CSEF
Centre for Studies in Economics and Finance

WORKING PAPER NO. 668

Simplistic Rhetoric and Poe's Law

Giovanni Andreottola and Elia Sartori

February 2023

This version July 2024

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS AND STATISTICS – UNIVERSITY OF NAPLES FEDERICO II
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unina.it
ISSN: 2240-9696

CSEF
Centre for Studies in Economics and Finance

WORKING PAPER NO. 668

Simplistic Rhetoric and Poe's Law

Giovanni Andreottola* **and Elia Sartori** †

Abstract
We study the use of simplistic arguments in political communication, developing a model of mobilization through rhetoric with naive and sophisticated voters. Politicians sometimes choose simplistic arguments to appear more competent, exploiting what we call Poe's Law, i.e., the uncertainty on whether the argument used by the politician reflects her competence or is 'degraded' to meet naive voters' preferences. We compare the Bayesian game with one where sophisticated voters conceptualize Poe's Law assuming that the politician communicates to a fully naive crowd, effectively dismissing their fellow citizens' cognitive abilities. Dismissal induces an overly simplistic political debate.

Keywords: Simplistic rhetoric, Dismissal, Poe's Law, Populism.

JEL classification: D72, D82, D83, D91.

Acknowledgments: We thank Alberto Bennardo, Federico Boffa, Gabriele Gratton, Giovanni Immordino, Barton Lee, Stephen Morris, Massimo Morelli, Paolo Piacquadio, Giorgio Saponaro, Karl Schlag for helpful discussions and comments. We also thank seminar audiences at University of Naples Federico II, the European University Institute, the University of Linz, Paris Dauphine University, the Lindau Meetings on Economic Sciences, the University of Vienna, the University of St. Gallen, the IBEOWorkshop in Alghero. Alessandro Sciacchetano and Francesco Crispano provided invaluable research assistance.

* Johannes Kepler University and CSEF. E-mail: giovanni.andreottola@gmail.com

† CSEF and Università di Napoli Federico II. E-mail: elia.sartori@unina.it

1 Introduction

Electoral campaigning is a balancing act between displaying competence and using rhetorical arguments which resonate with the electorate: whereas some voters are mostly interested in selecting a valent politician, i.e., someone who is able to govern effectively, there are also voters who are sensitive to the type of rhetorical argument employed by the politician, due to ideological or instinctive preferences. For example, such voters may want to hear a nationalistic argument, or alternatively an argument that stresses the politician's opposition to racism. Even if the possibility of targeting specific voters does exist, the bulk of electoral campaigns is carried out through mainstream media, or it consists of publicly available content which all voters can access.[1] Therefore, candidates need to disseminate a unified message that resonates well with both types of voters. Frequently, however, this may prove challenging: making a complex speech to prove one's valence may alienate part of the electorate, and analogously pursuing the 'gut feelings' of the instinctive electorate may alienate voters interested in valent policymakers.

Formalizing this trade-off is crucial to understand the incentives that shape the democratic debate and its recent transformations. Jordan et al. (2019) provide evidence that the political debate in the United States has become more simplistic over time: this trend is hard to reconcile as merely a consequence of either the general population or politicians becoming more ignorant, considering for example that the educational attainment of both has increased.[2]

In this paper we propose an explanation of simplism in the political debate based on candidates signaling valence to a heterogeneous audience, especially when voters are dismissive. At the heart of our analysis is the interpretation dilemma of a voter trying to gauge the candidate's valence: is the candidate (or politician, she) choosing the most complex argument that she is capable of defending, or is she instead choosing a simpler argument to 'get out the vote'? This dilemma closely resembles what is known as Poe's Law,[3] a term usually employed to describe interactions in the digital space, for example on social networks.

[1]Recent technological developments increased the ability of politicians to target voters, but at the same time also decreased their ability to ensure that a message is kept private: everything can be recorded and shared with wider audiences, and social media allow people with vastly different backgrounds to observe the content that each other engages with.

[2]For the educational attainment of US congressmen, see https://shorturl.at/hjzEZ; for the general population educational attainment, see https://shorturl.at/axyX7

[3]https://en.wikipedia.org/wiki/Poe%27s_law

Poe's Law: *Without a winking smiley or other blatant display of humor, it is utterly impossible to parody a Creationist in such a way that someone won't mistake for the genuine article.*

In the context of political communication, a *genuine* argument is used by a candidate who lacks the ability to defend a more complex one, while a *parody* argument is used to impress the instinctive electorate even though the candidate would be capable of employing a more complex one. In principle, simplistic arguments can be used both as a parody — by high valence candidates that are constrained by their instinctive electorate—, and genuinely — by low valence candidates that strategically abuse them to pool with high valence candidates and get a favorable valence update.

This naturally begs a set of questions that we aim to address in this paper: Does the type of rhetoric employed by politicians reflect the preferences of instinctive voters or the need to appear competent? Is it possible that political rhetoric overshoots the simplism demanded by instinctive voters? How does rhetorical simplism change with the share of voters who only reward the politician for her valence?

We address these questions in a model where candidates employ rhetorical arguments to mobilize a cognitively heterogeneous audience. Politicians are privately informed of their own valence, as well as of the preferences of the instinctive electorate. They campaign by choosing (i) what argument to defend and (ii) the degree of complexity with which to defend it. Arguments are ranked according to a complexity order, loosely defined as the valence needed to convincingly defend them or, equivalently, as the maximal valence that a politician can display while campaigning with them. A share of the electorate, which we call naive, takes the politician's speech at face value and gets instinctively mobilized by the type of argument presented.[4] The residual share of the electorate, which we call sophisticated, is instead aware of the purely rhetorical nature of the arguments employed and uses political communication only to infer the valence of the politician.

In the (Bayesian) equilibrium of this communication game a candidate chooses, given her private information, a speech to maximize support in her heterogeneous electorate; such support depends on how the argument chosen resonates with naive voters and on the updating by sophisticated voters, which has to be consistent with the candidate's strategy, creating a feedback loop that brings us back to Poe's Law. We show that in equilibrium some arguments must (but not all arguments can) be used strategically by candidates: a certain speech can be made both genuinely, by low valence candidates, and as a parody, by high valence candidates appealing to a simple-minded naive electorate. In equilibrium, the candidate enjoys the benefit of the doubt, as it (typically) cannot be ruled out that

[4]Djourelova (2022) shows that banning the use of a politically charged word affects the support of immigration policies. This provides evidence of the relevance of instinctive mobilization through the types of arguments employed.

she is better than the rhetoric she uses. In the words of Brennan (2015): "When you see politicians saying dumb things, remember that these politicians are not fools. [...] They say dumb things because they expect voters want to hear dumb things". Not only our model rationalizes this common view — predicting that some politicians degrade their rhetoric to match what voters want to hear — but it goes a step further: it predicts that some politicians exploit this perception and degrade their rhetoric even more than the naive voters require. This allows them to masquerade their poor valence behind the convincing defense of a simplistic argument. Paraphrasing the citation above, in our model politicians can be fools who say things that are as dumb as the sophisticated audience thinks the naive audience is. Of course, in equilibrium the possibilities to masquerade poor valence are limited by a consistency requirement that adjusts the inference drawn from each speech to the valence effectively possessed by the candidates that pursue it. The oversimplification that we pointed at is an interesting oddity that can occur in equilibrium (i.e., for *some* candidate types), but that is always reversed in the aggregate (i.e., averaging *across* candidate types). Valence signaling yields to a complexification of the political discourse, the more so the larger the share of sophisticated voters. This aggregate result is in line with the general intuition that, in order to be effective, signaling must be costly: equilibrium unravels if it is "too simple" for politicians to be perceived as competent.

Sophisticated voters make correct inference only insofar as they correctly conceptualize the communication game that the candidate is playing. Since a key element of this game is the share of sophisticated voters, a key implicit assumption is that each sophisticated voter knows how many other voters are also sophisticated. However, there are reasons to suspect that even voters that recognize the rhetorical nature of the political debate (those that we call sophisticated) might misperceive the share of the electorate having the same awareness. First, it is well known (see Hoorens (1993) and Kruger and Dunning (1999) for psychological evidence and Malmendier and Taylor (2015) for economic implications) that agents overestimate their cognitive abilities relative to others. Second, and more directly related to our application, there is evidence of a declining trust – shared across political sides – in the degree of fellow citizens' 'political widsom'.[5] Third, the mistrust in other citizens can be a direct consequence of populist narratives (Section 5) whereby a politician pretends to address a uniform crowd of 'common people' eager to be mobilized by their message, or it can result from the media depicting a candidate as exclusively targeting a given social group, such as working-class voters in the case of Donald Trump, as argued by Carnes and Lupu (2021).

