Modelling the effect of campaign advertising on US presidential elections

Maria Gallego* and Norman Schofield†
Wilfrid Laurier University and Washington University in Saint Louis

November 17, 2015

Abstract

We provide a stochastic electoral model of the US Presidential election. The availability of smart phone, social media coupled with large data on voters' personal information allows candidates to send targeted messages directly to voters taking their characteristics into account. In our model, candidates directly communicate with voters.

Prior to the election candidates announce their policies and advertising campaigns. Voters care about candidates' policies relative to their ideal, about messages candidates sent relative to their ideal message frequency, called the campaign tolerance level. The electoral mean is a strick (weak) local Nash equilibrium (LNE) of the election if the expected vote shares of all candidates are greater than the sufficient (necessary) pivotal vote shares. The sufficient pivotal vote share rises when voters give greater weight to the policy or advertising dimensions. The necessary pivotal vote share may increase or decrease in the importance votes give to the policy or advertising dimensions. If the expected vote share of at least one candidate is lower than the necessary pivotal vote share, then the electoral mean is not a LNE of the election.

Key words: stochastic vote mode, valence, local Nash equilibrium, pivotal probabilities, pivotal vote shares, state versus national campaigns

*Department of Economics, Wilfrid Laurier University, 75 University Avenue West, Waterloo, Canada N2L 3C5 (mgallego@wlu.ca). Tel: 519.884.0710 Ext 2635

†Department of Political Science, Washington University in St. Louis, Campus Box 1027, One Brookings Drive, St. Louis, MO 63130-4899.

1 Introduction

The collection of articles included in Hendricks and Kaid's (2011) edited volume *Techno Politics in Presidential Campaigning: New Voices, New Technologies and New Voters*, argue that the 2008 US presidential election was a landmark campaign because it elected the first African-American President and brought about dramatic changes in the way presidential campaigns would be conducted from 2008 onwards. We argue that the emergence of smart phones, social media[1] coupled with large data sets on voters' personal characteristics and public and private pre-election polls has had a profound effect on political campaign advertising and on US Presidential elections as new media and new technologies give voice to new candidates and voters. The congruence of these factors has revolutionized the way electoral campaigns are conducted. In particular, candidates now tailor their campaign message to voters taking their distinguishing sociodemographic characteristics and in their political preferences into account. Campaign strategists are now able to identify the core supporters of each candidate and undecided voters.

Ansolabehere et al (2003) examine campaign contributions in Congressional and Presidential elections and note that candidates, parties and organizations raised and spent about $3 billion in the 1999-2000 election cycle. Since the federal government, at that time, spent about $2 trillion, the prize from influencing politics was of considerable value. They argue that the reason so "little" is spent campaigning is that contributions are a consumption, rather than an investment, good. However, they note that the electoral motive is not insignificant as the marginal impact of $100,000 spent in a House race, ceteris paribus, is about 1% in vote. Thus, highlighting the powerful influence of ad campaigns on elections.

Moreover, by being able to directly communicate with potential supporters, candidates have greater control over their message than when using mass media (TV and radio). This highlights that models of US presidential elections must now take into account that candidates can directly communicate with voters. Political campaigns now use different means of communicating with voters. Not only have candidates used email, pod casts, RRS feeds and cell phone texting to communicate with voters but more recently new technologies such as blogs, YouTube, Twitter, Facebook have allowed candidates to sent messages directly to voters' smart phones and social accounts. The model developed in this paper assumes that candidates directly communicate with voters and that voters have preferences over the messages sent by candidates.

Voter judgments about candidates' competence are modeled by the notion of *valence*. The formal model can then be linked to Madison's understanding of the nature of the choice of Chief Magistrate (Madison, 1999 [1787]). Schofield (2002) suggests that Madison's argument on the "extended Republic" may have been influenced by Condorcet's "Jury Theorem" (Condorcet, 1994 [1785]) as it is based on the notion of electoral judgment rather than preference. Recent models involving valence contribute to a Madisonian conception of elections in representative democracies

[1] Barak Obama was he first candidate to use social media in his 2008 bid for the US presidency. Obama ran a very successful fund raising campaign with a portion of his funds having been donated by voters contributing small amounts through social media. Obama also used social media to contact voters directly during the campaign.

as methods of aggregation of both preferences and judgments.[2]

The standard spatial model assumes that it is only candidate *positions* that matter to voters. Within the context of the spatial model, controversy has arisen over whether rational candidates will converge to an electoral center, as suggested by Downs (1957) or whether elections are fundamentally unstable, as argued by Riker (1980, 1982, 1986).[3] However, as Stokes (1963, 1992) emphasizes, the non-policy evaluations, or *valences*, of candidates by the electorate are just as important as their policy preferences. Clarke et al (2009) compare a "Downsian" or spatial model of the 2000 US presidential election with a valence model of the same election, based on voters' perceptions of candidates' traits and find that "the two models have approximately equal explanatory power." As Sanders et al (2011) comment, valence theory is based on the assumption that "voters maximize their utilities by choosing the party that is best able to deliver policy success." Valence measures the *bias* in favor of one or other of the party leaders (McKelvey and Patty, 2006). From our work, we argue that neither the Downsian convergence result nor the instability results gives an accurate picture of democratic elections. Instead, both position and valence matter in a fundamental way.

We extend Schofield's (2007) electoral model to allow candidates to use two instruments–the policy platform and advertising (ad) campaigns–to maximize vote shares. Candidates' electoral campaign consists of their policies and ad campaigns. We assume that the effectiveness of the advertising campaigns varies across voters as they not only have policy preferences but also have preferences over the ad campaign messages send by candidates. In our model, voters are characterised by their preferred policy and campaign tolerance level,[4] with campaign tolerances varying across voters to capture differences voters' desire to be contacted by candidates.

The voting literature has also identified that voters assess candidates in various ways. Voters make decisions examining how close each candidate's policy is from theirs and take into account candidates' traits (popularity, charisma, age, gender, etc.), the *traits* valence; as well as their beliefs on candidates' ability to govern, the *competency valence*. It is also well known that voters with certain sociodemographic characteristics are also more prone to voting for one of the candidates,[5] the *sociodemographic valence*. Recent research has found evidence that differences in voting preferences depend not only on candidates' policy positions but also on the leaders traits and the sociodemographic and competency valences. Empirical findings show that the sociodemographic and competence valences are a major determinant of the electoral outcome (Schofield *et al.* 2011a,b; Gallego and Schofield 2014, 2015). There is evidence that voters' choice of candidate is influenced by voters' non-policy evaluation of candidates and by the distance between candidates' policy and voters' ideal policy (Schofield and Gallego, 2011). We incorporate the sociodemographic, traits and competence valences into voters' utility functions in our model.

[2] See Ansolabehere & Snyder, 2000; Groseclose, 2001; Aragones & Palfrey, 2002; Schofield, 2006; Zakharov, 2009.

[3] See McKelvey (1976); Schofield (1978); Saari (1997); Austen Smith and Banks (1999).

[4] A voter's campaign tolerance is given by the ideal messages a voter wants to receive from candidates.

[5] For example, African-American voters are much more likely to vote for the Democratic candidate than to vote for the Republican candidate. Thus, Democratic candidates have a higher average sociodemographic valence among African-American voters than do Republican candidates.

We assume that voters' sociodemographic, traits and competence valences vary around their mean values with the idiosyncratic component drawn from Type-I Extreme value Distributions. Since candidates do not observe voters' idiosyncratic valence components but know the distribution of these components, candidates view voters' utility as stochastic and cannot perfectly anticipate how each individual votes but can estimate their expected vote shares. The model examines how voters' preferences and characteristics affect candidates' policy and ad campaigns and how candidates' choices affect their electoral prospects.

Our results show that if candidates adopt the same electoral campaign, i.e., the same policy and ad campaign, they adopt the electoral mean[6] as their campaign strategy. We derive the conditions under which candidates converge to the electoral mean. The electoral mean (z_m, a_m) is a strong (weak) local Nash equilibrium (LNE) of the election if candidates' expected vote share at (z_m, a_m), V_j^m, is high enough, i.e., higher that the sufficient (necessary) pivotal vote share, $\mathcal{V}_{sp}^m$ $(\mathcal{V}_{np}^m)$. When candidate j's expected vote share at (z_m, a_m) is less than the necessary pivotal vote share, i.e., if $V_j^m < \mathcal{V}_{np}^m$, j's vote share is at a minimum or at a saddle point and j wants to move away from the electoral mean to increase its vote share. In this case, other candidates may also find it in their interest to move away from the electoral mean. The electoral mean is not a LNE of the election.

Suppose the electoral mean is a local Nash equilibrium (LNE) of the election. In equilibrium, candidates take voters' policy and campaign tolerances levels into account. The policy and advertising campaigns are the weighted average of voters' ideal policy and campaign tolerance levels. Even though we assume that the one-person-one-vote principle applies, candidates weight voters differently in their electoral campaign. In particular, candidates give little or no weight to to their core supporters or to voters who care not at all for them, giving maximal weight to undecided voters, i.e., voters who vote with probability close to a half for the candidate. Thus, candidates' equilibrium campaign reflects the policies and the campaign tolerance levels of undecided voters.

Our comparative statics results show that when voters give more importance to the policy or the advertising dimensions, that the sufficient pivotal vote share increases thus making it more difficult for the electoral mean to be a strict LNE of the election. Whether the necessary pivotal vote share increases or decreases depends on the relative variance of voters' ideals in one dimension compared to the aggregate variance of voters' preferences around the electoral mean.

Recent literature focuses on the effects of campaign expenditure on US elections. Meirowitz (2008) develops a model in which candidates select their campaign effort (time, energy and money) to win elections and shows that marginal asymmetries in costs or technology can explain incumbency advantage ignoring the policy side of the election. Herrera et al (2008) find that greater volatility in voters' policy preferences forces the two parties to spend more on the election. Ashworth and Bueno de Mesquita (2007) assume that candidates buy valence to increase their electoral chances. In our model, asymmetries between candidates are generated by the effectiveness of the policy and ad campaigns on each voter; moreover, the volatility of voters' policy and campaign preferences affect the pivotal vote shares and this affects the LNE of the election and thus the electoral outcome.

[6] The electoral mean is defined by the mean of voters' ideal policies and campaign tolerance levels.

The ad messages sent by candidates influence voters' choices but this effect varies across voters and so cannot be considered a valence.

Section 2 describes the preliminaries of the model. The stochastic electoral model of the election is presented in Section 3. Candidates' best response functions are derived in Section 5, a discussion of the common electoral campaign is presented in Section 6 with the conditions under which there is convergence to the electoral mean derived in Section 7. Comparactive staics– Section 8–shows how the parameters of the model affect convergence to the electoral mean. Final comments are given in Section 9. The derivation of the second order conditions for convergence and of the comparative statics are given in the Appendix in Section 10.

