(\rho) = \begin{cases} 1 \; if \quad \rho = \theta_i = G \\ 0 \; if \quad \rho = \theta_i = B \\ \frac{1}{2} \qquad \quad if \; \rho = \phi \end{cases}$$

Assume that if $\theta_i \in \{G, B\}$ is revealed out of equilibrium in a debate, then the voter assumes that both types G and B of candidate $-i$ have small positive probabilities η_G and η_B of having deviated and played N. This implies that off the equilibrium path,

$$B < \mu(\theta_{-i}|\theta_i) < G, \;\; \theta_i \in \{G, B\} \tag{5}$$

Next, note that in game $\mathcal{H}$, for action profiles that constitute the equilibrium, whenever any type θ of candidate i is revealed, there is a positive probability of both types G and B of candidate $-i$ to have attacked candidate i. Hence, along the equilibrium path too, by Bayes rule, we must have condition (5) satisfied. The next Lemma established this fact more generally for the game $\mathcal{G}(\{G, B\}, \alpha, c)$:

Lemma 2 *In the game $\mathcal{G}(\{G, B\}, \alpha, c)$, there is no symmetric equilibrium where for some type*

[14] To be exact, these games are specifed differently, so the equilibria can never be the same. But what I mean is that in equilibrium, the candidate and voter behaviour are the same in both games.

$\theta_i \in \{G, B\}$, *we have* $\mu(\theta_{-i}|\theta_i) = \theta_i$.

Proof. In Appendix B ■

From Lemma 3, we have for any equilibrium strategy profile of the game $\mathcal{G}(\{G, B\}, \alpha, c)$

$$B < \mu(\theta_{-i}|\theta_i) < G, \ \theta_i \in \{G, B\} \tag{6}$$

(5) and (6) imply that if $\Theta = \{G, B\}$, we must have

$$\begin{aligned} \mu(\theta_{-i}|\theta_i) &> \theta_i \ if \ \theta_i = B \\ \mu(\theta_{-i}|\theta_i) &< \theta_i \ if \ \theta_i = G \end{aligned}$$

From (4), the rational voting strategy supporting any equilibrium of $\mathcal{G}$, on and off the equilibrium path, is:

$$v_i(\rho) = \begin{cases} 1 \ if & \rho = \theta_i = G \\ 0 \ if & \rho = \theta_i = B \\ \frac{1}{2} & if \ \rho = \phi \end{cases}$$

Hence, passive voting is rational here. Therefore, we have the following proposition:

Proposition 5 *Given the out-of-equilibrium beliefs stated in* (5), *the solution to* $\mathcal{G}(\{G, B\}, \alpha, c)$ *is the same as the solution to the game* $\mathcal{H}(\alpha, c)$.

8.3 Continuous Type Space

This section demonstrates that all the results proved in the case of the binary type space extend naturally to a case where quality can vary over a continuum. Moreover, the continuum breaks into a “good" set and a “bad" set endogenously.

Normalise the type space Θ and consider it to be the unit interval $[0, 1]$. Suppose $F(\cdot)$ is a non-atomic prior distribution from which θ is drawn. Assume that $F(\cdot)$ has full support over Θ. Note that voter inference in this case is non-trivial. If some candidate is revealed to be of type $\theta \in (0, 1)$, then the attacker potentially can be of a type that is strictly better, equal or strictly worse. In the two-type model, both types were extreme - and therefore revelation of one type was enough for the voter to decide. What is interesting is that in the continuous model too, there is a class of equilibria in which rational voting looks very much like passive voting.

The strategy functions $p(\theta), q(\theta)$ and $r(\theta, \theta')$ are defined in the same way as in Section 3, except that.these functions are now probability density functions - assume them to be continuous except at a finite number of points.