To address this concern and verify the robustness of our equilibrium predictions, we consider a modification of the model in which sophisticated voters believe that all their

[5]See https://shorturl.at/iuwS8 for some survey evidence.

fellow citizens are naive. Such voters, which we call dismissive, update as econometricians of a misspecified game, where the politician has private information about the ideological preference of a fully naive electorate and strategically targets it. This interpretation is particularly appealing once we consider the widespread 'identification with the leader' occurring in populist parties, fostered by a considerable effort in convincing voters that the leader's arguments target the views of the 'common man' (see, e.g., (Mudde, 2004)). Comparing outcomes in the Bayesian and dismissive equilibria, we shed light on the importance of collective awareness in shaping rhetorical simplism: sophistication might become a propeller of populist rhetoric, rather than a countervailing force to it, if it is connected with an underappreciation of the level of sophistication in the electorate. To make this point, we compare three key outcomes.

The first is the *level* of average simplism. Recall that the Bayesian model predicts an average complexification, relative to naive voters' preferences, of the political debate. Under dismissive updating, exactly the opposite happens and political debate (weakly) exceeds the naive voters benchmark. This is natural: since dismissive voters ascribe any rhetorical simplism to naive voters' preferences rather than to the politician's own limitations, politicians have the incentive to employ very simplistic arguments.

Second, we compare how average simplism *changes* with the share of sophisticated voters. In the Bayesian model, the larger the share of sophisticated voters, the less simplistic the political debate becomes. Under dismissive updating, instead, simplism is non-monotonic in the share of sophisticated agents, first increasing and then reverting back to baseline. At first, the dismissive updating premium is exploited by low valence candidates that choose an argument they can effectively defend. This argument is typically simpler than what naive voters prefer, which drives the initial average simplification. Once all low valence politicians have exploited the updating premium, increasing the share of sophisticated voters only leads more valent candidates to complexify their argument, which eventually offsets the simplification created by low valence candidates.

Finally, we show that dismissal affects the occurrence and interpretation of *extreme simplism*. In the Bayesian model, the most simplistic arguments are always interpreted as harshly as possible: equilibrium cannot sustain an updating premium for overly simplistic arguments, hence a politician employing such arguments is outing herself as an incompetent. Dismissive voters, instead, grant an updating premium to all arguments, justifying the abuse of extreme ones. This difference is especially important when we try to rationalize why very extreme and toxic arguments emerge in the political debate. In the Bayesian equilibrium, those arguments are only used when naive voters have an intrinsic preference for them (and the candidate is not very valent); under dismissal, instead, they are also used, absent a demand by naive voters, by candidates who exploit the inconsistent updating of sophisticated voters. Therefore, if it is unrealistic that the instinctive electorate is mobilized by racist slur, then hearing it suggests that the sophisticated electorate might be dismissive. We further detail this interpretation in Section 5.

The rest of the paper is organized as follows. We conclude this introductory section discussing the relevant literature. Section 2 presents the communication game and solves its polar cases. Section 3 formalizes Poe's Law and studies the Bayesian equilibrium. Section 4 argues why dismissal is a valid concern in our setting and studies how a politician communicates to a dismissive audience. Section 5 presents alternative interpretations of our model, and Section 6 concludes. All proofs, as well as the formal development of some results, are relegated to the Appendix.

1.1 Literature Review

To study Poe's Law in a formal setting, we develop a simple but novel model of communication. The politician has a two-dimensional type and sends a two-dimensional message – an argument-complexity pair – which voters use to form expectations about her valence. As in a standard (Spence, 1978) signaling model, the cost of employing different arguments varies with the politician's type, specifically with the bliss point of the naive crowd she is facing. Along the complexity dimension, instead, the politician can send verifiable information about her valence,[6] as in the disclosure literature (Grossman, 1981). In our model, signaling and disclosure concerns interact – and generate Poe's Law – since the complexity that a politician can display is not only constrained by her type (valence) but also by the argument she chooses. Therefore, when choosing an argument the politician is determining both the signaling cost she pays and her ability to disclose valence. The presence of two dimensions of private information bears a resemblance with Frankel and Kartik (2019): Poe's Law can be thought of as the dilemma stemming from the muddling of information about valence with information about the preferences of naive voters. Another key aspect of our model is the heterogeneity of the audience. This connects it to Farrell and Gibbons (1989), who study a cheap talk game in which a sender deals with two receivers and compare public and private communication. In our model, which is not a standard cheap talk framework, the politician would always prefer private communication, and the trade-off arises precisely from the necessity of communicating publicly.

The key characteristic of rhetoric we focus on in our paper is complexity/simplism, which in turn is one of the most defining traits of populism, together with the conflict between different groups of citizens. Therefore, our paper contributes to the theoretical literature on populism (spurred by Acemoglu et al. (2013) and Di Tella and Rotemberg

[6]That is, using complexity v is hard evidence of possessing valence $\omega \geq v$. This requirement disciplines off-path updating.

(2018)), and specifically to a recent literature focusing on the interaction of simplism and populism. Levy et al. (2022) show that the presence of citizens with a misspecified simplistic worldview (reminding of our dismissive agents) creates cycles of simplistic/populist policymaking. In Morelli et al. (2021) there is no misspecification, but simplistic policies are a substitute for commitment power.[7]

Despite our focus being on simplism rather than extremism, the two are closely related, as we explain in Section 5. Hence, our paper is also related to the literature studying how extremist policy proposals can emerge as a consequence of candidates' strategic choices. This literature highlights several possible reasons for strategic extremism: two recurring ones are mobilization (of voters, as in Glaeser et al. (2005), or donors, as in (Alesina and Holden, 2008)) and valence signaling (Kartik and McAfee (2007) and Andreottola (2021)).[8] Despite also revolving around mobilization and valence signaling, our paper sheds light on a novel rhetorical foundation of strategic extremism, Poe's Law, which emerges as a consequence of the bound on the valence that each argument can convey.

When considering (Section 4) the robustness of our results to dismissive updating, we build on a large literature establishing that agents misperceive their cognitive abilities relative to others (see Malmendier and Taylor (2015) for an overview). Several contributions explore the impact of such overconfidence on various economic outcomes, ranging from political behavior (Ortoleva and Snowberg (2015)), to market entry (Camerer and Lovallo (1999)), innovation (Galasso and Simcoe (2011)) and financial markets. We contribute to this literature showing that dismissing others' cognitive abilities has effects on the type of rhetorical arguments employed in the political debate. We further discuss the behavioral assumptions behind dismissive updating in Section 4.1, describing also the connections with level-k thinking (Nagel, 1995) and cognitive hierarchy (Camerer et al., 2004) models. Finally, what we call dismissal might also be interpreted as a form of stereotype according to the notion of Bordalo et al. (2016): dismissive voters take the naive component as the representative type in the rest of the polity, ignoring the fact that some peers are sophisticated. By investigating how stereotyping affects political rhetoric, our study is somehow complementary to Bonomi et al. (2021), who instead focus on the consequences of stereotyping for identity politics and inter-group conflict. There are also other papers featuring naive or 'impressionable' voters. One is Grossman and Helpman (1996), where impressionable voters vote based on the campaign effort of candidates rather than the utility they would derive from the policies they propose; another one is Andonie and

[7]Policy simplism is also studied by Levy and Razin (2012), who find that simple policies are more likely to win when attention is scarce, but not too scarce. Our setup delivers a channel for simplistic policies to be successful even when voters are not subject to limited attention constraints.

[8]Other sources of strategic extremism are dynamic concerns of issue ownership (Eguia and Giovannoni, 2019) or (anti-)pandering incentives Kartik et al. (2015).

Diermeier (2019), where a voter's impression of a candidate consists of both objective components — such as the utility of the policies proposed — and random components.