2 Preliminaries

We develop a stochastic electoral model in which at least two candidates compete in the US presidential election (e.g., Obama, McCain and Nader ran in 2000 election). We model an election where candidates' *electoral campaigns* consist of their policy platform and advertising (or ad) campaign. Since voters perfectly observe candidates' policy announcements, the advertising potion of the campaign serves *only* to convince voters that they matter to candidates. The ad campaign consists of messages[7] candidates send to voters. We assume that as messages increase so do the campaign advertising costs regardless of the effectiveness of the ad campaign.[8] We study candidates' electoral campaign game in response to the anticipated electoral outcome.

Let $\mathcal{C} = \{1, .., j, .., c\}$ be the set of candidates competing in the election. Prior to the election, candidates simultaneously announce their electoral campaign. Candidate j's electoral campaign, for all $j \in \mathcal{C}$, consists of a policy, $z_j \in \mathcal{Z} \subseteq \mathbb{R}$, and ad messages, $a_j \in \mathcal{A} \subseteq \mathbb{R}$. Denote by $\mathcal{K}_j = \mathcal{Z} \times \mathcal{A}$ each candidate' electoral campaign space, and by $\mathcal{K} = \prod_{j \in \mathcal{C}}(\mathcal{K}_j)$ the campaign space of all candidates.

Let $\mathbf{z}$ and $\mathbf{a}$ be the electoral campaign profile of all candidates, i.e.,

$$\mathbf{z} \equiv (z_1, ..., z_c) \quad \text{and} \quad \mathbf{a} \equiv (a_1, ..., a_c). \tag{1}$$

The vector $\mathbf{z}_{-j}$ (respectively $\mathbf{a}_{-j}$) denotes the profile where all policies and ad messages *except j's policy* (*ad message*) are held constant. Let n and $\mathcal{N} = \{1, .., i, .., n\}$ represent the number and the set of voters in the country.

The utility voters derive from each candidate depends on candidates' policies and advertising campaigns. Voter i's utility is given by the vector of utilities, one for each candidate, i.e.,

$$u_i(\mathbf{z}, \mathbf{a}) = (u_{i1}(z_1, a_1), .., u_{ij}(z_j, a_j), ..., u_{ic}(z_c, a_c))$$

[7] Messages are measured in continuous, rather than discrete, terms so that messages of different lengths can be compared. If message 1 is twice as long as message 2, then message 1 is assigned a number twice as large as message 2. In this paper, we assume messages are one-dimensional.

[8] The other implicit assumption here is that candidates have sufficient funds to finance their advertising campaigns. We leave it to another paper to examine the effect that activists and their donations have on electoral campaigns.

where the utility i derives from candidate j, $u_{ij}(z_j, a_j)$, is given by

$$\begin{aligned} u_{ij}(z_j, a_j) &= -\beta(x_i - z_j)^2 - e(t_i - a_j)^2 + (\alpha_j + \xi_{ij}) + (\tau_j + \varsigma_{ij}) + (\lambda_j + \epsilon_{ij}) \\ &= u^*_{ij}(z_j, a_j) + \xi_{ij} + \varsigma_{ij} + \epsilon_{ij}. \end{aligned} \tag{2}$$

Voter i is characterised by her ideal policy $x_i \in \mathcal{Z}$ and by the messages she ideally wants to receive from any candidate, her *campaign tolerance level* $t_i \in \mathcal{A}$ with the (x_i, t_i) pair drawn from a joint distribution $\Gamma_x \times \Gamma_a$. Note that the dependence of u_{ij} on (x_i, t_i) is taken as understood and that u^*_{ij} in (2) measures the *observable* component of u_{ij}. The coefficients β and e in (2) are positive, the same for all voters, common knowledge and measure the importance voters give to j's policy and ad campaign. Voters policy preferences, $-\beta(x_i - z_j)^2$ in (2), says that the farther j's policy z_j is from i's ideal policy x_i, the lower the utility i derives from candidate j.

With policies being perfectly observable, the only reason candidates advertise in this model is to give voters a *further* impetus to vote for them and do so through ad messages. Since candidates directly contact voters, we assume voters have preferences over how often they want to be contacted by candidates, i.e., by their campaign tolerance level. Voters judge candidates' messages by examining the message frequency compared to their campaign tolerance level and since from the point of view of each voter there can be too much or too little advertising, we model voters as having quadratic preferences over ad campaigns. The *effectiveness of j's ad campaign on voter i*, $e(t_i - a_j)^2$ in (2), depends on the importance voters give to candidates' ad campaigns, e, *and* on the messages j sends, a_j, relative to i's campaign tolerance t_i.[9] When $a_j > t_i$, i believes j engaged in too much advertising, an irritant to voter i that leads to campaign fatigue[10] and if $a_j < t_i$, i believes that j's ad campaign reflects j's lack of concern for voters.

Voters sociodemographic characteristics affect their choice of candidate.[11] Voter i's *sociodemographic valence* for candidate j is modelled by $\alpha_j + \xi_{ij}$ in (2). Candidate j's *mean sociodemographic valence*, α_j, captures the idea that voters with similar sociodemographic characteristics (gender, age, class, education, financial situation, etc.) share a common evaluation or bias for j. Voters are identified by a d-vector $\boldsymbol{\eta}_i$ denoting their individual sociodemographic characteristics with mean $\boldsymbol{\eta} = \frac{1}{n}\sum_{i\in\mathcal{N}} \boldsymbol{\eta}_i$. The importance sociodemographic characteristics have on the utility i's derives from j are modelled by a d-vector $\boldsymbol{\gamma}_j$ *common* to all voters. The composition $\alpha_j = \{(\boldsymbol{\gamma}_j \cdot \boldsymbol{\eta})\}$ measures the *mean* sociodemographic valence for j. The effect that i's sociodemographic characteristics have on i's choice of candidate has an idiosyncratic component, ξ_{ij}, that varies around α_j according to a Type-I extreme value distribution, $E(0, \frac{\pi}{6})$, with mean zero and variance $\frac{\pi}{6}$.

The effect of candidate j's *traits* (race, gender, age, charisma, experience in public life, etc.) on voter i's choice is modelled by $\tau_j + \varsigma_{ij}$ in (2). The commonly known *mean trait valence* for j, τ_j, measures the mean influence j's traits have on voters' choice with the effect on voter i also

[9] Since advertising costs rise with the message frequency, i's campaign tolerance indirectly measure i's preferred campaign spending.

[10] As happens when voters are too frequently contacted by robo-calls or twitts.

[11] See footnote 5 on how African-Americans vote in the US.

depending an idiosyncratic component, ς_{ij}. Candidates are identified by a t-vector $\mathbf{t}_j$ capturing their individual traits. The *average* importance voters give to these traits is given by a t-vector $\boldsymbol{\varpi}_j$ *common* to all voters. The composition $\tau_j = \{(\mathbf{t}_j \cdot \boldsymbol{\varpi}_j)\}$ measures the mean effect that j's traits have on voters. We assume that voters' beliefs on the effect that candidates traits have on their utility are *not* observable. Voter i's private component , ς_{ij}, varies around τ_j and is drawn from a Type-I extreme value distribution, $T(0, \frac{\pi}{6})$, with zero mean and variance $\frac{\pi}{6}$.

The term $\lambda_j + \epsilon_{ij}$ in (2) represents i's belief on candidate j's competence. The *mean competence valence*, λ_j, measures the *common belief* voters have on j's ability to govern. The private portion of i's competence signal, ϵ_{ij}, varies around λ_j and is drawn from a Type-I extreme value distribution, $F(0, \frac{\pi}{6})$ with mean zero and variance $\frac{\pi}{6}$.

We assume that the three idiosyncratic components in voters' utility are drawn independent of each other. Moreover, Train (2003) has shown that the sum of Type-I Extreme Value Distributions all with the same mean and variance has a Type I Extreme Value Distribution. Thus, since $\xi_{ij} \sim E(0, \frac{\pi}{6})$, $\varsigma_{ij} \sim T(0, \frac{\pi}{6})$ and $\epsilon_{ij} \sim F(0, \frac{\pi}{6})$ are Type I Extreme Value Distributions then their sum $\varkappa_{ij} = \xi_{ij} + \varsigma_{ij} + \epsilon_{ij} \sim \mathcal{X}(0, \frac{\pi}{6})$ also has a Type I Extreme value Distribution. Under these circumstances, voter i's utility in (2) can be re-written as follows:

$$\begin{aligned} u_{ij}(z_j, a_j) &= -\beta(x_i - z_j)^2 - e(t_i - a_j)^2 + \alpha_j + \tau_j + \lambda_j + \varkappa_{ij} \\ &= u^*_{ij}(z_j, a_j) + \varkappa_{ij}. \end{aligned} \tag{3}$$

Since candidates do not observe the idiosyncratic sociodemographic, traits and competence components of voters' utilities but know that these components are drawn from Type I Extreme Value Distributions, voters' utilities are stochastic. Thus, given policies and advertising levels $(\mathbf{z}, \mathbf{a})$ and the distributions E, T, F and $\mathcal{X}$, the probability that voter i chooses candidate j is given by

$$\rho_{ij}(\mathbf{z}, \mathbf{a}) = \Pr[u_{ij}(z_j, a_j) > u_{ih}(z_h, a_h), \text{ for all } h \neq j \in \mathcal{C}]$$

where Pr is the probability operator generated by the distribution assumptions on E, T, F and $\mathcal{X}$ and where $u_{ij}(z_j, a_j)$ is given by (3). The probability that i votes for j is given by the probability that $u_{ij}(z_j, a_j) > u_{ih}(z_h, a_h)$ for all $j, h \in \mathcal{C}$, i.e., given by the probability that i gets a higher utility from j than from any other candidate.

Since the sociodemographic, traits and competence signals and their sum follow Type-I extreme value distributions, the probability that i votes for j has a logit specification,

$$\rho_{ij} \equiv \rho_{ij}(\mathbf{z}, \mathbf{a}) \equiv \frac{\exp[u^*_{ij}(z_j, a_j)]}{\sum_{h=1}^{c} \exp[u^*_{ih}(z_h, a_h)]} = \left[\sum\nolimits_{h=1}^{c} \exp[u^*_{ih}(z_h, a_h) - u^*_{ij}(z_j, a_j)]\right]^{-1} \tag{4}$$

for all $i \in \mathcal{N}$ and $j \in \mathcal{C}$ with the dependence of ρ_{ij} on $(\mathbf{z}, \mathbf{a})$ sometimes omitted.

Candidates' objective is to maximize their vote shares. Since candidates do not observe the idiosyncratic components of voters' utility, they cannot perfectly anticipate how voters will vote but can estimate their *expected* vote shares. Candidate j's *expected vote share* is the average of the voting probabilities of all voters in the country, i.e.,

$$V_j(\mathbf{z}, \mathbf{a}) \equiv \frac{1}{n} \sum\nolimits_{i \in \mathcal{N}} \rho_{ij}(\mathbf{z}, \mathbf{a}) \tag{5}$$

where ρ_{ij} is given by (4) with the expected vote shares adding to 1, $\sum_{j \in \mathcal{C}} V_j(\mathbf{z}, \mathbf{a}) = 1$.