Next, define $h(\theta, \theta')$ as the probability density of the event that conditional on candidates of types θ and θ' respectively, the one with type θ employs message P. This is the equivalent of $\theta P \theta'$ as defined in (1) in the discrete set-up. Formally,

$$h(\theta, \theta') = (1 - p(\theta))q(\theta) + p(\theta)r(\theta, \theta'), where\ \theta, \theta' \in [0, 1]^2 \tag{7}$$

Using (7), define $g(\theta, \theta')$ as the probability density that type θ is revealed through a fruitful debate against type θ'. In other words, $g(\theta, \theta')$ is the probability density of the event that conditional on type θ facing type θ', type θ plays message action P and type θ' plays message N.

$$g(\theta, \theta') = h(\theta, \theta') \left(1 - h(\theta, \theta')\right), \text{ where } \theta, \theta' \in [0, 1]^2 \tag{8}$$

Next, define $e(\theta)$ as the expected type of candidate $-i$ when candidate i has been revealed to be of type θ. Formally, $e(\theta) = \mu(\theta_{-i} | \theta_i = \theta)$. When a type θ is revealed in equilibrium with positive probability, using (8), $e(\theta)$ can be calsulated as:

$$e(\theta) = \frac{\int_0^1 \theta' g(\theta, \theta') dF\left(\theta'\right)}{\int_0^1 g(\theta, \theta') dF\left(\theta'\right)},\ when\ \int_0^1 g(\theta, \theta') dF\left(\theta'\right) > 0 \tag{9}$$

If some type θ is not revealed in equilibrium with a positive probability, then $e(\theta)$ has to be determined by an appropriate specification of out of equilibrium beliefs.

Signaling games with continuous types often admit multiple equilibria. Both Banks (1990) and Callander and Wilkie (2005) use the refinement of universal divinity to select equilibria. Here, the main interest is in the link between passive and rational beliefs, and specifically in the existence of equilibria that can be supported by passive beliefs. To avoid equilibria that are too dependent on beliefs off the equilibrium path, assume the following restriction on the set of equilibria:

In equilibrium, all types $\theta \in \Theta$ should be revealed with positive probability, i.e.

$$\int_0^1 g(\theta, \theta') dF\left(\theta'\right) > 0 \text{ for all } \theta \in \Theta \tag{10}$$

This guarantees that $e(\theta)$ is defined by (9) for all θ.

Next, separate the type space into disjoint sets $\mathbf{G}, \widetilde{\mathbf{M}}$ and $\mathbf{B}$ such that

$$\begin{aligned} \mathbf{G} &= \{\theta : e(\theta) < \theta\} \\ \mathbf{B} &= \{\theta : e(\theta) > \theta\} \\ \widetilde{\mathbf{M}} &= \{\theta : e(\theta) = \theta\} \end{aligned}$$

Since we are looking for equilibria that are similar to those in the two-type case, consider equilibria where $\widetilde{\mathbf{M}}$ is a collection of a finite number of points, and thus has measure zero. Henceforth, ignore the set $\widetilde{\mathbf{M}}$ and look only at $\mathbf{G}$ and $\mathbf{B}$.

If candidate i is revealed to be of type $\theta \in \mathbf{G}$, $\mu(\theta_{-i}|\theta \in \mathbf{G}) = e(\theta) < \theta$, implying that the voter will vote in favour of candidate i. Conversely, if candidate i is revealed to have type $\theta \in \mathbf{B}$, he is voted against. Hence, if there is an equilibrium with restriction (9), it would look like the equilibrium with binary types and passive strategies. Note however that in Section 2, the prior in favour of the good type was exogenous, and in this section the measure of the set $\mathbf{G}$ is determined endogenously. Formally, call

$$\int_{\theta \in \mathbf{G}} dF(\theta) = \alpha$$

We need to show that in equilibrium, we must have $\alpha \in (0, 1)$.

To put more structure on the set of equilibria, consider another restriction. This restriction is not necessary for the existence of binary equilibria with passive voting strategies. However, it selects equilibria among those with such strategies that lead to a "natural" interpretation of good and bad. This is a weaker form of the monotonicity restriction in Polborn and Yi (2006)[15]. Define the expected utility in equilibrium for type θ of player i as $U_i(\theta)$, and stipulate that in equilibrium

[15] Polborn and Yi (2006) assumes that expected utility strictly increases in type - here weak monotonicity is assumed.

a type must not have a strictly greater expected utility than a higher type, i.e.

$$\text{For } \varphi, \varphi' \in \Theta^2, \ \varphi < \varphi' \Rightarrow U_i(\varphi) \leq U_i(\varphi') \tag{11}$$