2 The Model

An office motivated candidate (she) chooses among a set of arguments with the aim of maximizing the support of a cognitively heterogeneous population of voters. Some voters (naive) exclusively respond to the type of argument that the candidate uses, while the rest of the voters (sophisticated) look at the candidate's speech to infer her valence ω. A speech is an argument-complexity pair (l, v), interpreted as choosing to defend argument l making a case of complexity v. The candidate makes a single speech, meaning that she cannot target her message to the cognitively heterogeneous crowd she is facing.[9] Therefore, she needs to find a speech that resonates well with naive voters and at the same time convinces sophisticated ones that she is a valent candidate. At the heart of our analysis is the inferential problem of a sophisticated voter (he). In his attempt to back out the valence of the candidate, the voter must take into account that she is subject to a twofold constraint: she can neither display a valence that exceeds what she is *given by nature*, nor she can display one that exceeds the bound associated to the argument that she *chooses to defend*. Upon observing the succesful defence of a relatively simple argument, the voter is faced with an interpretative dilemma that very much resembles Poe's Law; the resolution of such dilemma determines the incentives of the politician and feeds back into the updating, creating a complex interaction whose analysis is at the heart of Sections 3 and 4.

In the current section we formalize the model, present the interpretative dilemma (Poe's Law), and define the simplism metrics that will be used throughout the paper as the relevant outcomes of political communication. To build intuition we also discuss the outcomes in the polar economies that are populated by only naive ($k = 0$) and only sophisticated ($k = 1$) voters.

In terms of notation, we use Greek letters (λ, ω) to denote the politician's type, Roman letters (l, v) to denote possible actions (speeches) and boldface notation $(\boldsymbol{l}, \boldsymbol{v})$ to denote the policy function, i.e. the mapping from types to actions.

[9] If the politician could send a different message to the naive and sophisticated crowds, she would face two independent problems, each corresponding to one of the polar cases that we discuss at the end of this section. We focus our analysis on the more interesting – and arguably more realistic when it comes to modeling communication choices of large parties – case where politicians have to speak through a single platform and cannot tailor communication to the cognitive ability of their audience. Sticking to Poe's Law's wording, politicians cannot put a "winking smiley" on their speech to the naive agents and independently signal their valence to the sophisticated ones.

2.1 Primitives

Type and Action Space: The politician is characterized by a two-dimensional type $(\lambda, \omega) \in [0,1]^2$, which is her private information. The location type λ represents the naive voters' bliss point, while ω is the politician's valence. We will assume that ω and λ are independently and uniformly distributed. Given her type, the candidate chooses a *speech*, namely an *argument-complexity* pair (l, v). In choosing such pair, she is free to choose any argument l, but she is subject to two constraints: first, she cannot make a speech whose complexity v exceeds her own valence ω. Second, each location l is associated with an upper bound $b(l) = 1 - l$ on the complexity that can be used while defending argument l. The bound is decreasing, reflecting the idea that the more simplistic an argument, the less complexity it can accommodate, and the lower the valence that be conveyed by the speech.[10] Formally, she chooses from:

$$\mathcal{A}(\omega) = \{(l, v) \in [0,1]^2 \text{ s.t. } v \leq \min\{\omega, b(l)\}\} \tag{1}$$

Notice that action sets are ordered in valence, meaning that $\omega_H > \omega_L$ implies $\mathcal{A}(\omega_L) \subset \mathcal{A}(\omega_H)$. The space of possible outcomes of political communication, $\mathcal{A} = \bigcup_{\omega \in [0,1]} \mathcal{A}(\omega) = \mathcal{A}(1)$ is therefore the lower triangle in the unit square (Figure I).

Due to the qualitative difference in the inference problem they induce, we partition the space $\mathcal{A}$ in speeches that lie on the $b(\cdot)$ boundary (effective speaking strategies) and in its interior (ineffective speaking strategies). We use the following taxonomy to classify the politician's strategies.

Definition 1 (Policy Taxonomy).

- *A politician* ***speaks effectively*** *if she chooses a speech on the bound, i.e.*

$$(l, v) \in \mathcal{A}_{ES} := \{(l, v) : v = b(l)\}.$$

  - *A politician* ***matches her valence*** *if she speaks effectively and* $v = \omega$.
  - *A politician* ***degrades her valence*** *if she speaks effectively and* $v < \omega$.

- *A politician* ***speaks ineffectively*** *if she chooses a speech below the bound, i.e.*

$$(l, v) \in \mathcal{A}_{\neg ES} := \{(l, v) : v < b(l)\}.$$

  - *A politician* ***exposes her valence*** *if she speaks ineffectively and* $v = \omega$.

[10] The linear specification is used for expositional convenience. More precisely, it will be clear from the analysis that the bound is not separately identified from the joint distribution of (λ, ω), and in particular of the conditional expectations $\mathbb{E}[\omega | \omega > b(l)]$ for $l \in [0,1]$.

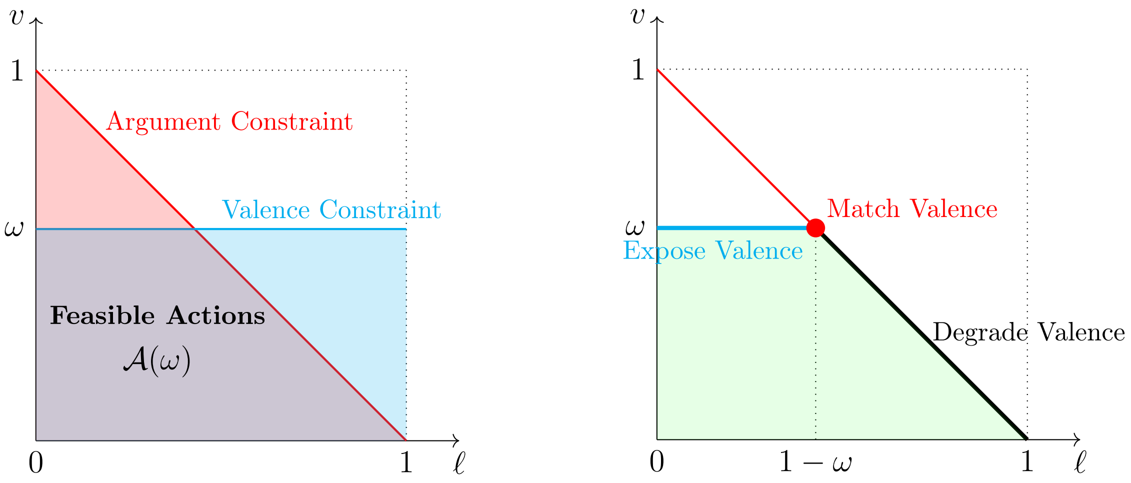


Figure I: Feasible action set and taxonomy of strategies for a politician with valence ω. The green area (including Expose valence) denotes the feasible ineffective speeches, while the black line (Degrade valence) and red point (Match valence) are the feasible effective speeches.

Definition 1 suggests an equivalent interpretation of the bound $b(l)$ as the minimal valence required to *effectively defend* argument l. Although candidates can attempt the defense of any type of argument, they can succeed only at arguments with a sufficiently high l: the larger is the valence of the candidate, the broader is the set of arguments she can effectively defend. What is crucial for our results is that anyone that looks critically at the political speech (in particular sophisticated voters, as discussed in the following paragraph) is able to discern whether an argument could have been defended more effectively or not.[11]

Voters: There are two types of voters (or citizens/agents). A proportion k of sophisticated voters and a proportion $1-k$ of naive voters. The latter type is characterized by a bliss point $\lambda \in [0,1]$: throughout the paper, we stick to the interpretation of λ as a measure of the prevalence of preferences for simplistic political arguments. In Section 5 we offer alternative interpretations.[12] Naive voters respond solely to the location chosen by the politician, supporting her with probability:

$$\pi_N = 1 - (l - \lambda)^2 \tag{2}$$

[11] This amounts to requiring that the bound $b(l)$ is common knowledge in the game between the politician and sophisticated voters.

[12] We assume that all naive agents have the same bliss-point λ, and that λ is known to the politician, just for ease of exposition. Indeed, we only need that λ is the conditional expectation of naive voters' bliss point given the information available to the candidate, and that such information is not available to sophisticated voters. The analysis extends if the true bliss point is stochastic given λ, with conditional variance σ^2. In that case, support from naive voters could be split in bias and variance components, to yield $\pi_N = 1 - (l-\lambda)^2 - \sigma^2$ where the variance term is just a constant in the maximization.