Candidate j's chooses its policy and ad campaign to maximize its expected vote share, i.e.,

$$\max_{z_j \in \mathcal{Z}, a_j \in \mathcal{A}} V_j(\mathbf{z}, \mathbf{a}) \quad \text{for all} \quad j \in \mathcal{C}.$$

3 The Stochastic Electoral Model

The timing of events in this stochastic election model is as follow:

1. Candidates simultaneously announce their policy platforms and advertising campaigns.
2. After observing platforms and advertising levels, each voter observes the idiosyncratic components of her utility function.
3. The election takes place.
4. The President elect implements the announced policy platform.

We now formally define the election model.

Definition 1 *The stochastic election model can be represented by a game in normal form* $\Delta\langle \mathcal{C}, \mathcal{K}, \Gamma_x \times \Gamma_a, \mathbf{V} \rangle$*, where*

1. ***Players**: $\mathcal{C} = \{1, .., j, .., c\}$ is the set of candidates competing in the election.*
2. ***Strategies**: $\mathcal{K}_j = \mathcal{Z} \times \mathcal{A}$ denotes the electoral campaign space of candidate j with $\mathcal{Z}$ denoting the policy space and $\mathcal{A}$ the ad messaging space, and where the electoral campaign space of all candidates is given by $\mathcal{K} = \Pi_{j \in \mathcal{C}}(\mathcal{K}_j)$.*
3. ***Voters**: Voters are characterized by their ideal policy and campaign tolerance level, $(x_i, t_i) \in \mathcal{Z} \times \mathcal{A}$ for $i \in \mathcal{N}$ and by the mean sociodemographic, traits and competence valences, $(\alpha_j, \tau_j, \lambda_j)$. Voters' ideal policies and campaign tolerance levels are drawn from the joint distribution $\Gamma_x \times \Gamma_a$. The utility voter i derives from candidate j is given by (2) and subject to three random shocks generated by Type-I Extreme Value Distributions, $E(0, \frac{\pi}{6})$, $T(0, \frac{\pi}{6})$ and $F(0, \frac{\pi}{6})$ whose sum $\mathcal{X}(0, \frac{\pi}{6})$ also follows a Type-I Extreme value distribution.*
4. ***Candidates' Payoff Functions**: For any candidate $j \in \mathcal{C}$, j's expected vote share function is given by $V_j : \mathcal{K}_j \to [0, 1]$ in (5). Let $\mathbf{V}(\mathbf{z}, \mathbf{a}) = ((V_j(\mathbf{z}, \mathbf{a})$ for all $j \in \mathcal{C})$ represent the payoff function profile of all candidates.*

Let us find the *local Nash equilibria* (LNE) of this stochastic election model.

4 The equilibrium of the stochastic election model

The electoral campaign profile $(\mathbf{z}^*, \mathbf{a}^*)$ is a local Nash equilibrium (LNE), if when candidates choose their policies in $\mathbf{z}^*$ and ad campaigns in $\mathbf{a}^*$, each candidate is at a local maximum in its expected vote share functions. That is, the vector $(\mathbf{z}^*, \mathbf{a}^*)$ is such that no candidate may either shift its policy *or* advertising level by a *small* amount—ceteris paribus—to increase its vote share. Formally,

Definition 2 *A strict (weak) local Nash equilibrium of the stochastic electoral campaign model* $\Delta\langle \mathcal{C}, \mathcal{K}, \Gamma_x \times \Gamma_a, \mathbf{V}\rangle$ *is a vector of candidate policies and advertising campaign levels,* $(\mathbf{z}^*, \mathbf{a}^*)$ *such that for each candidate* $j \in \mathcal{C}$*:*

1. *There exists a small neighborhood of* z_j^**,* $B(z_j^*) \subset \mathcal{Z}$*, such that*

$$V_j(\mathbf{z}^*, \mathbf{a}^*) > (\geqslant) V_j(z_j', \mathbf{z}_{-j}^*, \mathbf{a}^*) \quad \textit{for all} \quad z_j' \in B(z_j^*) - \{z_j^*\}.$$

2. *There exists a small neighborhood of* a_j^**,* $B(a_j^*) \subset \mathcal{A}$*, such that*

$$V_j(\mathbf{z}^*, \mathbf{a}^*) > (\geqslant) V_j(\mathbf{z}^*, a_j', \mathbf{a}_{-j}^*) \quad \textit{for all} \quad a_j' \in B(a_j^*) - \{a_j^*\}.$$

Since the observable component of voter i's utility for candidate j, $u_{ij}^*(z_j, a_j)$ in (2) for all $i \in \mathcal{N}$, and the distributions of voters' ideal policies and campaign tolerance levels, $\Gamma_x \times \Gamma_a$, are continuously differentiable, then so are ρ_{ij} in (4) and $V_j(\mathbf{z}, \mathbf{a})$ in (5). We can then estimate how j's expected vote share changes if j *marginally* adjusts its policy *or* ad campaign, ceteris paribus.

Remark 1 *If in Definition 2 we can substitute* $\mathcal{Z}$ *for* $B(z_j^*)$ *and* $\mathcal{A}$ *for* $B(a_j^*)$*, then a local Nash equilibrium is also a pure strategy Nash equilibrium (PNE) of the election.*

Remark 2 *In political models, the critical equilibrium may be characterized by positive eigenvalues for the Hessian of one of the political parties. As a consequence the expected vote share functions of such a candidate fails pseudo-concavity. Therefore, none of the usual fixed point arguments can be used to assert existence of a "global" pure Nash equilibrium (PNE). For this reason, we use the concept of a "critical Nash equilibrium" (CNE), namely a vector of strategies which satisfies the first-order condition for a local maximum of candidates' expected vote share functions. Standard arguments based on the index, together with transversality arguments can be used to show that a CNE will exist and that, generically, it will be isolated. A local Nash equilibrium (LNE) satisfies the first-order condition, together with the second-order condition that the Hessians of all candidates are negative (semi-) definite at the CNE. Clearly, the set of LNE will contain the PNE.*

The parameters $(\beta, e, \alpha_j, \tau_j, \lambda_j)$ for all $j \in \mathcal{C}$ in voter i's utility function in (2) as well as the distributions Γ_x and Γ_a of voters' ideal policies and campaign tolerance levels, x_i and t_i, are exogenously given in the model. Using these parameters and the distribution of voters' preferences, we can relate any vector of candidate policies and advertising levels, $(\mathbf{z}, \mathbf{a})$, to a vector of expected vote share functions, i.e., we can estimate the vector of expected voters shares $\mathbf{V}(\mathbf{z}, \mathbf{a})$ given by

$$\mathbf{V}(\mathbf{z}, \mathbf{a}) = (V_j(\mathbf{z}, \mathbf{a}) \text{ for all } j \in \mathcal{C}).$$

Using simulations we can determine if one of the LNE is a Nash equilibrium of the election.

Candidate j's expected vote share function is at a local maximum if the following *first* and *second* order conditions are satisfied. Candidate j's *first order necessary conditions (FONC)* determine j's *best response* policy and advertising functions. There are then $2 \times p$ best response functions and their associated $2 \times p$ *critical values*,[12] one for each policy and ad campaign for each candidate.

We then examine whether at these critical values the expected vote shares functions of each candidate are at a *maximum, minimum or a saddle point*. To do so, we use the *Hessian matrix of second order partial derivatives of the vote share functions* evaluated at these critical values as the Hessian determines the *local curvature* of the vote share functions at these critical values. The *sufficient* (*necessary*) second order condition (SOC) for j's expected vote share to be at a maximum at the critical values is that the Hessian be negative (semi-) definite at these critical values which happens only when the eigenvalues of the Hessian of the vote shares are all negative (non-positive) at these critical values.

Denote by z_j^C and a_j^C the *critical* (C) policy and advertising campaigns profiles of candidate j. The critical policy and campaign advertising profile of all candidates is given by

$$\mathbf{z}^C \equiv (\mathbf{z}_1^C, .., \mathbf{z}_j^C, .., \mathbf{z}_c^C) \quad \text{and} \quad \mathbf{a}^C \equiv (\mathbf{a}_1^C, .., \mathbf{a}_j^C, .., \mathbf{a}_c^C) \tag{6}$$

The profile $(\mathbf{z}^C, \mathbf{a}^C)$ represent a *prospect*[13] LNE of the election.

5 Candidates' best response functions (FONC)

To derive the first order necessary condition (FONC) of candidates' vote shares, we first examine how *voters* respond to changes in candidate j's policies or ad messages then examine how changes in positions and messaging levels affect candidates' *expected vote shares.*

The *marginal* impact of a change in j's policy or message, z_j or a_j, on the probability that i votes for j is given by *partial* derivative of ρ_{ij} in (4) with respect to (*wrt*) to each choice variable, ceteris paribus, i.e.,

$$\mathbf{D}\rho_{ij}(\mathbf{z}, \mathbf{a}) \equiv \begin{pmatrix} \frac{\partial \rho_{ij}}{\partial z_j} \\ \frac{\partial \rho_{ij}}{\partial a_j} \end{pmatrix} = 2\rho_{ij}(1 - \rho_{ij}) \begin{pmatrix} \beta(x_i - z_j) \\ e(t_i - a_j) \end{pmatrix}. \tag{7}$$

The effect of a marginal change in one of j's choice variables on ρ_{ij} depends on the *endogenously* determined probability that i votes for j, ρ_{ij} in (4), and for any other candidate, $(1 - \rho_{ij})$; on the

[12] Called critical as we do not yet know whether at these values the expected vote shares functions are at a maximum.

[13] We say *prospect* because we do not know yet if candidates' critical choices satisfy the SOC for a maximum.

importance voters give to j's choice variable β or e; and on how far i's ideal policy or tolerance level, x_j or t_j, is from j's corresponding choice z_j or a_j.

5.1 Candidates' best response functions

To find j's best response function take the partial derivative of j's vote share function in (5) wrt the corresponding variable, z_j or a_j and set it equal to zero, i.e.,

$$\mathbf{D}V_j(\mathbf{z}_j,\mathbf{a}_j) = \frac{1}{n}\sum\nolimits_{i\in\mathcal{N}}\mathbf{D}\rho_{ij} = \frac{1}{n}\sum\nolimits_{i\in\mathcal{N}} 2\rho_{ij}(1-\rho_{ij})\begin{pmatrix}\beta(x_i - z_j)\\ e(t_i - a_j)\end{pmatrix} = 0 \tag{8}$$

where the third term follows from (7) and ρ_{ij} is given by (4). This FONC is satisfied when

$$\sum\nolimits_{i\in\mathcal{N}}\rho_{ij}(1-\rho_{ij})\begin{pmatrix}x_i - z_j\\ t_i - a_j\end{pmatrix} = 0.$$

After some manipulation j's *best response functions* are then given by

$$\begin{pmatrix}z_j^C\\ a_j^C\end{pmatrix} = \sum\nolimits_{i\in\mathcal{N}}\mu_{ij}^C\begin{pmatrix}x_i\\ t_i\end{pmatrix} \quad \text{where} \quad \mu_{ij}^C \equiv \frac{\rho_{ij}^C(1-\rho_{ij}^C)}{\sum_{i\in\mathcal{N}}\rho_{ij}^C(1-\rho_{ij}^C)}. \tag{9}$$

We use superscript C to indicate variables evaluated at $(\mathbf{z}^C,\mathbf{a}^C)$ and in short hand use $\mu_{ij}^C \equiv \mu_{ij}(\mathbf{z}^C,\mathbf{a}^C)$ and $\rho_{ij}^C \equiv \rho_{ij}(\mathbf{z}^C,\mathbf{a}^C)$ in (9).