In the candidate equilibrium, all $\theta \in \mathbf{G}$ and all $\theta \in \mathbf{B}$ play the same strategy, and therefore,

$$\left.\begin{array}{l} \varphi \in \mathbf{G}, \varphi' \in \mathbf{G} \Rightarrow U_i(\varphi) = U_i(\varphi') \\ \varphi \in \mathbf{B}, \varphi' \in \mathbf{B} \Rightarrow U_i(\varphi) = U_i(\varphi') \\ \varphi \in \mathbf{G}, \varphi' \in \mathbf{B} \Rightarrow U_i(\varphi) \geq U_i(\varphi') \end{array}\right\} \text{ where } \varphi, \varphi' \in \Theta^2 \tag{12}$$

Also, unless there is complete pooling in equilibrium, for any $\varphi \in \mathbf{G}$ and $\varphi' \in \mathbf{B}$, we must have $U_i(\varphi) > U_i(\varphi')$. Therefore, the restriction (11) and conditions (12) imply that for such a candidate equilibrium, if all types do not pool on the same strategy, there must exist some $\theta^* \in (0, 1)$ such that

$$\varphi \in \mathbf{G}, \varphi' \in \mathbf{B} \Leftrightarrow \varphi' < \theta^* < \varphi \tag{13}$$

In other words, $\mathbf{G} = \{\theta : \theta > \theta^*\}$ and $\mathbf{B} = \{\theta : \theta < \theta^*\}$. With this definition of the sets $\mathbf{G}$ and $\mathbf{B}$, there indeed exists a class of equilibria with the requisite properties. Moreover, any number in the range $(c, 1 - c)$ can serve as θ^*.

Proposition 6 is a formal statement of the existence and characterisation of equilibria in the continuous type case. It demonstrates that the type space endogenously breaks into two sets which conform to the “natural" definition of good and bad, and that the equilibrium with the binary type space is replicated.

Proposition 6 *Consider any* $\theta^* \in (c, 1-c)$, *and define sets* $\mathbf{G}$ *and* $\mathbf{B}$ *as in* (13). *Then the following* *is an equilibrium obeying restrictions* (10) *and* (11) :

$$p(\theta) = \begin{cases} p(G) \text{ if } \theta \in \mathbf{G} \\ p(B) \text{ if } \theta \in \mathbf{B} \end{cases}$$

$$q(\theta) = \begin{cases} q(G) \text{ if } \theta \in \mathbf{G} \\ q(B) \text{ if } \theta \in \mathbf{B} \end{cases}$$

$$r(\theta, \theta') = \begin{cases} 1 \text{ if } \theta \in \mathbf{G}, \theta' \in \mathbf{G} \\ r(G) \text{ if } \theta \in \mathbf{G}, \theta' \in \mathbf{B} \\ r(B) \text{ if } \theta \in \mathbf{B}, \theta' \in \mathbf{G} \\ 0 \text{ if } \theta \in \mathbf{B}, \theta' \in \mathbf{B} \end{cases}$$

where, if $\alpha = 1 - F(\theta^*)$, *the quantities* $p(G)$, $p(B)$, $q(G)$, $q(B)$, $r(G)$ *and* $r(B)$ *are given by Proposition* 1(ii).

The fact that the strategies mentioned in the proposition constiute an equilibrium conforming to the aforementioned restrictions can be easily checked.

This proposition states that there is always a monotonic equilibrium with rational voting that is supported by passive voting strategies. There is a cut-off type above which all types are deemed to be “good" and below which all types are deemed to be “bad". All good types behave like type G in the discrete case, and all bad types behave like type B. The voter simply votes in favour of a candidate if he is revealed to be a good type and against him if he is revealed to be a bad type. Note however that this cutoff type is not unique. In fact, any number in the range $(c, 1 - c)$ can serve as the cut-off type Thus, it is possible that in two electorates where the prior distributions from which the two candidates are drawn are exactly the same, the definition of a good candidate and a bad candidate are different. Comparative static conclusions are difficult to draw because of the multiplicity of equilibria, but it can be said that the supportable set of cut-off types expands as the search cost decreases.

9 Appendix B: Proofs

To define the equilibrium strategies in this setting, some basic notation needs to be introduced.