Sophisticated agents do not 'instinctively respond' to the location chosen by the politician; on the contrary, they support the candidate with probability

$$\pi_S = \mathbb{E}\left[\omega \,|(l, v)\right]$$

equal to the expected valence of the candidate, given the speech she chose. Since the operator $\mathbb{E}$ depends both on equilibrium objects as well as primitive behavioral assumptions, we postpone its precise characterization to the respective Sections (3 for rational and 4 for dismissive voters) in the analysis.

Communication: Given her type (λ, ω) the candidate chooses a speech to solve

$$\Pi\left(\lambda, \omega\right) = \max_{(l,v)\in\mathcal{A}(\omega)} (1-k)\pi_N\left(\lambda, l\right) + k\pi_S\left(l, v\right). \tag{3}$$

We define $(\boldsymbol{l}, \boldsymbol{v}) : [0,1]^2 \to \mathcal{A}$ the policy function of the problem (3). Through this function, political communication maps the unit square (the politician's type space) into the lower triangle defined by the bound $b\left(\cdot\right)$. Notice how the two coordinates of the politician's type enter her problem: only the naive preferred argument λ directly affects her payoff (through π_N), while only the valence ω determines the set of speeches $\mathcal{A}\left(\omega\right)$ that are feasible. Likewise, only the choice of argument l determines the support from naive voters, while the whole speech (l, v) is used by sophisticated voters in their inference.

Simplism: Recall our motivation is to study the prevalence of simplistic arguments in the political debate. To this end, we define a metric that measures how much each argument is strategically exploited by the politician.

Definition 2 (Simplism metrics). *For a given argument l, the associated argument-wise simplism is*

$$\bar{s}_l = l - \mathbb{E}[\lambda|\boldsymbol{l}(\lambda, \omega) = l].$$

Aggregate simplism is

$$\bar{s} = \mathbb{E}[\bar{s}_l]$$

Argument l is a ***catalyst*** *(resp. a* ***repellent****) of simplism if $\bar{s}_l > 0$ ($\bar{s}_l < 0$). Political communication catalyzes (resp. repels) simplism if $\bar{s} > 0$ ($\bar{s} < 0$).*

In words, the value of $\bar{s}_l$ answers the following question: if we observe argument l, to what extend are we exceeding, on average, the simplism required by naive agents? It therefore measures the excess of simplification induced by strategic communication relative to a demand-driven benchmark. An argument l is a catalyst (resp. repellent) of simplism if it is adopted by politicians who, on average, face a naive-preferred argument

λ lower (resp. higher) than l. Analogously, $\bar{s} = \mathbb{E}[l] - \mathbb{E}[\lambda]$ measures the average degree of simplification in the politician's argument compared to what is on average required by naive voters. We now highlight the main forces at play and study the outcomes of interest in the two polar cases of all-naive voters ($k = 0$) and all-sophisticated voters ($k = 1$).

It is worth pointing out that characterizing the simplism of the political speech may be relevant not only from a positive perspective, but also from a normative one: for example, we may assign a negative welfare value (which in our model is internalized neither by the politician, nor by voters) to simplism, which may involve discrimination, toxic language, and which may instigate acts of bullying and violence such as hate crimes.

2.2 Fully Naive Benchmark ($k = 0$)

When $k = 0$, the politician solves the simple location problem

$$\max_{l \in [0,1]} 1 - (l - \lambda)^2 \tag{4}$$

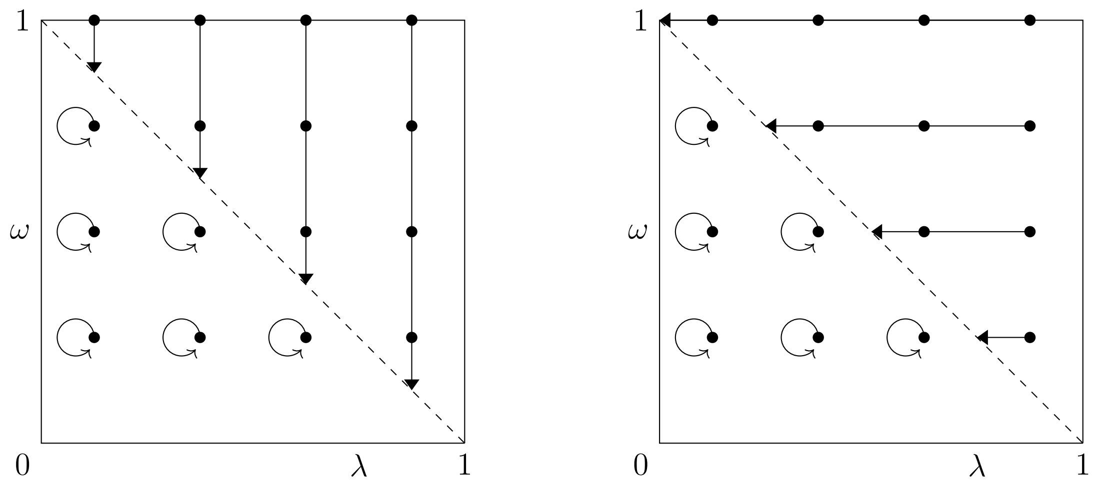
that gives $l = \lambda$ as the optimal policy: when all agents are naive the politician tracks their preference and is supported with probability 1. In this environment the degree of valence displayed does not affect the politician's payoff.[13] However, as anticipated above, the natural assumption – which also makes the equilibrium robust to small perturbations in k – is to select the equilibrium in which a politician of type (λ, ω) chooses to display $\min\{\omega, b(\lambda)\}$. With $k = 0$, therefore, the policy function is $(\boldsymbol{l}, \boldsymbol{v})(\lambda, \omega) = (\lambda, \min\{\omega, b(\lambda)\})$. As highlighted in Figure IIa, the chosen location coincides pointwise with the naive-preferred argument λ. Therefore, recalling the terminology introduced by Definition 1, we have that $\bar{s}_l = 0$ for all $l \in [0, 1]$ and a fortiori $\bar{s} = 0$: no argument can either be a catalyst or a repellent of simplism, simply because politicians track the preferences of naive voters one for one. This benchmark case corresponds to the purely demand-driven explanation of simplism in political communication.

An important observation related to this fully naive benchmark regards the type of inference a rational outsider ("econometrician") would make about the politicians' valence. After observing an effective speech of complexity v, i.e., speech $ES(v) = (1 - v, v)$, an econometrician would only know that the candidate has at least valence v and would therefore attribute the candidate an expected valence of

$$\mathbb{E}[\omega | ES(v)] = \mathbb{E}[\omega | \omega \geq v] = \frac{1 + v}{2}. \tag{5}$$

[13]Indeed, naive voters perceive political communication as one-dimensional (the argument employed) and instinctively respond to it. Mobilizing a naive audience makes the communication problem a trivial tracking exercise as, quoting Rachman (2020) "For the simplist, the slogan is the platform".

Effective speeches grant an updating premium because when all agents are naive the politician is "forced" to choose $l = \lambda$, essentially constraining her valence-showcasing possibilities. In light of this, one natural foundation of dismissive updating (see Section 4) is that the agent behaves as if he was able to make sophisticated valence inference from a game with only naive agents (i.e., $k = 0$) and therefore uses (5) as his updating rule.



(a) Location-complexity choices for $k = 0$.

(b) Location-complexity choices for $k = 1$.

Figure II: Optimal location-complexity choices in the polar cases $k = 0$ and $k = 1$. Points represent types, arrows represent actions.

2.3 Fully Sophisticated Benchmark ($k = 1$)

The opposite polar case occurs when $k = 1$, that is, when all voters are sophisticated. In this case politicians do not care about mobilizing naive voters and choose speeches of the form (l, ω), for some argument $l \leq 1 - \omega$ that allow them to fully display their valence. Communicating to a sophisticated audience is a trivial valence-display exercise where arguments only serve as vehicles to display valence in full. Indeed, an equilibrium is that all politicians choose the most complex argument $l = 0$, and expose their valence by speaking ineffectively.[14] However, similar to the case of $k = 0$, it is more natural to select $l = \min\{\lambda, 1 - \omega\}$, i.e., to assume that politicians maximize valence updating breaking ties in favor of locations that are closer to λ. By simple computations we obtain that $\bar{s}_l = \frac{l-1}{4} \leq 0$ and $\bar{s} = -\frac{1}{6}$, meaning that all arguments are repellents of simplism and so a fortiori is political communication. Repeating the "econometrician" exercise, inference conditional on effective speaking is

$$\mathbb{E}[\omega | ES(v)] = v, \tag{6}$$

[14] As in the hypothetical contingency described by Brennan (2015): "If voters were well-informed, dispassionate policy-wonks, then political campaigns would resemble peer-reviewed economics journals."