The result in (9) states that j's best response policy or ad messaging functions, (z_j^C, a_j^C), are a *weighted* average of voters' ideal policy or campaign tolerance levels where the *endogenously* determined *weight* candidate j gives to *voter* i is given by μ_{ij}^C in (9). The weight μ_{ij}^C depends on how likely is i to vote for j, ρ_{ij} in (4) evaluated at $(\mathbf{z}^C,\mathbf{a}^C)$, ρ_{ij}^C, and to vote for any other candidate, $(1-\rho_{ij}^C)$, *relative* to all voters.[14] Since ρ_{ij}^C and μ_{ij}^C depend on $(\mathbf{z}^C,\mathbf{a}^C)$, so do j's best response functions, that is, z_j^C and a_j^C depend on the policy and advertising choices of all *other* candidates. Note that voters receive the *same* weight in j's policy or advertising campaign.

The weight μ_{ij}^C in (9) is non-monotonic in ρ_{ij}^C. To see this, denote the numerator of μ_{ij}^C by $W_{ij}^C \equiv \rho_{ij}^C(1-\rho_{ij}^C)$. When i votes for j with probability one ($\rho_{ij}^C = 1$ so that i is a *core* supporter) or when i votes for j with probability zero ($\rho_{ij}^C = 0$), then $W_{ij}^C = 0$ and j gives these two voters zero weight in its critical policy and ad campaign choices, i.e., $\mu_{ij}^C = 0$. When i votes for j with probability $\rho_{ij}^C = \frac{1}{2}$, i is *undecided*, and since $W_{ij}^C = \frac{1}{4}$, j gives i the highest weight in its critical choices. Thus, even though the one-person-one-vote principle applies in this model, candidates weight voters differently in their electoral campaigns, e.g., $\mu_{ij}^C(\rho_{ij}^C = 0) = \mu_{ij}^C(\rho_{ij}^C = 1) = 0 < \mu_{ij}^C(\rho_{ij}^C = \frac{1}{2})$. In particular, j caters to undecided voters by giving them a higher weight in its best response functions.

Thus, (9) says that when holding $(\mathbf{z}_{-j}^C, \mathbf{a}_{-j}^C)$ constant, j's policy and advertising best response

[14] Suppose voters are equally likely to vote for j with probability $\rho_{ij} = v$. The weight j gives each voter is $\mu_{ij} = \dfrac{1}{n}$, as intuitively expected j weights each voter according to the inverse of the voting population.

functions, (z_j^C, a_j^C), are a weighted average of the ideal policies or campaign tolerance levels of voters where the weight j gives i along the policy and advertising dimensions is given by μ_{ij}^C. Note that when μ_{ij}^C varies across candidates, candidates adopt different campaigns.

Candidate j's critical ad campaign, a_j^C, determines the number of messages j sends to voters. We assume that developing and sending these messages cost money:[15] the more messages j's sends, the more money j spends on advertising.[16]

The weight j gives voters in its critical choices, μ_{ij}^C in (9), is endogenous since μ_{ij}^C is the crucial factor determining candidates' policies and ad campaigns *and* depends on the probability voter i chooses candidate j, ρ_{ij}^C *which in turn* depends on candidates' positions and advertising campaign levels. These weights change between elections as the parameters $(\beta, e, \alpha_j, \tau_j, \lambda_j)$ for all $j \in \mathcal{C}$ in voter i's utility function in (2) as well as the distributions Γ_x and Γ_a of voters' ideal policies and campaign tolerance levels, x_i and t_i for all $i \in \mathcal{N}$, vary across elections.

6 Is there a common campaign?

The objective is to find candidates' local Nash equilibrium (LNE) policies and advertising campaigns. We now examine whether candidates *adopt or converge to* a common campaign.

Definition 3 *Let (z_m, a_m) represent electoral mean, i.e., the mean of voters' ideal policies and campaign tolerance levels,*

$$\begin{pmatrix} z_m \\ a_m \end{pmatrix} \equiv \frac{1}{n} \sum\nolimits_{i \in N} \begin{pmatrix} x_i \\ t_i \end{pmatrix} \tag{10}$$

and define $(\mathbf{z}_m, \mathbf{a}_m) = ((z_m, a_m), .., (z_m, a_m))$ as the joint electoral mean for all candidates.

Lemma 1 *If candidates adopt a common campaign, they adopt the electoral mean as their electoral campaign strategy.*

Proof. The proof goes as follows: assume candidates adopt the same campaign. Given identical campaigns, show that since all voters vote with the same probability for each candidate they receive the *same* weight in candidates' critical policy and ad campaigns. Finally, show that candidates' best response is to adopt the electoral mean as their campaign strategy.

Suppose candidates adopt the same campaign, i.e., $(z_j, a_j) = (z^O, a^O)$ for $j \in \mathcal{C}$. The probability that voter i chooses candidate j, ρ_{ij}^O in (4) at (z^O, a^O) depends on the *difference* between the observable component of i's utility from candidates h and j which using (2) is given by

$$u_{ih}^{*O} - u_{ij}^{*O} \equiv u_{ih}^*(z^O, a^O) - u_{ij}^*(z^O, a^O) = (\alpha_h - \alpha_j) + (\tau_h - \tau_j) + (\lambda_h - \lambda_j) \tag{11}$$

since under identical campaigns the policy and ad campaign components cancel out.

[15] E.g., by buying personal data on voters that campaign strategists use to develop the message send to voters.

[16] Recall that j's campaign effectiveness is determined by voters' campaign tolerance and the importance voters give to the ad campaign [see (2)] and not by how much j spends on advertising.

Using (11), the probability that voter i chooses candidate j is then independent of candidates' policy and ad campaign since ρ_{ij} in (4) when evaluated at $(\mathbf{z}^O, \mathbf{a}^O)$ reduces to

$$\rho_j^O = \left[\sum_{h \neq j, h=1}^{c} \exp[u_{ih}^{O*} - u_{ij}^{O*}] \right]^{-1} = \left[\sum_{h \neq j, h=1}^{c} \exp[(\alpha_h - \alpha_j) + (\tau_h - \tau_j) + (\lambda_h - \lambda_j)] \right]^{-1}, \quad (12)$$

i.e., ρ_j^O depends only on the difference between the mean sociodemographic, traits and competence valences of the two candidates and not on their campaign strategies. In addition, (12) implies that voters choose j with *equal* probability as ρ_j^O does not depend on voters' ideal policies or campaign tolerance levels, i.e., $\rho_{ij}^O = \rho_j^O$ for all $i \in N$. In this case, ρ_j^O also represents j's expected vote share at $(\mathbf{z}^O, \mathbf{a}^O)$ since using (5), $V_j(\mathbf{z}^O, \mathbf{a}^O)$ is given by

$$V_j^O \equiv V_j(\mathbf{z}^O, \mathbf{a}^O) \equiv \frac{1}{n} \sum\nolimits_{i \in N} \rho_j^O = \rho_j^O. \quad (13)$$

With voters choosing j with the same probability, $\rho_{ij}^O = \rho_j^O$ for all $i \in N$ in (19), j gives voters the same weight in its policy and ad campaign as μ_{ij}^C in (9) reduces to

$$\mu^O \equiv \frac{\rho_j^O(1-\rho_j^O)}{\sum_{i \in N} \rho_j^O(1-\rho_j^O)} = \frac{1}{n}, \quad (14)$$

i.e., j weights each voter according to the inverse of the voting population.

In this case, j's critical *choices,* (z_j^C, a_j^C) in (9) using (14), are given by

$$\begin{pmatrix} z_j^C \\ a_j^C \end{pmatrix} = \sum\nolimits_{i \in N} \mu^O \begin{pmatrix} x_i \\ t_i \end{pmatrix} = \frac{1}{n} \sum\nolimits_{i \in N} \begin{pmatrix} x_i \\ t_i \end{pmatrix} \equiv \begin{pmatrix} z_m \\ a_m \end{pmatrix} \quad (15)$$

where the last term follows from (10). This says that if *all other* candidates adopt (z^O, a^O) as their campaign strategy, then j's best response is to adopt the electoral mean $(\mathbf{z}_m, \mathbf{a}_m)$ as its campaign strategy. As this is true for all candidates, the Lemma is proven. ■

Using (13), define V_j^m as *j's expected vote share*[17] *at the electoral mean*, $(\mathbf{z}_m, \mathbf{a}_m)$, i.e.,

$$V_j^m \equiv V_j^m(\mathbf{z}_m, \mathbf{a}_m) = \rho_j(\mathbf{z}_m, \mathbf{a}_m) \equiv \rho_j^m \quad (16)$$

where ρ_j^m is ρ_j^O in (12) evaluated at $(\mathbf{z}_m, \mathbf{a}_m)$. $V_j^m = \rho_j^m$ is exogenously given since from (12) ρ_j^m depends only on the exogenously given mean sociodemographic, traits and competence valences. Note that when candidates adopt the electoral mean as their campaign strategy, the campaign profile $(\mathbf{z}_m, \mathbf{a}_m)$ satisfies the FONC for *all* candidates.

Moreover, we can rank candidates' expected vote shares at the electoral mean. From (12) it is clear that to rank V_j^m for all $j \in \mathcal{C}$, the mean sociodemographic, traits and competence valences,

[17] We identify functions evaluated at $(\mathbf{z}_m, \mathbf{a}_m)$ with a subscript or superscript m, e.g., the weight given to votes at $(\mathbf{z}_m, \mathbf{a}_m)$ is given by $\mu^m = \frac{1}{n}$.

$(\alpha_j, \tau_j, \lambda_j)$ in (2), of all candidates need to be ranked. To do so, define the *expected composite valence of candidate* j, κ_j, as

$$\kappa_j = \alpha_j + \tau_j + \lambda_j \quad \text{for all } j \in \mathcal{C}. \tag{17}$$

Let candidate 1 (candidate c) be the one with the *lowest* (*highest*) expected composite valence so that candidates' expected composite valences satisfy

$$\kappa_1 \leq ... \leq \kappa_j \leq ... \leq \kappa_c. \tag{18}$$

Using κ_j in (17), the probability that voter i chooses candidate j in (12) when candidates adopt the electoral mean as their campaign strategy is given by

$$\rho_j^m = \left[\sum\nolimits_{h\neq j, h=1}^{c} \exp[(\alpha_h + \tau_h + \lambda_h) - (\alpha_j + \tau_j + \lambda_j)]\right]^{-1} = \left[\sum\nolimits_{h\neq j, h=1}^{c} \exp[\kappa_h - \kappa_j]\right]^{-1}. \tag{19}$$

Given the composite valence ranking in (18), candidates' expected vote shares at the electoral mean V_j^m in (16) satisfy

$$V_1^m = \rho_1^m \leq ... \leq V_j^m = \rho_j^m \leq ... \leq V_c^m = \rho_c^m. \tag{20}$$

At the electoral mean, (z_m, a_m), candidate 1 with the lowest composite valence, κ_1, is also the candidate with the *lowest* expected vote share.