Define $Eu_i^{\theta}(M|p,q,r,I)$ as the expected utility from debate to type $\theta \in \{G,B\}$ of player i from playing message $M \in \{P,N\}$ when player $-i$ is using strategies $p(\cdot), q(\cdot)$ and $r(\cdot,\cdot)$, and the informaion available to the player is $I \in \{G,B,\phi\}$. If there is search and the type of the rival is known then $I=G$ or $I=B$, else, $I=\phi$. This expected utility is constructed from the debate payoff $u_i(\mathbf{M},\boldsymbol{\theta})$, taking expectation over the possible messages of the rival (from the strategies) and if the rival type is not known, then over possible types of the rival too.

Define as $Eu_i^{\theta}(m|p,q,r,I)$ the expected utlity when type θ of player i plays message P with a probability $m \in [0,1]$.

Define $EU_i^{\theta}(S|p,q,r) = E_{(\theta')}\left[\underset{m\in[0,1]}{\arg\max} Eu_i^{\theta}(m|p,q,r,\theta')\right] - c$ as the extected utlity from search, taking into account the optimal message choice post search, and taking expectation over rival types and $EU_i^{\theta}(NS,m|p,q,r) = Eu_i^{\theta}(m|p,q,r,\phi)$ as the expected utility from not searching and playing a mix m of messages.

Equilibrium strategies is a triad of functions $\{p^*(\cdot), q^*(\cdot), r^*(\cdot,\cdot)\}$ for $\theta \in \{G,B\}$ and $\theta' \in \{G,B\}$ such that:

1. $r^*(\theta,\theta') = \underset{m\in[0,1]}{\arg\max} Eu_i^{\theta}(m|p^*,q^*,r^*,\theta')$,
2. $q^*(\theta) = \underset{m\in[0,1]}{\arg\max} Eu_i^{\theta}(m|p^*,q^*,r^*,\phi)$, and
3. $p^*(\theta) = \underset{p\in[0,1]}{\arg\max}\left\{pEU_i^{\theta}(S|p^*,q^*,r^*) + (1-p)EU_i^{\theta}(NS,q^*|p^*,q^*,r^*)\right\}$.

When a candidate of type θ knows that his rival is also of type θ, then the candidate has a strictly dominant message. Denote this message by $D(\theta)$. Denote by $\overline{D(\theta)}$ the other available action in the message space, i.e. $\overline{D(\theta)} = \{P,N\}\backslash D(\theta)$. In the same way, denote by $\overline{\theta}$ the type different from θ, i.e. $\overline{\theta} = \{G,B\}\backslash\theta$.

We shall first prove a few results in the form of claims and use those results to find the equilibrium strategies for different parameter values.

Claim 1 $Eu_i^{\theta}(D(\theta)|p,q,r,\theta) - Eu_i^{\theta}(\overline{D(\theta)}|p,q,r,\theta) = 1$

Proof. For $\theta=G$, $D(\theta)=P$, and $\overline{D(\theta)}=N$

$Eu_i^{G}(P|p,q,r,G) - Eu_i^{G}(N|p,q,r,G) = \{GPG_{-i} + 2(1-GPG_{-i})\} - \{(1-GPG_{-i})\} = 1$

For $\theta=B$, $D(\theta)=N$, and $\overline{D(\theta)}=P$

$Eu_i^B(N|p,q,r,B) - Eu_i^B(P|p,q,r,B) = \{(1-BPB_{-i}) + 2BPB_{-i}\} - \{BPB_{-i}\} = 1$ ■

The above claim establishes that if type θ finds it strictly dominant to use message $D(\theta)$ when he knows that the rival is of type θ.

Claim 2 *If $p_i(\theta) > 0$, we must have* $Eu_i^\theta(D(\theta)|p,q,r,\overline{\theta}) < Eu_i^\theta(\overline{D(\theta)}|p,q,r,\overline{\theta})$

Proof. Suppose not. Hence, $Eu_i^\theta(D(\theta)|p,q,r,\overline{\theta}) \geq Eu_i^\theta(\overline{D(\theta)}|p,q,r,\overline{\theta})$.