In clear contrast with the polar opposite $k = 0$ case, updating is now maximally harsh: whenever a politician speaks effectively, the econometrician knows she has exactly the displayed valence, which is indeed the minimal valence consistent with the effective defense of the employed argument. Figure II displays the policies chosen by politicians in the polar cases.

3 Bayes Nash Equilibrium

In this section we characterize the equilibria of the Bayesian game where the candidate privately observes (λ, ω) and then chooses a speech to solve (3). Because of the triviality of updating following an ineffective speech (Lemma 1), the key endogenous variable that pins down the equilibrium is the function $f(v)$ that associates each effective speech with the expected valence that sophisticated voters attribute to candidates that make that speech. We show how candidates best respond given f (Lemma 2), compute the expectation consistent with politicians' actions (Poe's Law, Lemma 3) and by a fixed point argument

$$f \overset{\text{Optimality}}{\longrightarrow} (\boldsymbol{l}, \boldsymbol{v}) \overset{\text{Consistency}}{\longrightarrow} f$$

we characterize (Proposition 1) equilibrium expectations: the function f corresponds to the identity function below a triviality threshold $v^{\star}$, and then it jumps, coinciding from that point onwards with the solution to an initial value problem parametrized by k. The set of feasible thresholds increases with k and for large values only the degenerate all-sophisticated equilibrium (Sec. 2.3) survives. Proposition 2 studies the comparative statics of the rational model, showing that simplism is always lower than the demand-driven benchmark, and that it decreases with the share of sophisticated agents.

3.1 Equilibrium Definition

An equilibrium requires that (every type of) politician optimizes given the voters' update and that the voters' update is consistent with the politician's choice. Formally,[15]

Definition 3. *An equilibrium is a pair of functions:*

$$\tilde{f} : \mathcal{A} \to [0,1] \textit{ and } (\boldsymbol{l}, \boldsymbol{v}) : [0,1]^2 \to \mathcal{A} \textit{ such that:}$$

[15] Notation reminder: Greek letters (λ, ω) denote politician's types. Boldface notation $(\boldsymbol{l}, \boldsymbol{v})$ denotes a policy function (type to speech map) and plain notation $a = (l, v)$ denotes a generic speech (we use a as a generic speech when it is not necessary to distinguish between coordinates).

1. ***Optimality**: for every (λ, ω), $(\boldsymbol{l}, \boldsymbol{v})(\lambda, \omega)$ solves*

$$\max_{(l',v')\in\mathcal{A}(\omega)} (1-k)(1-(l'-\lambda)^2) + k\tilde{f}((l',v')) \tag{7}$$

2. ***Consistency**: The following holds:*

   i. For all $(l, v) \in \mathcal{A}$, $\tilde{f}((l, v)) \geq v$

   ii. Let $(\boldsymbol{l}, \boldsymbol{v})^{-1}(a) = \{(\lambda, \omega) : (\boldsymbol{l}, \boldsymbol{v})(\lambda, \omega) = a\}$. If $(\boldsymbol{l}, \boldsymbol{v})^{-1}(a) \neq \emptyset$, then

$$\tilde{f}(a) = \frac{\int\int_{(\boldsymbol{l},\boldsymbol{v})^{-1}(a)} \omega d\omega d\lambda}{\int\int_{(\boldsymbol{l},\boldsymbol{v})^{-1}(a)} d\omega d\lambda} \tag{8}$$

The consistency condition requires that sophisticated agents use the Bayes formula upon hearing speeches that are played with positive probability (point *ii.*) and that even for actions not played, they update consistently with the fact that using complexity v is hard evidence of possessing valence $\omega \geq v$ (point *i.*). Leveraging on the latter requirement we can make a first step toward a characterization of the equilibria: unless they speak effectively, politicians are 'honest', meaning that they choose the argument-complexity pair that equals their type. Consequently, voters' inference upon observing an ineffective speech is trivial and yields the conclusion that the politician has valence equal to the showcased complexity (i.e., she is exposing her valence), as formalized in Lemma 1.

Lemma 1 (**Ineffective Speeches Expose Valence**). *All ineffective speaking strategies expose the politician's valence and induce harsh update, namely*

$$(l, v) \in \mathcal{A}_{\neg ES} \Rightarrow \{(\lambda, \omega) : (\boldsymbol{l}, \boldsymbol{v})(\lambda, \omega) = (l, v)\} = (l, v) \text{ and } \tilde{f}((l, v)) = v$$

The result is intuitive: a politician that chooses an ineffective speaking strategy must be constrained by her valence, because she could stay in the same location (thus obtaining the same support from naive agents), yet report a higher valence and gain on the sophisticated.[16] By virtue of Lemma 1 it is sufficient to characterize behavior in the – one dimensional – subset of effective speeches. An equilibrium is fully characterized by the self-map $f : [0, 1] \to [0, 1]$ given by $f(v) = \tilde{f}((1-v, v))$, i.e., mapping the complexity of an effective speech to the expected valence of the politician making such speech. Denoting

[16]This unraveling argument only requires that the complexity of a speech is used as hard evidence that the candidate possesses at least that much valence, which is exactly point *i.* of our consistency requirement and is typical of models of information disclosure.

$ES(v)^{-1} = \{(\lambda, \omega) : (\boldsymbol{l}, \boldsymbol{v})(\lambda, \omega) = (1 - v, v)\}$, we write the consistency condition as

$$f(v) = \frac{\int\int_{ES(v)^{-1}} \omega \mathrm{d}\omega \mathrm{d}\lambda}{\int\int_{ES(v)^{-1}} \mathrm{d}\omega \mathrm{d}\lambda}, \quad \forall v \in [0, 1] \tag{9}$$

As for optimality, a politician (λ, ω) computes the best among feasible effective speaking strategies, i.e., she solves

$$\Pi_{ES}(\lambda, \omega) = \max_{v \leq \omega} u^{\lambda}_{ES}(v) = \max_{v \leq \omega} (1 - k)(1 - (1 - v - \lambda)^2) + kf(v) \tag{10}$$

and compares its value with that of the expose valence alternative which, by Lemma 1, is

$$\Pi_{EV}(\lambda, \omega) = \mathbb{I}_{[\omega < 1 - \lambda]} \left((1 - k) + k\omega\right) \tag{11}$$

3.2 Optimality

The first step is to find the politicians' best response given the updating function f. In doing this step, we assume that the expectation function f gives rise to well-behaved indifference *loci*. We verify such assumption after completing the consistency step.

Lemma 2 (**Optimality**). *Take an expectation function f such that both*

$$\underline{v}(\lambda) = 1 - \lambda - \sqrt{\frac{k}{1 - k}(f(\underline{v}(\lambda)) - \underline{v}(\lambda))} \tag{12}$$

and

$$\overline{v}(\lambda) = 1 - \lambda + \frac{k}{2(1 - k)} f'(\overline{v}(\lambda)) \tag{13}$$

are well-defined and strictly decreasing functions, with $\overline{v}(\lambda) \geq \underline{v}(\lambda)$. Then, the policy functions $(\boldsymbol{l}, \boldsymbol{v}) : [0, 1]^2 \to \mathcal{A}$ determine the following partition of the type space:

$$(\boldsymbol{l}, \boldsymbol{v})(\lambda, \omega) = \begin{cases} (\lambda, \omega) & \omega < \underline{v}(\lambda) & \textbf{\textit{Expose Valence}} \\ (1 - \omega, \omega) & \omega \in [\underline{v}(\lambda), \overline{v}(\lambda)] & \textbf{\textit{Match Valence}} \\ (1 - \overline{v}(\lambda), \overline{v}(\lambda)) & \omega > \overline{v}(\lambda) & \textbf{\textit{Degrade Valence}} \end{cases} \tag{14}$$

Figure III provides a graphical intuition of how $\underline{v}$ and $\overline{v}$ partition the type space according to the policy – see Definition 1 for the taxonomy – that politicians adopt. Always following Figure III, we fix the naive-preferred argument to $\lambda = \frac{1}{2}$ and see how the candidate changes her speech as her valence ω increases. Candidates with low valence $\omega < \underline{v}(\lambda) \leq 1 - \lambda$ (type A in Figure III) maximize the turnout of naive voters by choosing their preferred argument λ and defend it in the best way they can, namely with complexity ω. Because of their low valence, such politicians have no feasible effective speaking strategies that do better than (11).