Let us now determine the conditions under which candidates adopt the electoral mean as their campaign strategy.

7 Convergence conditions

We now derive the conditions under which j *adopts the electoral mean as its* critical campaign, $(z_j^C, a_j^C) = (z_m, a_m)$, holding the electoral campaign of all other candidates at the electoral mean.

The Appendix shows that two pivotal vote shares characterize the *necessary* and *sufficient* conditions under which the electoral mean is a local Nash equilibrium (LNE) of the election, i.e., the conditions under which all candidates adopt (z_m, a_m) as their electoral campaign strategy.

Definition 4 *The necessary (n) and sufficient (s) pivotal (p) vote shares at* (z_m, a_m) *are given by*

$$\mathcal{V}_{sp}^m \equiv \mathcal{V}_{sp}(\mathbf{z}_m, \mathbf{a}_m) \equiv \frac{1}{2} - \frac{1}{4}\frac{1}{Tr\left[(\mathbf{B})^{-1}\boldsymbol{\nabla}^m\right]} = \frac{1}{2} - \frac{1}{4}\frac{1}{\frac{1}{\beta}var(x) + \frac{1}{e}var(t)} \tag{21}$$

$$\mathcal{V}_{np}^m \equiv \mathcal{V}_{np}(\mathbf{z}_m, \mathbf{a}_m) \equiv \frac{1}{2} - \frac{1}{4}\frac{Tr(\mathbf{B})}{Tr\left(\boldsymbol{\nabla}^m\right)} = \frac{1}{2} - \frac{1}{4}\frac{\beta + e}{var(x) + var(t)} \tag{22}$$

where the importance voters give to each dimension are represented by $\mathbf{B}$ *defined as*

$$\mathbf{B} \equiv \begin{pmatrix} \beta & 0 \\ 0 & e \end{pmatrix} \quad \Rightarrow \quad Tr(\mathbf{B}) = \beta + e \tag{23}$$

so that its trace, $Tr(\mathbf{B})$, measures the aggregate importance voters give to the policy and advertising dimensions.

The matrix $\boldsymbol{\nabla}^m$ in (21) and (22) given by

$$\boldsymbol{\nabla}^m \equiv \frac{1}{n}\sum\nolimits_{i\in\mathcal{N}} \begin{pmatrix} \beta(x_i - z_m)^2\beta & \beta(x_i - z_m)(t_i - a_m)e_k^S \\ \beta(x_i - z_m)(t_i - a_m)e & e(t_i - a_m)^2 e \end{pmatrix} \tag{24}$$

measures the weighted variance-covariance matrix of the distribution of voters preferences around the electoral mean (z_m, a_m).

Let $var(x)$ represent the variance of voters' ideal policies, i.e.,

$$var(x) \equiv \frac{1}{n}\sum\nolimits_{i\in\mathcal{N}} \beta(x_i - z_m)^2\beta. \tag{25}$$

Similarly, let $var(t)$ denote the variance of voters' campaign tolerance levels given by

$$var(t) \equiv \frac{1}{n}\sum\nolimits_{i\in\mathcal{N}} e(t_i - a_m)^2 e. \tag{26}$$

Using (24), (25) and (26), the trace of $\boldsymbol{\nabla}^m$, $Tr\left(\boldsymbol{\nabla}^m\right)$, given by

$$\begin{aligned} Tr\left(\boldsymbol{\nabla}^m\right) &= \frac{1}{n}\sum\nolimits_{i\in\mathcal{N}} \beta(x_i - z_m)^2\beta + \frac{1}{n}\sum\nolimits_{i\in\mathcal{N}} e(t_i - a_m)^2 e \\ &= var(x) + var(t), \end{aligned} \tag{27}$$

measures the "aggregate" variance of voters' distribution around the electoral mean and is always positive. The trace of $(\mathbf{B})^{-1}\boldsymbol{\nabla}^m$, $Tr\left[(\mathbf{B})^{-1}\boldsymbol{\nabla}^m\right]$ in (21), using (25) and (26),

$$\begin{aligned} Tr\left[(\mathbf{B})^{-1}\boldsymbol{\nabla}^m\right] &= \frac{1}{\beta}\frac{1}{n}\sum\nolimits_{i\in\mathcal{N}} \beta(x_i - z_m)^2\beta + \frac{1}{e}\frac{1}{n}\sum\nolimits_{i\in\mathcal{N}} e(t_i - a_m)^2 e \\ &= \frac{1}{\beta}var(x) + \frac{1}{e}var(t), \end{aligned} \tag{28}$$

scales the weighted variance in each dimension by the importance voters give respectively to the policy, β, and advertising, e, dimensions and is always positive.

Lemma 2 *It is always the case that the sufficient pivotal vote share is higher than the necessary one, i.e., $\mathcal{V}_{sp}^m > \mathcal{V}_{np}^m$ for all $j \in \mathcal{C}$.*

Proof: *See Appendix A Section 10.2.*

Lemma 2 proves that the sufficient condition for convergence to the electoral mean is more stringent than the necessary one, i.e., $\mathcal{V}_{sp}^m > \mathcal{V}_{np}^m$ for all $j \in \mathcal{C}$.

The following proposition gives the sufficient and necessary conditions for the joint electoral mean, $(\mathbf{z}_m, \mathbf{a}_m)$, to be an equilibrium of the election.

Proposition 1 *Convergence to the electoral mean*

Assume that candidates adopt the electoral mean (z_m, a_m) as their campaign strategy. Let voters' idiosyncratic sociodemographic, traits and competency components of the valences, $\xi_{ij} \sim E(0, \frac{\pi}{6})$, $\varsigma_{ij} \sim T(0, \frac{\pi}{6})$ and $\epsilon_{ij} \sim F(0, \frac{\pi}{6})$ and their sum $\varkappa_{ij} = \xi_{ij} + \varsigma_{ij} + \epsilon_{ij} \sim \mathcal{X}(0, \frac{\pi}{6})$ follow Type-I extreme value distributions.

- *If $\mathcal{V}_{sp}^m \leqslant V_j^m$ for all $j \in \mathcal{C}$, the* ***sufficient*** *condition for convergence to (z_m, a_m) has been met by all candidates. The joint electoral mean, $(\mathbf{z}_m, \mathbf{a}_m)$, constitutes a* ***strict*** *LNE of the election.*

- *If $\mathcal{V}_{np}^m \leqslant V_j^m < \mathcal{V}_{sp}^m$ for some $j \in \mathcal{C}$ and $\mathcal{V}_{sp}^m \leqslant V_h^m$ for $h \neq j \in \mathcal{C}$, the* ***necessary*** *but not the sufficient condition for convergence to (z_m, a_m) is met by candidate j. The joint electoral mean, $(\mathbf{z}_m, \mathbf{a}_m)$, is a* ***weak*** *LNE of the election.*

- *If $V_j^m < \mathcal{V}_{np}^m$ for some $j \in \mathcal{C}$, the* ***necessary*** *condition for convergence to (z_m, a_m) is* ***not*** *met for candidate j. The joint electoral mean, $(\mathbf{z}_m, \mathbf{a}_m)$, is* ***not*** *a LNE of the election and at least one candidate adopts a campaign that differs from the electoral mean.*

Proof: *See Appendix, Section 10.1.*

The Appendix shows that when all candidates' electoral campaigns, except j's, are held constant, the *sufficient* condition for j to adopt the electoral mean, (z_m, a_m), as its electoral campaign, also given in the Appendix in (53), is

$$\mathcal{V}_{sp}^m \leqslant V_j^m \tag{29}$$

where $\mathcal{V}_{sp}^m$ is given by (21) and V_j^m by (16). Condition (29) says that j's expected vote share at $(\mathbf{z}_m, \mathbf{a}_m)$, V_j^m, must be *high* enough for j to adopt (z_m, a_m) as its electoral campaign.

Note that the Left hand side (LHS) of (29), also given in (21), is independent of j's policy and ad campaign and is the same for all candidates. Thus, (29) gives the *sufficient condition* for all candidates to converge to the electoral mean. When (29) is satisfied for all $j \in \mathcal{C}$, the joint electoral mean $(\mathbf{z}_m, \mathbf{a}_m)$ is a *strict* LNE of the election.

The *necessary* convergence condition for candidate j to adopt (z_m, a_m) as its electoral campaign–also given in the Appendix in (48)–is that

$$\mathcal{V}_{np}^m \leqslant V_j^m \tag{30}$$

where $\mathcal{V}_{np}^m$ is given by (22) and V_j^m by (16). Condition (30) says that the necessary condition for j to adopt (z_m, a_m) as its electoral campaign is that its expected vote share at $(\mathbf{z}_m, \mathbf{a}_m)$, V_j^m, is *high* enough, i.e., is greater than $\mathcal{V}_{np}^m$. When (30) is satisfied for some $j \in \mathcal{C}$ while $\mathcal{V}_{sp}^m \leqslant V_h^m$ for $h \neq j \in \mathcal{C}$, the joint electoral mean $(\mathbf{z}_m, \mathbf{a}_m)$ is a *weak* LNE of the election.

If $V_j^m < \mathcal{V}_{np}^m$, j's expected vote share does *not* meet the necessary convergence condition in (30), j moves away from (z_m, a_m) to increase its votes since j's expected vote share is at a minimum or at a saddle-point at $(\mathbf{z}_m, \mathbf{a}_m)$. Once j moves away from (z_m, a_m), other candidates may also find it in their interest move away from the electoral mean. The joint electoral mean, $(\mathbf{z}_m, \mathbf{a}_m)$, is *not* a LNE of the election.

8 Comparative statics

We are interested in finding how the sufficient and necessary pivotal vote shares–given by (21) and (22)–change when the parameters of the model change. The comparative statics carried out are such that candidates continue to adopt the electoral mean as their campaign strategy. Changes in the importance voters give to each dimension change the pivotal vote shares and imply changes in the conditions leading to the LNE of the election. We examine changes in the importance given to the policy (ad campaign) dimension then changes in the valences.

Corollary 1 *Comparative Statics*

Suppose that the electoral mean is a LNE of the election.

An increase in the importance voters give to the policy and ad campaign dimensions, β and e, ceteris paribus,

- *increases the sufficient pivotal vote share $\mathcal{V}_{sp}^m$ thus making it more difficult for the electoral mean to be a strick LNE of the election*
- *has an ambiguous effect on the necessary pivotal vote share $\mathcal{V}_{np}^m$. The necessary pivotal vote share increases (decreases) if $\frac{\beta+e}{\beta} var(x) > (<) \frac{1}{2} [var(x) + var(t)]$.*

A mean preserving increase in $var(x)$ or $var(t)$, ceteris paribus,

- *increases the sufficient and necessary pivotal vote shares.*

When voters give greater importance to the policy (ad campaign) dimension relative to the other dimension and to the valences, i.e., when β increases (e increases), holding every thing else constant, $Tr(\mathbf{B})$ increases. In this case, the sufficient pivotal increases (see Appendix) increases. As a consequence, the sufficient condition for convergence to the electoral mean becomes more difficult to satisfy as candidates' expected vote shares must be greater than the now higher sufficient pivotal vote share. Thus, the greater the importance voters give one of the two dimensions the more difficult it becomes for candidates to converge to the electoral mean.