Since $p_i(\theta) > 0$, we must have

$$EU_i^\theta(S|p,q,r) \geq \max\left[EU_i^\theta(NS,P|p,q,r), EU_i^\theta(NS,N|p,q,r)\right]$$

Using claim 1, we can rewrite this as

$$\begin{aligned} & \Pr(\theta)Eu_i^\theta(D(\theta)|p,q,r,\theta) + \Pr(\overline{\theta})\max\left[Eu_i^\theta(D(\theta)|p,q,r,\overline{\theta}), Eu_i^\theta(\overline{D(\theta)}|p,q,r,\overline{\theta})\right] - c \\ > \ & \max\left[EU_i(NS,D(\theta)|p,q,r), EU_i^\theta(NS,\overline{D(\theta)}|p,q,r)\right] \end{aligned}$$

From our supposition,

$$\begin{aligned} LHS &= \Pr(\theta)Eu_i^\theta(D(\theta)|p,q,r,\theta) + \Pr(\overline{\theta})Eu_i^\theta(D(\theta)|p,q,r,\overline{\theta}) - c \\ &= EU_i^\theta(NS,D(\theta)|p,q,r) - c \end{aligned}$$

Now, $RHS = \max\left\{\begin{array}{c} \Pr(\theta)Eu_i^\theta(D(\theta)|p,q,r,\theta) + \Pr(\overline{\theta})Eu_i^\theta(D(\theta)|p,q,r,\overline{\theta}), \\ \Pr(\theta)Eu_i^\theta(\overline{D(\theta)}|p,q,r,\theta) + \Pr(\overline{\theta})Eu_i^\theta(\overline{D(\theta)}|p,q,r,\overline{\theta}) \end{array}\right\}$

By claim 1 and our supposition, $RHS = EU_i^\theta(NS,D(\theta)|p,q,r) > EU_i^\theta(NS,D(\theta)|p,q,r) - c = LHS$, which is a contradiction. ■

The above claim establishes that whenever there is search with a positive probability, type θ uses message $\overline{D(\theta)}$ when he knows that the rival is of type $\overline{\theta}$. Claims 1 and 2 determine what actions will be played by a type when the rival type is known. Note that the choice of message post search is independent of the strategy of the rival type. Thus,

$$EU_i^\theta(S|p,q,r) = \Pr(\theta)Eu_i^\theta(D(\theta)|p,q,r,\theta) + \Pr(\overline{\theta})Eu_i^\theta(\overline{D(\theta)}|p,q,r,\overline{\theta}) - c \tag{14}$$

Claim 3 $EU_i^\theta(S|p,q,r) - EU_i^\theta(NS,\overline{D(\theta)}|p,q,r) = \Pr(\theta) - c$

Proof. By equation (14), $EU_i^\theta(S|p,q,r) - EU_i^\theta(NS, \overline{D(\theta)}|p,q,r)$ equals

$$\begin{aligned} &\Pr(\theta)\left[Eu_i^\theta(D(\theta)|p,q,r,\theta) - Eu_i^\theta(\overline{D(\theta)}|p,q,r,\theta)\right] \\ &+\Pr(\overline{\theta})\left[Eu_i^\theta(\overline{D(\theta)}|p,q,r,\overline{\theta}) - Eu_i^\theta(\overline{D(\theta)}|p,q,r,\overline{\theta})\right] - c \end{aligned}$$

By claim 1, the above expression equals $\Pr(\theta) - c$ ■

Note that $c = \alpha$ or $c = 1 - \alpha$ are not considered in our range of parameter values. Thus, between and no search with $\overline{D(\theta)}$, one always strictly dominates the other, based on the values of the parameter. Most importantly, if type θ searches with positive probability, he will not play the message $\overline{D(\theta)}$ conditional on not searching. If the probability of search is strictly between 0 and 1, then $D(\theta)$ will be played by θ conditional on playing action $X = NS$.

Claim 4 $EU_i^\theta(S|p,q,r) - EU_i^\theta(NS, D(\theta)|p,q,r) = \Pr(\overline{\theta})(2\overline{\theta}P\theta_{-i} - 1) - c$

Proof. By equation (14), $EU_i^\theta(S|p,q,r) - EU_i^\theta(NS, \overline{D(\theta)}|p,q,r)$ equals:

$$\begin{aligned} &\Pr(\theta)\left[Eu_i^\theta(D(\theta)|p,q,r,\theta) - Eu_i^\theta(D(\theta)|p,q,r,\theta)\right] \\ &+\Pr(\overline{\theta})\left[Eu_i^\theta(\overline{D(\theta)}|p,q,r,\overline{\theta}) - Eu_i^\theta(D(\theta)|p,q,r,\overline{\theta})\right] - c \end{aligned}$$

By simple algebra, the above expression equals $\Pr(\overline{\theta})(2\overline{\theta}P\theta_{-i} - 1) - c$. ■

This claim, along with claim 3, establishes that when there is search by type θ of player i, we must have $\Pr(\overline{\theta})(2\overline{\theta}P\theta_{-i} - 1) - c \geq 0$. If there is indifference between search and no search, then the inequality must be satisfied as an equality. Note that this depends on the strategy of the rival candidate.