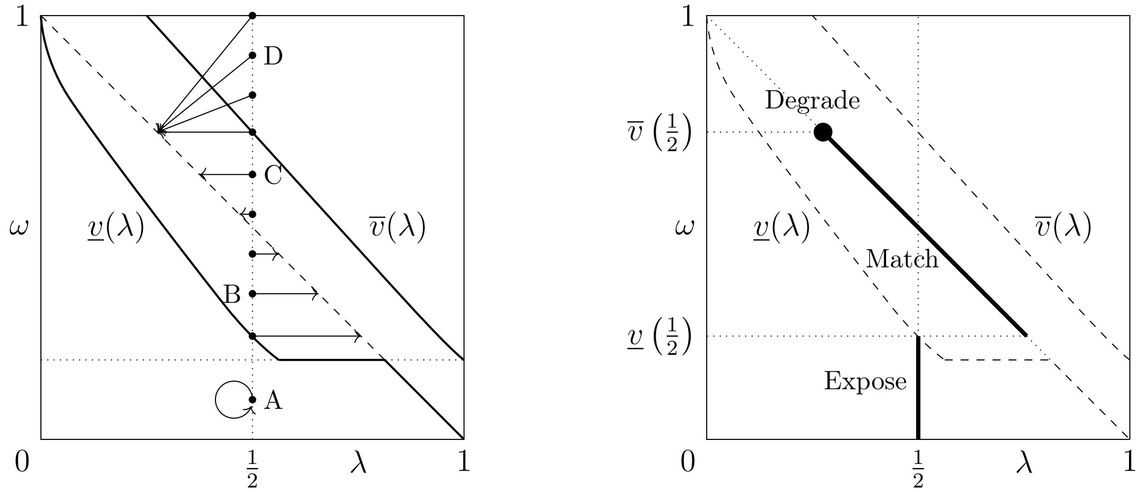


Figure III: Optimal speech for politician with $\lambda = \frac{1}{2}$, varying valence.

For intermediate values of valence $\omega \in [\underline{v}(\lambda), \overline{v}(\lambda)]$ politicians depart from the naive-preferred argument λ and pursue a valence matching policy, meaning that they pick the most complex argument their valence allows them to effectively defend. Despite displaying the same degree of valence as under the optimal expose valence strategy, speaking effectively allows politicians to gain the updating premium $f(v) - v$. The threshold $\underline{v}(\lambda)$ represents the valence for which a λ-politician is indifferent between exposing and matching her valence. We indeed obtain (12) from the indifference conditions $\Pi_{EV}(\underline{v}(\lambda), \lambda) = u^{\lambda}_{ES}(1 - \underline{v}(\lambda))$.[17] The switch from a "pure naive targeting" to a "pure sophisticated targeting" policy induces a non-monotonicity of displayed complexity as a function of actual valence: as ω crosses the threshold $\underline{v}(\lambda)$, the argument chosen becomes (discontinuously) more simplistic and then decreases along the effective speaking bound. There is a significant difference on the reason for (and the implications of) a match valence strategy among politicians that choose it, based on whether they lie above or below the effective speaking bound. Those with $\omega < 1 - \lambda$ (type B in Figure III) escape the harsh update following an expose valence strategy by choosing the best – though still more simplistic than the naive-preferred λ – argument they can effectively defend: they can't satisfy their audience, so they pretend they face a less demanding one. Conversely, politicians with $\omega \in [1 - \lambda, \overline{v}(\lambda)]$, (type C in Figure III) do not have an expose valence alternative because the bound prevents them from displaying their valence in full at $l = \lambda$: they are "too good" for their naive audience and choose to sacrifice some naive support

[17] To be precise, politicians located along the flat part of $\underline{v}(\lambda)$ that can be seen in Figure III are not indifferent, but strictly prefer effective speaking.

to maximally exploit their ability to mobilize sophisticated voters. As valence grows, the policy of choosing the argument based on one's own valence becomes more and more costly in terms of forgone support from the naive audience. Therefore, as ω crosses $\overline{v}(\lambda)$, ignoring the bliss point λ is no longer optimal and politicians strike a compromise between mobilizing sophisticated voters through the valence update and naive ones through the type of argument chosen. $\overline{v}(\lambda)$ represents a bound on the degree of complexity politicians are willing to display given λ. If interior,[18] $\overline{v}(\lambda)$ solves the first-order condition of the effective speaking problem (10), resulting in condition (13). All politicians with $\omega > \overline{v}(\lambda)$ (type D in Figure III) degrade their valence by speaking effectively at an argument whose complexity is strictly below their valence.

3.3 Consistency and Equilibrium

We now need to characterize the valence update induced by Lemma 2. To this end, think of the inference problem faced by a sophisticated agent upon observing the effective speech $ES(v)$. In light of Lemma 2, there are two rationalizations for such behavior, portrayed by Figure IV. The first rationalization, captured by the horizontal segment, is that the politician has valence type exactly equal to v, and that the preference type λ is such that pursuing the best possible valence update is optimal.[19] The second rationalization, captured by the vertical segment, is that the politician has preference type λ such that $v = \overline{v}(\lambda)$ and high valence that she is optimally degrading by speaking at v. Such rationalizations imply radically different updates: in the former scenario, the valence of the politician is equal to v, hence the update corresponds to the harshest possible one given effective speaking. This is the same outcome as in the benchmark case of $k = 1$ (Section 2.3). The second rationalization leads instead to a valence update of $\frac{1+v}{2}$, that is, a benevolent update driven by the valence degradation of high-valence politicians. In this case, the update corresponds to that in the $k = 0$ benchmark (Section 2.2).

As a result, the consistency condition (9) is equivalent to a point-wise characterization of f as a convex combination of the expectations in the polar cases $k \in \{0, 1\}$.

Lemma 3 (Poe's Law). *Given a partition according to Lemma 2, the consistency con-*

[18] That is, which is obtained assuming that the valence constraint in (10) is not binding. In general, the best effective speech is

$$\Pi_{ES}(\lambda, \omega) = \begin{cases} u^{\lambda}_{ES}(\omega) & \omega \leq \overline{v}(\lambda) \\ u^{\lambda}_{ES}(\overline{v}(\lambda)) & \omega > \overline{v}(\lambda) \end{cases}$$

[19] This requires that $\lambda \in [\underline{v}^{-1}(v), \overline{v}^{-1}(v)]$, i.e., that it is neither too low – otherwise a politician of valence v would expose valence – neither too high – otherwise a politician with valence v would degrade it.

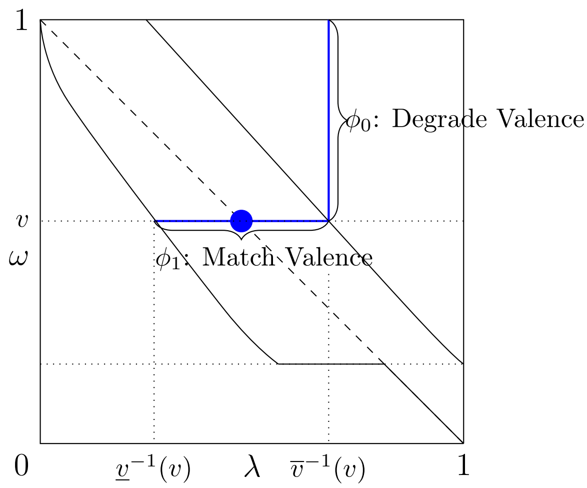


Figure IV: **Poe's Law:** Candidate's types who choose $ES(v)$

dition (9) *becomes*

$$f(v) = v\frac{\phi_1(v)}{\phi_0(v) + \phi_1(v)} + \mathbb{E}[\omega|\omega > v]\frac{\phi_0(v)}{\phi_0(v) + \phi_1(v)}, \tag{15}$$

where $\phi_1(v) = \overline{v}^{-1}(v) - \underline{v}^{-1}(v)$ *and* $\phi_0(v) = \mathbb{I}_{[v \geq \overline{v}(1)]}(1 - v)$.