The effect of an increase in β or in e, holding every thing else constant, may increase or decrease the necessary pivotal vote. The effect depends on whether

$$\begin{aligned} \frac{\beta + e}{\beta} var(x) &\gtrless \frac{1}{2} [var(x) + var(t)] \\ \frac{\beta + e}{e} var(t) &\gtrless \frac{1}{2} [var(x) + var(t)] . \end{aligned}$$

This says that the $\mathcal{V}_{np}^m$ rises as β or e increase when the variance of voters' ideal policies weighted by $\frac{\beta+e}{\beta}$ is larger than half the aggregate variance of voters' preferences, $Tr\left(\boldsymbol{\nabla}^m\right)$, i.e., greater than half the aggregate variance of voters' preferences, $var(x) + var(t)$. This says that the more dispersed voters are along either dimension the more difficult it is for candidates' to adopt the electoral mean as their electoral campaign strategy.

These results show that when β or e increase, the sufficient condition for convergence to the electoral mean may be harder to meet. In this case, it becomes harder for the electoral mean to be a *strick* LNE of the election. When the necessary pivotal vote share deceases as β or e increase, it is more likely that the necessary condition for convergence to the electoral mean is satisfied for a broader range of expected vote shares for all candidates and thus the electoral mean is a weak LNE of the election for a wider range of expected vote share.

When voters' ideal policies or campaign tolerance levels become more dispersed along the corresponding dimension, a higher $var(x)$ or $var(t)$, the pivotal vote shares $\mathcal{V}_{sp}^m$ and $\mathcal{V}_{np}^m$ rise, thus making it more difficult for candidates to converge to the electoral mean.

Suppose that there is an increase in the mean sociodemographic, traits or competence valence of candidate j. The mean valence components do not affect the pivotal vote shares or the electoral mean, however, these mean components increase j's expected vote share while decreasing the expected vote shares of all other candidates and so affect the conditions under which there is convergence to the electoral mean.

9 Concluding comments

We introduce campaign advertising into Schofield's (2007) stochastic election model to examine how candidates choose their advertising campaign strategy by endowing voters with preferences over policies and ad campaigns. Voters evaluate candidates not only by how far the candidate's policy is from their ideal but also by the frequency of candidates' messages relative to their campaign tolerance level. Their choice of candidate is also affected by their sociodemographic, traits and competence valences.

Lemma 1 shows that if candidates adopt the same campaign they adopt the electoral mean as their campaign strategy. Proposition 1 extends Schofield's (2007) result to include candidates' ad campaigns and voters' campaign tolerance levels. In this extension, we adapt Schofield's results to one where candidates contact voters directly in an effort to give voters' a further impetus to vote for them. The sufficient and necessary conditions for convergence to the electoral mean are presented in terms of pivotal vote shares rather than the convergence coefficient used by Schofield (2007). The pivotal vote shares give intuitive results: candidates' expected vote shares must be high enough for candidates to want to rationally adopt the electoral mean, $(\mathbf{z}_m, \mathbf{a}_m)$, as their campaign strategy. The sufficient condition for convergence to the electoral mean requires that all candidates' expected vote shares at the electoral mean be higher than the sufficient pivotal vote share $\mathcal{V}_{sp}^m$; the necessary condition that they be higher the necessary pivotal vote share $\mathcal{V}_{np}^m$. If the expected vote share of at least one candidate is below $\mathcal{V}_{np}^{m}$, then that candidate moves away from the electoral mean to increase its votes and so the electoral mean cannot be a LNE of the election.

The result on advertising campaign is new as it shows that the variation in voters' campaign tolerances levels not only determines candidates' advertising campaigns but also affects the electoral outcome. The presence of voters' campaign tolerance levels shows that if a candidate's advertising campaign alienates too many voters by either contacting them too often or not often enough, the utility these voters get from this candidate decreases, thus decreasing the probability that they vote for the candidate. Candidates have then an incentive to find their optimal advertising campaign, one that alienates the fewest number of voters taking into account the effect that their choice of policy has on the expected vote share. In our model candidates use two instruments to maximize their vote shares, policies and ad campaigns. Thus, inour model adjustments in policy may be smaller than if the ad campaign were not present.

The comparative statics results shows that where voters to give greater importance to one of the two dimensions (an increase in β or e holding every thing else constant) the more difficult it is for the expected vote shares to meet the sufficient condition for converge to the electoral mean. If the weighted variance along one dimension is more than half the aggregate variance of voters' preferences around the electoral mean, an increase in β or e makes it easier for candidates to meet the necessary condition for convergence to the electoral mean.

An increase in the mean sociodemographic, traits and/or competence valences of candidate j increases j's expected vote share, decreases the vote share of all the other candidates but does not affect the pivotal vote shares or the electoral mean, thus affecting the convergene conditions to the electoral mean.

Predictions emanating from these results are that if candidates and their campaign strategists realize that either candidates' positions or ad campaigns are ineffective, in the sense of convincing voters to vote for them, then adjustments to positions and/or ad campaigns are made in an effort to increase the probability that these voters vote for the candidate. This result is important because it shows that when candidates have two campaign instruments–policy and advertising campaigns–at their disposal they may not necessarily adjust their policy position. That is, they may instead adjust their advertising campaign to increase their vote share.

References

[1] ANSOLABEHERE S, de FIGUEIREDO JM, SNYDER J. 2003. Why Is There so Little Money in U.S. Politics? *The Journal of Economic Perspectives* 17.1 : 105–130.

[2] ARAGONES E, PALFREY T. 2002. Mixed Equilibrium in a Downsian Model with a Favored Candidate. *Journal of Economic Theory*, 103: 131–161.

[3] ASHWORTH S, BUENO DE MESQUITA E. 2009. Elections with Platform and Valence Competition. *Games Econonomic Behavior* 67: 191–216.

[4] AUSTEN-SMITH D, BANKS JS. 1999. Cycling of simple rules in the spatial model. *Social Choice and Welfare*, 16(4): 663-672.

[5] CONDORCET J. 1984 [1785]. Essai sur l'application de l'analyse a la probabilite des decisions rendus a la pluralite des voix. Imprimerie Royale, Paris.

[6] CLARKE, H D, KORNBERG A, SCOTTO, TJ. 2009. *Making Political Choices*. Toronto: Toronto University Press.

[7] DOWNS, A. 1957. *An Economic Theory of Democracy*. New York: Harper and Row

[8] GALLEGO, M, SCHOFIELD, N. 2014. The Convergence Coefficient across political systems. *Probability and Statistics*.

[9] GALLEGO, M, SCHOFIELD. 2015 Do all parties converge to the center in political systems. *The Journal of Theoretical Politics* (in press).

[10] GROSECLOSE T. A Model of Candidate Location When One Candidate Has a Valence Advantage. *American Journal of Political Science* 45.4 (2001): 862–886.

[11] HENDRICKS Ja, KAID LL. 2011. *Techno politics in presidential campaigning : new voices, new technologies, and new voters*. New York : Routledge.

[12] MADISON J. 1999[1787]. The Federalist No. 10., In J. Rakove (Ed.), James Madison: Writings. New York: The Library of America.

[13] MCKELVEY RD. 1976. Intransitivities in multidimensional voting models and some implications for agenda contro. *Journal of Economic Theory*, 12(3): 472-482.

[14] MCKELVEY RD, PATTY J. 2006. A Theory of Voting in Large Elections. *Games Econonic Behavior* 57: 155–180.

[15] MEIROWITZ A. 2008. Electoral Contests, Incumbency Advantages and Campaign Finance. *Journal of Politics* 70: 681-699.

[16] Riker WH. 1980. Implications From the Disequilibrium of Majority Rule For the Study of Institutions. *Amererican Political Science Review* 74: 432–46.

[17] Riker WH 1982. *Liberalism against Populism*. San Francisco, Freeman.

[18] Riker WH 1986. *The Art of Political Manipulation*. New Haven, Yale University Press.

[19] SAARI D. 1997 The Generic existence of a core for q-Rules. *Economic Theory* 9:219–260.

[20] SANDERS D, CLARKE H, STEWART MC, WHITELEY P. 2011. Downs, Stokes and the Dynamics of Electoral Choice. *British Journal of Political Science* 41(2): 287-314.

[21] SCHOFIELD N. 1978. Instability of simple dynamic games. *Review of Economic Studies* 45: 575–594.

[22] SCHOFIELD N. 2002. Evolution of the Constitution. *British Journal of Political Science* 32:1–20.

[23] SCHOFIELD N. 2007. The mean voter theorem: Necessary and sufficient conditions for convergent equilibrium. *Review of Economic Studies* 74, 965-980.

[24] SCHOFIELD, N, N, GALLEGO M. 2011. *Leadership or Chaos*, Springer: Berlin

[25] SCHOFIELD N, GALLEGO M, JEON JS. 2011a. Leaders, Voters and Activists in Elections in the Great Britain 2005 and 2010. *Electoral Studies* 30(3) :484-496.

[26] SCHOFIELD N„ JEON JS, MUSKHELISHVILI, M. OZDEMIR, U. TAVITS, . 2011b. Modeling Elections in Post-Communist Regimes: Voter Perceptions, Political Leaders and Activists, in: Schofield, N., Caballero G. *(Eds.) The Political Economy of Institutions, Democracy and Voting*, Springer, Berlin, pp.259-301.

[27] STOKES D 1963. Spatial models and party competition. *AmericanPolitical Science Review* 57:368–377.

[28] STOKES D. 1992. Valence Politics. In D. Kavanagh (Ed.), *Electoral Politics*. Oxford: Clarendon Press.

[29] TRAIN K. 2003. *Discrete Choice Methods for Simulation*. Cambridge, U.K., Cambridge University Press.

[30] ZAKHAROV AV. 2009. A Model of Candidate Location with Endogenous Valence. *Public Choice* 138: 347-366.

10 Appendix

10.1 Second order conditions (SOC)

Section 5 derived candidates' best response policies and ad campaign functions and the critical values that satisfy the first order conditions. Lemma 1 proved that if candidates adopt the same electoral campaign, they adopt the electoral mean, $(\mathbf{z}_m, \mathbf{a}_m)$, as their campaign strategy. To evaluate if candidates' vote share functions are at a maximum at $(\mathbf{z}_m, \mathbf{a}_m)$ we use the *Hessian matrix of second order partial derivatives* of the vote share functions of each candidate evaluated at $(\mathbf{z}_m, \mathbf{a}_m)$.