9.1 Proof of Lemma 1

When the rival is known to be bad, the expected payoff of type G from using P is $BPG + 2(1 - BPG) - c$, and that from using message N is $2BPG + (1 - BPG) - c$. Thus, the net gain from using P instead of N is $(1 - 2BPG)$, which is strictly positive if $BPG < \frac{1}{2}$ and negative if $BPG > \frac{1}{2}$. This proves the first part of (a). When the rival is unknown, the payoff to type G from message P is $\alpha\{GPG + 2(1 - GPG)\} + (1 - \alpha)\{BPG + 2(1 - BPG)\}$ and that from playing N is $\alpha\{(1 - GPG)\} +$

$(1-\alpha)\{2BPG+(1-BPG)\}$, and the gain from playing P instead of N is $[1-2(1-\alpha)BPG]$, and the second part of (a) follows. The proof of (b) is similar.

With the four basic claims and Lemma 1, we can find the equilibrium, i.e. prove proposition 1.

9.2 Proof of proposition 1

Part (i): $\alpha < c$

By claim 4, $\alpha < c \Rightarrow EU_i^B(S|p,q,r) < EU_i^B(NS,N|p,q,r)$

By claim 3, $1-\alpha > c \Rightarrow EU_i^B(S|p,q,r) > EU_i^B(NS,P|p,q,r)$

Thus, for type B of either player, using the negative message and not searching strictly dominates other strategies.

Hence, $p(B)=q(B)=0 \Rightarrow BPG=0$. By Lemma 1(a), $q(G)=r(G)=1$

Also, $BPG=0$ implies

$$EU_i^G(S|p,q,r) - EU_i^G(NS,P|p,q,r) = -c < 0 \Rightarrow p(G)=0.$$

Part (iii): $\alpha > 1-c$

By claim 4, $1-\alpha < c \Rightarrow EU_i^G(S|p,q,r) < EU_i^G(NS,P|p,q,r)$

By claim 3, $\alpha > c \Rightarrow EU_i^G(S|p,q,r) > EU_i^G(NS,N|p,q,r)$

Thus, for type G of either player, using the positive message while not searching strictly dominates other strategies.

Hence, $p(G)=0$ and $q(G)=1 \Rightarrow GPB=1$. By Lemma 1(b), $q(B)=r(B)=1$

By claim 2, $r(B)=1 \Rightarrow p(B)=0$

Part (ii): $c < \alpha < 1-c$

Here, $\min(\alpha, 1-\alpha) > c$. Therefore, by claim 3, for $\theta \in \{G,B\}$,

$EU_i^\theta(S|p,q,r) - EU_i^\theta(NS,\overline{D(\theta)}|p,q,r) > 0.$

Thus, whenever the action NS is played by type θ with positive probability, $D(\theta)$ is the message employed, or $p(G)<1 \Rightarrow q(G)=1$ and $p(B)<1 \Rightarrow q(B)=0$

Next, we claim that we cannot have search with certainty for either type.

Claim 5 *We cannot have $p(\theta)=1$ in equilibrium for $\theta \in \{G,B\}$.*

Proof. It has already been established that $p(\theta) = 0$ in equilibrium for $\theta \in \{G, B\}$ as long as $\alpha < c$ or $\alpha > 1 - c$. We need to prove this claim only for the case $c < \alpha < 1 - c$. Suppose $c < \alpha < 1 - c$ and $p(G) = 1$ for i. Therefore, $GPB_i = r(G)$, and by Claim 2, $r(G) = 0$. This implies by Lemma 1(b) that for $-i$, $r(B) = q(B) = 0 \Rightarrow BPG_{-i} = 0$ (by Lemma 1). Using this in claim 3, we have a contradiction, since

$$EU_i^G(S|p,q,r) - EU_i^G(NS,P|p,q,r) = -c < 0 \Rightarrow p(G) = 0.$$

Now, suppose $c < \alpha < 1 - c$ and $p(B) = 1$ for i. Therefore, $BPG_i = r(B) = 1$. Using this in claim 3, we have a contradiction, since