Expression (15) is the representation of Poe's Law in our setting: following the original statement of the law, the "genuine creationist statement" is in our model the politician speaking effectively to the best of her capabilities (ϕ_1 types), whereas the "parody of a creationist statement" is a politician degrading her valence to mobilize naive voters who require a simplistic argument (ϕ_0 types).

The latter rationalization need not be part of the equilibrium, as $\phi_0 > 0$ only if $v \geq \overline{v}(1)$. The lowest v that admits this rationalization is the one pursued by valent politicians that face naive voters with the highest possible bliss point $\lambda = 1$. For this reason $\overline{v}(1) =: v^*$ which, by (13), is given by

$$f'(v^*) = \frac{2(1-k)}{k}v^*. \tag{16}$$

The threshold v^* determines a qualitative change in the equilibrium expectation function f. Below v^*, (15) gives that $f(v) = v$, that is the same harsh update of the valence-exposing strategies. Above v^*, instead, $\phi_0 > 0$ and therefore we have a more benevolent update $f(v) > v$: the fact that the politician might be degrading her valence results in an updating premium. We can plug the analytic expressions for ϕ_0 and ϕ_1 obtained inverting the $\overline{v}, \underline{v}$ functions (12)-(13) into (15) to arrive at the following differential equation:

$$f'(v) = 2\frac{1-k}{k}\left[\frac{(1-v)\left(\frac{1+v}{2} - f(v)\right)}{f(v)-v} - \sqrt{\frac{k}{1-k}(f(v)-v)}\right] \tag{17}$$

Now, fixing $v^* \in [0,1]$, the pair (16)-(17) constitutes an initial value problem. Let $\hat{f}_{v^*,k} : [v^*,1] \to [v^*,1]$ be the unique solution to this problem.

Proposition 1. *For $k \in [0,1]$, the function*

$$f(v) = \begin{cases} v & v < v^* \\ \hat{f}_{v^*,k}(v) & v \geq v^* \end{cases} \tag{18}$$

constitutes an equilibrium if and only if $v^* \in V_k$, an interval with the following properties:

i. $V_k > \{0\}$ for $k > 0$,

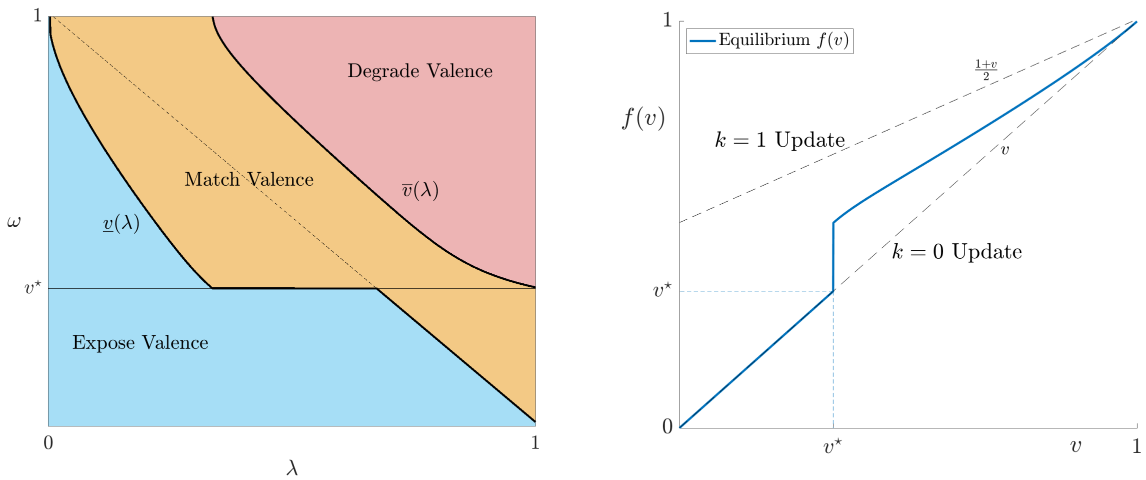
ii. V_k is strictly increasing (in the set-inclusion order) for $k < \frac{2}{3}$, and

iii. $V_k = \{1\}$ $\forall k \geq 2/3$.

For a typical equilibrium, Figure Va and Figure Vb represent, respectively, the communication policy that each type of candidate pursues and the valence update conditional on effective speaking. Following the construction that leads to Proposition 1, effective speeches of complexity above v^* attract both candidates who are 'too valent for their crowd' and candidates who are 'not valent enough for their crowd'. For the set of effective speeches that admit both rationalizations, i.e. for which Poe's Law is non-trivial ($\phi_0 > 0$ in Lemma 3), the differential equation (17) guarantees that these two forms of attraction are correctly accounted for in the equilibrium expectations. What remains to be determined is when the construction implicit in (17) is valid, namely what $v^\star$ are legitimate lower bounds for the complexity of speeches that can be used as a valence degrading strategy. By Proposition 1, there is an interval $V_k = [\underline{V}_k, \overline{V}_k]$ that contains the values of v^* for which (18) constitutes an equilibrium. Both $\underline{V}_k, \overline{V}_k$ are increasing in k, $\overline{V}_0 = \underline{V}_0 = 0$ and $\overline{V}_k = \underline{V}_k = 1$ for all $k \geq \frac{2}{3}$. We now give an intuition behind the determination of the functions $\underline{V}_k$ and $\overline{V}_k$, formalized and discussed further in the Appendix.

The condition $v^* < \overline{V}_k$ derives from the requirement that types in the triangle above the bound but below v^* (refer to Figure Va) follow their equilibrium prescription and, even without an updating premium, choose valence matching. If this condition failed, then some politician would use the argument as part of a valence-degrading strategy and updating $f(v) = v$ would become 'incoherently' harsh. The condition $v^* > \underline{V}_k$ derives instead from the requirement that the first-order condition (16) characterizes a maximum, and not a minimum of $u^1_{ES}(v)$. If this condition failed, then not even candidates facing $\lambda = 1$ would use $v^\star$ as part of a valence degrading strategy (preferring more complex ones) and any non-trivial Poe's Law would entail 'incoherently' benevolent updating.

The bounds on the feasible v^* also constitute the equilibria that emerge from the fixed point procedure (guess f, compute $\overline{v}, \underline{v}$ through Lemma 2, update f through Lemma 3) initialized at the polar equilibrium updates. Initializing at $f_0(v) = v$, we converge to the equilibrium with the lowest possible threshold $v^* = \underline{V}_k$, while initializing at $f_0(v) = \frac{v+1}{2}$ we converge to the equilibrium with the highest possible threshold $v^* = \overline{V}_k$. This suggests that we can justify equilibrium selection through an appropriately initialized learning dynamics, e.g. sophisticated voters learning the "mistakes from dismissal" (see Section 4). We do not embrace any such selection and instead characterize the full set of Bayesian equilibria.



(a) Bounds $\underline{v}(\lambda)$, $\overline{v}(\lambda)$ and strategy choice in (λ, ω) space, for $k = 0.4$, $v^* = \overline{V}_{0.4}$.

(b) Expectations conditional on effective speaking $f(v)$, for $k = 0.4$ and $v^* = \overline{V}_{0.4}$.

Figure V: Equilibrium policy characterization (left) and consistent expectations (right)

The properties of V_k are important to shed light on the drivers of extreme simplism.[20] In a given equilibrium, v^* represents the lower bound to the simplicity of the argument that a politician can make while still pretending to be constrained by the simple-mindedness of naive agents. All arguments whose complexity falls short of v^* are defended effectively only by politicians who, even without an updating premium, give up some naive support to provide the 'hard evidence' that they possess valence v. For $k \in (0, \frac{2}{3})$, V_k is a strict

[20] Our model with continuous action space is particularly well-suited for addressing this question. A version of the model with small (binary) arguments set would most likely replicate our results on average simplism but would rule out, by construction, any variation on the type of arguments that are used strategically in the Bayesian (and dismissive) equilibrium. It is only in a model where this adjustment is active that we can point to dismissal as a valid rationalization for extreme arguments and the properties of V_k in equilibrium are a key step making this claim.

subset of $[0, 1]$, implying that in any (Bayesian) equilibrium *some argument must, but not all arguments can, be used strategically.* In particular, the existence of a lower bound implies that there are limits to the ability of politicians to strategically employ simplistic arguments. In the words of the Poe's Law, extreme arguments are never used as a parody. Moreover, the set of arguments that can be used as a parody (i.e. those whose bound exceed v^*) shrinks with the share of sophisticated voters. When $k \geq \frac{2}{3}$ the equilibrium collapses to the fully sophisticated benchmark (cases with $k < 1$ justify the selection criterion): when mobilizing sophisticated voters becomes the preponderant objective of politicians, it is impossible for any argument to be used as a parody and, effectively, there is no Poe's Law.