Before proceeding recall from Section 6 that when all candidates adopt the electoral mean as their campaign strategy, the probability that i votes for j, ρ_{ij} in (4) when evaluated at $(\mathbf{z}_m, \mathbf{a}_m)$, is independent of voters' ideal policies and campaign tolerance levels as well as independent of candidates' policies and ad campaigns and gives j's expected vote share so that from (16),

$$\rho_{ij}^m = \rho_j^m = V_j^m \text{ for all } i \in \mathcal{N}. \tag{31}$$

10.1.1 Proof of Proposition 1

We use the *Hessian matrix of second order partial derivatives of* candidate j to determine whether j's vote share function, V_j in (5), is at a maximum at (z_m, a_m) holding the electoral campaigns of all other candidates at (z_m, a_m). To find this Hessian we need the second order partial derivatives of the probability that i votes for j, $\mathbf{D}^2\rho_{ij}$ evaluated at $(\mathbf{z}_m, \mathbf{a}_m)$, $\mathbf{D}^2\rho_{ij}^m$, using ρ_{ij} in (4). Then using $\mathbf{D}^2\rho_{ij}^m$ we find the Hessian of j's vote share, $\mathbf{D}^2 V_j$ at $(\mathbf{z}_m, \mathbf{a}_m)$.

Suppose all candidates adopt the electoral mean, (z_m, a_m) as their campaign strategy. Using (31), the partial derivatives of (7) wrt z_j and a_j evaluated at (z_m, a_m), i.e., $\mathbf{D}^2\rho_{ij}^m$, is given by

$$\begin{aligned} \mathbf{D}^2\rho_{ij}^m &\equiv \mathbf{D}^2\rho_{ij}(\mathbf{z}_m, \mathbf{a}_m) \equiv \begin{pmatrix} \frac{\partial^2\rho_{ij}}{\partial(z_j)^2} & \frac{\partial^2\rho_{ij}}{\partial z_j \partial a_j} \\ \frac{\partial^2\rho_{ij}}{\partial z_j \partial a_j} & \frac{\partial^2\rho_{ij}}{\partial(a_j)^2} \end{pmatrix} \\ &= 2\rho_j^m(1-\rho_j^m)\left[2(1-2\rho_j^m)\mathbf{B}\mathbf{E}^{ij\mathbf{T}}_{(\mathbf{z}_m,\mathbf{a}_m)}\mathbf{E}^{ij}_{(\mathbf{z}_m,\mathbf{a}_m)}\mathbf{B} - \mathbf{B}\right] \\ &= 2\rho_j^m(1-\rho_j^m)\left[2(1-2\rho_j^m)\boldsymbol{\nabla}_i^m - \mathbf{B}\right] \end{aligned} \tag{32}$$

$$\text{where} \qquad \mathbf{B} \equiv \begin{pmatrix} \beta & 0 \\ 0 & e \end{pmatrix}, \quad \mathbf{E}^{ij}_{(\mathbf{z}_m,\mathbf{a}_m)} \equiv \left[\begin{array}{cc} (x_i - z_m) & (t_i - a_m) \end{array}\right]. \tag{33}$$

and where a vector with a $\mathbf{T}$ denotes the transpose of that vector. The diagonal of $\mathbf{B}$ shows the importance voters give the policy and advertising dimensions and $\mathbf{E}^{ij}_{(\mathbf{z}_m,\mathbf{a}_m)}$ the differences between i's ideals (x_i, t_i) and j's electoral campaign when at the electoral mean (z_m, a_m). The matrix $\boldsymbol{\nabla}_i^m$ in (32) defined by

$$\boldsymbol{\nabla}_i^m \equiv \mathbf{B}\mathbf{E}^{ij\mathbf{T}}_{(\mathbf{z}_m,\mathbf{a}_m)}\mathbf{E}^{ij}_{(\mathbf{z}_m,\mathbf{a}_m)}\mathbf{B} \equiv \begin{pmatrix} \beta(x_i - z_m)^2\beta & \beta(x_i - z_m)(t_i - a_m)e \\ \beta(x_i - z_m)(t_i - a_m)e & e(t_i - a_m)^2 e \end{pmatrix} \tag{34}$$

gives the *weighted variance-covariance matrix* of *i's ideals around the electoral mean* (z_m, a_m).

Using $\mathbf{D}^2\rho_{ij}^m$ in (32), the *Hessian* of second order partial derivatives of V_j in (5) wrt z_j and a_j, evaluated at $(\mathbf{z}_m, \mathbf{a}_m)$ is given by

$$\begin{aligned} \mathbf{D}^2 V_j^m &\equiv \mathbf{D}^2 V_j(\mathbf{z}_m, \mathbf{a}_m) \equiv \frac{1}{n}\sum\nolimits_{i\in\mathcal{N}} \mathbf{D}^2\rho_{ij}^m \\ &= \frac{1}{n}\sum\nolimits_{i\in\mathcal{N}} 2\rho_j^m(1-\rho_j^m)\left[2(1-2\rho_j^m)\boldsymbol{\nabla}_i^m - \mathbf{B}\right] \\ &= 2\rho_j^m(1-\rho_j^m)\left\{2(1-2\rho_j^m)\left(\frac{1}{n}\sum\nolimits_{i\in\mathcal{N}} \boldsymbol{\nabla}_i^m\right) - \mathbf{B}\right\}. \end{aligned} \tag{35}$$

Note that averaging $\boldsymbol{\nabla}_i^m$ in (34) across all voters, we obtain the matrix

$$\boldsymbol{\nabla}^m \equiv \frac{1}{n}\sum\nolimits_{i\in\mathcal{N}} \boldsymbol{\nabla}_i^m, \tag{36}$$

i.e., *the weighted variance-covariance matrix* of the distribution of voters' ideal policies and campaign tolerance levels around the *electoral mean* where the weights are given by $\mathbf{B}$ in (33).

After substituting $\boldsymbol{\nabla}^m$ from (36) into (35), we get that j's hessian matrix is given by

$$\mathbf{D}^2 V_j^m = 2\rho_j^m(1-\rho_j^m)\left\{2(1-2\rho_j^m)\boldsymbol{\nabla}^m - \mathbf{B}\right\} \tag{37}$$

Candidate j adopts the electoral mean, (z_m, a_m), as its campaign strategy, iff j's vote share is at a maximum at (z_m, a_m). Hence, to show that (z_m, a_m) for all $j \in \mathcal{C}$ is a *strict (weak) local Nash equilibrium* [S(W)LNE] of the election we must find the conditions under which j's Hessian matrix, $\mathbf{D}^2 V_j$ in (35) when evaluated at (z_m, a_m), $\mathbf{D}^2 V_j^m$, is negative definite (semi-definite) for all $j \in \mathcal{C}$ which happens only when its eigenvalues are negative (non-positive) for all $j \in \mathcal{C}$, that is, iff the trace and the determinant of $\mathbf{D}^2 V_j^m$ are respectively negative and positive as in this case the eigenvalues of $\mathbf{D}^2 V_j^m$ are both negative.

Before deriving the trace and the determinant of $\mathbf{D}^2 V_j^m$ and to simplify the calculations, we give the following definitions.

Definition 5 *Let $\mathbf{C}_j^m \equiv \mathbf{C}_j(\mathbf{z}_m, \mathbf{a}_m)$ be the characteristic matrix of candidate j when all candidates adopt the electoral mean as their campaign strategy $(\mathbf{z}_m, \mathbf{a}_m)$, i.e.,*

$$\mathbf{C}_j^m \equiv \mathbf{C}_j(\mathbf{z}_m, \mathbf{a}_m) \equiv A_j^m \boldsymbol{\nabla}^m - \mathbf{B} \tag{38}$$

$$\textit{where} \quad A_j^m \equiv A_j^m(\mathbf{z}_m, \mathbf{a}_m) \equiv 2(1-2\rho_j^m) \tag{39}$$

is the characteristic factor of $\mathbf{C}_j^m$, $\boldsymbol{\nabla}^m$ is given by (36), ρ_j^m by (31) and $\mathbf{B}$ by (33).

Let q_j^m be the joint probability that voters vote for j and for any other candidate at $(\mathbf{z}_m, \mathbf{a}_m)$,

$$q_j^m \equiv q_j(\mathbf{z}_m, \mathbf{a}_m) \equiv \rho_j^m(1-\rho_j^m). \tag{40}$$

Using (39) and (40), $\mathbf{D}^2 V_j^m$ in (37) can be expressed as a function of $\mathbf{C}_j^m$ in (38), i.e.,

$$\mathbf{D}^2 V_j^m = 2q_j^m\left\{A_j^m \boldsymbol{\nabla}^m - \mathbf{B}\right\} \equiv 2q_j^m \mathbf{C}_j^m. \tag{41}$$

So that the trace of $\mathbf{D}^2 V_j^m$ in (38) is then given by

$$Tr\left(\mathbf{D}^2 V_j^m\right) = 2q_j^m Tr(\mathbf{C}_j^m). \tag{42}$$

Since q_j^m in (40) is always positive, $Tr\left(\mathbf{D}^2 V_j^m\right) < 0$ iff $Tr(\mathbf{C}_j^m) < 0$.

The trace of $\mathbf{C}_j^m$ in (38) is given by

$$Tr(\mathbf{C}_j^m) \equiv Tr\left[A_j^m \boldsymbol{\nabla}^m - \mathbf{B}\right] = A_j^m Tr\left(\boldsymbol{\nabla}^m\right) - Tr(\mathbf{B}) \tag{43}$$

where $Tr(\boldsymbol{\nabla}^m)$ is given by (27), $Tr(\mathbf{B})$ by (23) and A_j^m by (39). From (43), $Tr(\mathbf{C}_j^m) < 0$ when

$$A_j^m < \frac{Tr(\mathbf{B})}{Tr\left(\boldsymbol{\nabla}^m\right)}. \tag{44}$$

After substituting A_j^m from (39) into (44) and some manipulation, we obtain

$$\frac{1}{2} - \frac{1}{4}\frac{Tr(\mathbf{B})}{Tr\left(\boldsymbol{\nabla}^m\right)} < \rho_j^m \tag{45}$$

The left hand side (LHS) of (45) is independent of candidates' policies and ad campaigns and is thus the same for all candidates. Define the LHS of (45) as the *necessary pivotal (np) vote share* at the electoral mean, (z_m, a_m), i.e.,

$$\mathcal{V}_{np}^m \equiv \frac{1}{2} - \frac{1}{4}\frac{Tr(\mathbf{B})}{Tr\left(\boldsymbol{\nabla}^m\right)} = \frac{1}{2} - \frac{1}{4}\frac{\beta + e}{\frac{1}{n}\sum_{i\in\mathcal{N}}\beta(x_i - z_m)^2\beta + \frac{1}{n}\sum_{i\in\mathcal{N}} e(t_i - a_m)^2 e} \tag{46}$$

where the numerator in the last term follows from (23) and the denominator from (27).

Since ρ_j^m in (31) is independent of candidates' policies and ad campaigns and is the same for all voters then ρ_j^m also gives j's expected vote share at the electoral mean, i.e., $V_j^m = \rho_j^m$. Therefore, condition (45) can be re-written as

$$\mathcal{V}_{np}^m \equiv \frac{1}{2} - \frac{1}{4}\frac{Tr(\mathbf{B})}{Tr\left(\boldsymbol{\nabla}^m\right)} < \rho_j^m = V_j^m \tag{47}$$

This condition says that if j's expected vote share at $(\mathbf{z}_m, \mathbf{a}_m)$, V_j^m, is *higher* than the necessary pivotal vote share, $\mathcal{V}_{np}^m$, then $Tr(\mathbf{C}_j^m) < 0$. If for *some* candidate $j \in \mathcal{C}$, $\mathcal{V}_{np}^m > \rho_j^m = V_j^m$, then for this candidate $Tr(\mathbf{C}_j^m) > 0$.