$$EU_i^G(S|p,q,r) - EU_i^G(NS,P|p,q,r) = (1-\alpha) - c > 0 \Rightarrow p(G) = 1$$

and we have just shown that we cannot have $p(G) = 1$ in equilibrium. ■

Equation (15) and claim 5 together establish that in equilibrium with $c < \alpha < 1 - c$, we must have $q(G) = 1$ and $q(B) = 0$. Using this in claim 4 and Lemma 1(b), we can show similarly that neither type will have $p(\theta) = 0$. Therefore, we must have $p(\theta) \in (0, 1)$ for both types and for both players. By claim 2, $r(B) = 1$ and $r(G) = 0$.

Also, from claim 4, we must have for indifference between search and no search for type θ,

$$\Pr(\overline{\theta})(2\overline{\theta}P\theta - 1) = c \Rightarrow \overline{\theta}P\theta = \frac{1}{2}\left(1 + \frac{c}{\Pr(\overline{\theta})}\right)$$

Since $BPG = p(B)$ and $GPB = 1 - p(G)$, we are done.

9.3 Proof of Proposition 3

Since $p(G) = 1$, $GPB = r(G)$. From claim 3,

$$EU_i^B(S|p,q,r) - EU_i^B(NS,P|p,q,r) = (1-\alpha) - c > 0 \tag{15}$$

Thus, along the equilibrium path, we must have $q(B) = 0$, and hence, $BPG = p(B)r(B)$. If $BPG > \frac{1}{2}$, $r(G) = 0$ by Lemma 1(a). Therefore, $GPB = 0 \Rightarrow r(B) = 0$ by Lemma 1(b). As this implies $p(B)r(B) = 0$, it is a contradiction.

On the other hand, if $BPG < \frac{1}{2}$, $r(G) = 1$ by Lemma 1(a). Therefore, $GPB = 1 \Rightarrow r(B) = 1$ by Lemma 1(b). Also, by claim 4,

$$EU_i^B(S|p,q,r) - EU_i^B(NS,N|p,q,r) = \alpha - c > 0. \tag{16}$$

From (16) and (17), we must have $p(B) = 1$. Since $r(B) = 1$, we have $p(B)r(B) = 1$, which is again a contradiction.

Therefore, we must have $p(B)r(B) = \frac{1}{2}$. Since $p(B) > 0$, we must have $r(B) = 1$, by claim 2. Therefore, $p(B) = \frac{1}{2}$. Also, $p(B) \in (0,1)$ implies that $EU_i^G(S|p,q,r) = EU_i^G(NS,P|p,q,r)$, which, by claim 4, implies that $r(G) = \frac{1}{2}\left(1 + \frac{c}{\alpha}\right)$.

9.4 Proof of Lemma 3

Suppose the assertion is false, and we have $\mu(\theta_{-i}|\theta_i = G) = G$. This must be on the equilibrium path due to assumption (5) on off-equilibrium beliefs. Note that $v_i(G) = \frac{1}{2}$. This implies that irrespective of the message profile, whenever two good types meet, they both get a debate utility of 1. This implies that $X = S$ is strictly dominated. To see that, suppose the good type of a candiate plays a mix m_B of messages against the bad type and m_G of messages against the good type conditional on search. Since any strategy against the good type of the rival fetches a payoff of 1, the candidate can be better off (by an amount c) by not searching and employing a mix m_B.

Now, for $\mu(\theta_{-i}|\theta_i = G) = G$ to be true we need the good type to have fruitful debate with good type with positive probability, i.e. $GPG \in (0,1)$ and type B never to reveal type G, which happens when either $BPG = 1$ or $GPB = 0$.

Since $p(G) = 0$, $GPG = GPB = q(G)$. If $GPG > 0$, we can rule out $GPB = 0$.

If $BPG = 1$, we must have $q(G) = 0$, i.e. $GPG = 0$, which is a contradiction.

Thus, it is established that we cannot have $\mu(\theta_{-i}|\theta_i = G) = G$ in equilibrium. In the same fashion, we can show that we cannot have $\mu(\theta_{-i}|\theta_i = B) = B$ in any equilibrium with rational voting.