Therefore, the *necessary condition* for j to adopt (z_m, a_m) as its electoral campaign, when all other candidates adopt the electoral mean as their electoral campaign, is

$$\mathcal{V}_{np}^m < \rho_j^m = V_j^m. \tag{48}$$

That is, candidate j must expect a *high* enough vote share at $(\mathbf{z}_m, \mathbf{a}_m)$ for j to adopt the electoral mean as its campaign strategy.

The *sufficient SOC* for j to converge to (z_m, a_m) is that the eigenvalues of $\mathbf{D}^2 V_j^m$ in (37) be *both* negative implying that the *determinant* of $\mathbf{D}^2 V_j^m$ must be positive at $(\mathbf{z}_m, \mathbf{a}_m)$, i.e., $\det\left(\mathbf{D}^2 V_j^m\right) > 0$. From (37), the determinant of $\mathbf{D}^2 V_j$ when evaluated at $(\mathbf{z}_m, \mathbf{a}_m)$ is given by

$$\det\left(\mathbf{D}^2 V_j^m\right) = \det\left[2q_j^m\left\{A_j^m \boldsymbol{\nabla}^m - \mathbf{B}\right\}\right] = \det\left[2q_j^m A_j^m \boldsymbol{\nabla}^m - 2q_j^m \mathbf{B}\right]$$

After some manipulation and using (23) and (27), $\det(\mathbf{D}^2 V_j^m)$ becomes

$$\begin{aligned}\det(\mathbf{D}^2 V_j^m) &= 4\left[q_j^m A_j^m\right]^2 \det(\boldsymbol{\nabla}^m) \\ &\quad +4\left[q_j^m\right]^2 \det(\mathbf{B}) \times \left\{1 - A_j^m Tr\left((\mathbf{B})^{-1}\boldsymbol{\nabla}^m\right)\right\}\end{aligned} \tag{49}$$

From the triangle inequality, the determinant of the variance-covariance matrix of voters' policies and campaign tolerance levels around the electoral mean, $\det(\boldsymbol{\nabla}^m)$ is always non-negative. The determinant of $\mathbf{D}^2 V_j^m$ is positive iff the last term in (49) is positive. Since $\det(\mathbf{B}) = \beta e > 0$ and since from (40) $q_j^m > 0$, the determinant of $\mathbf{D}^2 V_j^m$ is positive iff the term in curly brackets is positive, i.e., iff

$$A_j^m < \frac{1}{Tr\left[(\mathbf{B})^{-1}\boldsymbol{\nabla}^m\right]}. \tag{50}$$

After substituting A_j^m from (39) into (50) and some manipulation, we obtain

$$\frac{1}{2} - \frac{1}{4}\frac{1}{Tr\left[(\mathbf{B})^{-1}\boldsymbol{\nabla}^m\right]} < \rho_j^m. \tag{51}$$

Note that the LHS of (51) is the same for all candidates. Define the LHS of (51) as the *sufficient pivotal (sp) vote share* at $(\mathbf{z}_m, \mathbf{a}_m)$, i.e.,

$$\mathcal{V}_{sp}^m \equiv \frac{1}{2} - \frac{1}{4}\frac{1}{Tr\left[(\mathbf{B})^{-1}\boldsymbol{\nabla}^m\right]} = \frac{1}{2} - \frac{1}{4}\frac{1}{\frac{1}{\beta}\frac{1}{n}\sum_{i\in\mathcal{N}}\beta(x_i - z_m)^2\beta + \frac{1}{e}\frac{1}{n}\sum_{i\in\mathcal{N}}e(t_i - a_m)^2 e} \tag{52}$$

where the denominator in the last term follows from (28).

As before recall that j's expected vote share at the electoral mean is $V_j^m = \rho_j^m$. Therefore, using (51) we have that the *sufficient condition* for j to adopt (z_m, a_m) as electoral campaign—when all other candidates adopt (z_m, a_m) as their electoral campaign—is that

$$\mathcal{V}_{sp}^m < \rho_j^m = V_j^m, \tag{53}$$

i.e., V_j^m must be *high* enough at $(\mathbf{z}_m, \mathbf{a}_m)$ for j to adopt to the electoral mean as its electoral campaign. ■

For the necessary part of the proof, assume $(\mathbf{z}_m, \mathbf{a}_m)$ is a *weak* local Nash equilibrium of the election. Then for all $j \in \mathcal{C}$, j's Hessian matrix evaluated at (z_m, a_m) must be negative semi-definite. This implies that $Tr(\mathbf{D}^2 V_j^m) \leq 0$ which is true if and only if $Tr(\mathcal{C}_j^m) \leq 0$ for all $j \in \mathcal{C}$. If $\mathcal{V}_{np}^m > V_j^m$ for some $j \in \mathcal{C}$ then it must be the case that $Tr(\mathcal{C}_j^m)$ must be strictly positive violating the weak Nash equilibrium condition. This completes the proof of necessity. ■

10.2 Proof of Lemma 2

We only need the proof for candidate j. Using (52) and (46), $\mathcal{V}_{sp}^m > \mathcal{V}_{np}^m$ when

$$\frac{1}{2} - \frac{1}{4}\frac{1}{Tr\left[(\mathbf{B})^{-1}\boldsymbol{\nabla}^m\right]} > \frac{1}{2} - \frac{1}{4}\frac{Tr(\mathbf{B})}{Tr\left(\boldsymbol{\nabla}^m\right)}$$

This translates into

$$Tr(\mathbf{B})Tr\left[(\mathbf{B})^{-1}\boldsymbol{\nabla}^m\right] > Tr\left(\boldsymbol{\nabla}^m\right)$$

which using (23), (28) and (27) and after some manipulation translates into

$$\left(1+\frac{e}{\beta}\right)var(x) + \left(1+\frac{\beta}{e}\right)var(t) > var(x) + var(t).$$

After multiplying this inequality by $[\beta e]^{-1} > 0$ and some manipulation, we get

$$var(x) + var(t) > 0,$$

so that $\mathcal{V}_{sp}^m > \mathcal{V}_{np}^m$ always as required and this holds for all $j \in \mathcal{C}$. ■

10.3 Comparative statics

To find the effect of β on the *sufficient* pivotal vote share, when all candidates adopt the electoral mean $(\mathbf{z}_m, \mathbf{a}_m)$ as their campaign strategy, $\mathcal{V}_{sp}^m$ in (22), take the derivative of $\mathcal{V}_{sp}^m$ wrt β, ceteris paribus, i.e.,

$$\frac{\partial \mathcal{V}_{sp}^m}{\partial \beta} = \frac{1}{4}\frac{\frac{1}{\beta^2}var(x)}{\left[\frac{1}{\beta}var(x) + \frac{1}{e}var(t)\right]^2} > 0,$$

so that $\mathcal{V}_{sp}^m$ increases as β increases. A similar analysis would show that $\mathcal{V}_{sp}^m$ increases when e increases.

The effect of β on the *necessary* pivotal vote share, $\mathcal{V}_{sp}^m$ in (22), i.e., the derivative of $\mathcal{V}_{np}^m$ wrt β, ceteris paribus, is given by

$$\begin{aligned}
\frac{\partial \mathcal{V}_{np}^m}{\partial \beta} &= -\frac{1}{4}\frac{var(x) + var(t) - (\beta+e)\frac{1}{n}\sum_{i\in\mathcal{N}} 2\beta(x_i - z_m)^2}{[var(x) + var(t)]^2} \\
&= -\frac{1}{4}\frac{var(x) + var(t) - 2var(x) - 2\frac{1}{n}\sum_{i\in\mathcal{N}} \beta(x_i - z_m)^2 e}{[var(x) + var(t)]^2} \\
&= -\frac{1}{4}\frac{-var(x) + var(t) - 2\frac{e}{\beta}var(x)}{[var(x) + var(t)]^2} \\
&= \frac{1}{4}\frac{(1+2\frac{e}{\beta})var(x) - var(t)}{[var(x) + var(t)]^2}
\end{aligned}$$

where the fourth line follows after multiplying by $\frac{\beta}{\beta}$. The necessary pivotal vote share increases when β increases iff the numerator in the last line of the above equation is positive, i.e., iff

$$(1+2\frac{e}{\beta})var(x) > var(t)$$

Adding $var(x)$ on both sides of the above inequality gives

$$2\frac{\beta+e}{\beta}var(x) > var(x)+var(t)$$

The RHS of this inequality equals $Tr\left(\boldsymbol{\nabla}^m\right)$, given in (27). Thus, when β increases the necessary pivotal vote share increases when

$$\frac{\beta+e}{\beta}var(x) > \frac{1}{2}Tr\left(\boldsymbol{\nabla}^m\right),$$

i.e., when the variance of voters' ideal policies weighted by $\frac{\beta+e}{\beta}$ is larger than half the aggregate variance of voters' preferences about the electoral mean.

A similar analysis shows that $\mathcal{V}_{np}^m$ rises as e increases iff

$$\frac{\beta+e}{e}var(x) > \frac{1}{2}Tr\left(\boldsymbol{\nabla}^m\right).$$

Suppose voter's preferences become more dispersed along the policy dimension while maintaining the electoral mean (z_m, a_m) constant, so that the distribution of voters' ideal policies undergoes a mean preserving spread. This means that $var(x)$ increases but the electoral mean (z_m, a_m), $var(t)$, β and e remain unchanged.

We now examine what happens to the pivotal vote shares as $var(x)$ or $var(t)$ increase. The effect of higher $var(x)$ or $var(t)$ on the sufficient pivotal vote share is given by

$$\begin{aligned}
\frac{\partial\mathcal{V}_{sp}^m}{\partial var(x)} &= \frac{1}{4}\frac{1}{\beta}\frac{1}{\left[\frac{1}{\beta}var(x)+\frac{1}{e}var(t)\right]^2} \\
\frac{\partial\mathcal{V}_{sp}^m}{\partial var(t)} &= \frac{1}{4}\frac{1}{e}\frac{1}{\left[\frac{1}{\beta}var(x)+\frac{1}{e}var(t)\right]^2}
\end{aligned}$$

which are both always positive. That is, a mean preserving increase in $var(x)$ or $var(t)$ increases the sufficient pivotal vote share. Thus, making it more difficult for the electoral mean to be a strick LNE of the election.

The effect of a mean preserving increase in $var(x)$ or $var(t)$ on the *necessary* pivotal vote in (22) is given by

$$\frac{\partial\mathcal{V}_{np}^m}{\partial var(x)} = \frac{\partial\mathcal{V}_{np}^m}{\partial var(t)} = \frac{1}{4}\frac{(\beta+e)}{[var(x)+var(t)]^2} > 0.$$

The necessary pivotal vote share also increases in the variance of voters' preferences along any dimension. The necessary condition for convergence to the electoral mean is then harder to meet.