3.4 Simplism Comparative Statics

We now discuss how the simplism metrics (see Definition 2) vary across equilibria. In Definition 4, we account explicitly for the multiplicity of equilibria (one for each feasible v^*) for $k < \frac{2}{3}$.

Definition 4. *For $k \in [0, 1]$ and $v^* \in V_k$, let $\bar{s}^{k,v^*}$ be the average simplism computed – according to Definition 2 – under the (k, v^*)−equilibrium policies. Let also the correspondence $\bar{s} : [0, 1] \to \mathbb{R}$ be given by $\bar{s}(k) = \bigcup_{v^* \in V_k} \bar{s}^{k,v^*}$.*

In words, $\bar{s}(k)$ associates k with all the possible levels of *aggregate* simplism that can result in an equilibrium when the share of sophisticated voters is k. Recall that $\bar{s}(0) = 0$ and that for $k \geq \frac{2}{3}$ there is a unique equilibrium that collapses to the fully sophisticated benchmark $k = 1$. Therefore, $\bar{s}(k) = -\frac{1}{6}$ for all $k \geq \frac{2}{3}$. For $k \in (0, 2/3)$, $\bar{s}(k)$ is a proper set, as depicted in Figure VIa, with the following properties.

Proposition 2. *The following properties hold:*

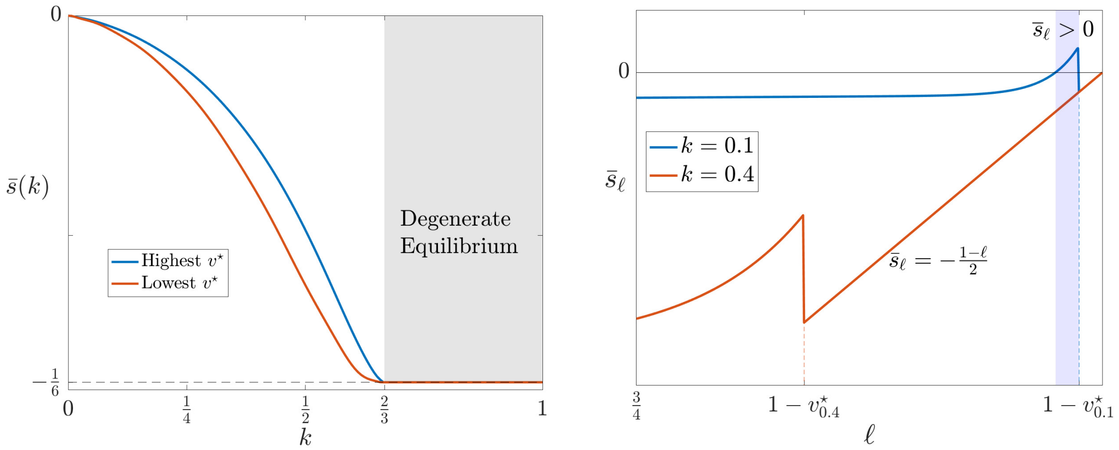
i) ***Average Complexification*** *$\bar{s}(k) < 0$ for all $k > 0$.*

ii) ***Possible Catalysts*** *For low k, $\bar{s}_l^{k,\underline{V}_k} > 0$ for l in a left neighborhood of $1 - \underline{V}_k$.*

Moreover, by numerically solving the model we can formalize the idea that, in equilibrium, the share of sophisticated voters reduces simplism.

Fact 1. *$\bar{s}(k)$ is decreasing in k in set order.*

Point $i)$ of Proposition 2 establishes that in *all* equilibria aggregate simplism falls short of the fully naive benchmark: in equilibrium, candidates use on average arguments that are more complicated than warranted by the preferences of naive voters. Fact 1 considers the comparative statics, adding that the larger the share of sophisticated voters k, the



(a) Average simplism $\bar{s}(k)$, for the highest and lowest possible values of v^*.

(b) Argument-wise simplism and catalyst arguments as per Prop 2-*ii*).

Figure VI: Average and Argument-Wise simplism in BNE

more complex political communication is on average (meaning also complexity strictly exceeds naive preferences when $k > 0$).

At first sight, these results are natural: as complex arguments signal valence, we would expect politicians that care about their perceived valence (the larger k is, the more important perceived valence becomes) to exploit this by using more complex arguments on average. However, the benefits from using effective speeches might be abused by politicians with valence that falls short of the bound implied by the naive-preferred argument, which indeed happens in equilibrium: if valence signaling concerns were absent, all politicians with $\underline{v}(\lambda) \leq \omega \leq b(\lambda)$ would choose a less simplistic argument.

The presence of politicians who 'simplify their speech' in order to be perceived as more competent thanks to the pooling with valent types facing a very simplistic audience (in the language of Figure III, type B politicians) is an interesting phenomenon that can shed new light on the behavior of demagogues, a good case in point being Donald Trump. In response to the simplism of his argumentations, it is often maintained that Trump is not really that ignorant, but simply pretends to be so in order to pander to his ignorant audience. In the words of Kraus (1990): "The secret of the demagogue is to appear as dumb as his audience, so that these people can believe themselves as smart as he". This resembles what politicians who degrade their valence (type C in Figure III) do, but our model presents an intriguing alternative. A demagogue using simplistic arguments may really be ignorant, and especially he may be speaking in a way that is even more ignorant than his ignorant audience, so as to appear as a smart politician appeasing an ignorant audience. This is possible because the politician can, to an extent, manipulate sophisticated voters, making them believe that the naive audience is more ignorant than it really is. Therefore, paraphrasing the above-mentioned words of Karl Kraus, our model suggests that the secret of the demagogue is to appear as dumb as you think his audience is, so that you believe he is as smart as you are.[21]

This force pushes $\bar{s}$ upwards, but it is never strong enough to offset the argument complexification (i.e., choosing an argument l lower than λ to display higher valence) pursued by politicians above the bound. Despite never affecting the negative sign of aggregate simplism $\bar{s}$, Point *ii*) establishes that the abuse of simplistic arguments by politicians below the bound $b(l)$ does not only lead to the *type-wise* failure of argument complexification, but it is severe enough to also cause the failure of *argument-wise* complexification: that is, $\bar{s}_l$ can be positive. In other words, it is possible that the effective defense of a relatively extreme argument is pursued by politicians who on average face a naive-preferred argument of greater complexity. This effectively means that even in a rational equilibrium, some simplistic arguments might be catalysts for further simplism. This occurrence is displayed in Figure VIb.

Summary: Simplism in BNE So far, we have developed an equilibrium model where some politicians choose simplistic arguments to exploit Poe's Law and appear more competent than they are. However, in the aggregate the BNE model cannot explain any excess of simplism in political debates. First, extremely low quality arguments (below $v^{\star}$) are effectively defended only by low-quality politicians facing a demand for simplistic arguments, leading to a failure of the Poe's Law and to harsh updating from sophisticated voters. Second, despite the fact that some arguments might be catalysts of simplism, the average complexity of the political discourse is improved by the presence of sophisticated voters in the audience. Third, increasing the share of sophisticated agents (as a consequence, for example, of a growing awareness of the purely rhetorical nature of campaign messages) pushes towards an increase in the average complexity of the political debate. In the next section we show that these quantitative results are driven by assumptions, implicit in the BNE model, that are somehow at odds with empirical evidence. Therefore, the insights we just obtained should be applied only after assessing, in the setting under consideration, the behavioral plausibility of equilibrium analysis, in particular of equilibrium awareness among sophisticated voters.

[21] We thank an anonymous referee for pointing us to the quote of Karl Kraus and for suggesting the effective paraphrasis which we report here.

4 Dismissive Update

Motivation It often happens that the relationship of a candidate with one specific group catalyzes the bulk of the media and public attention. A case in point is Donald Trump and the working class: using the words of Carnes and Lupu (2021), "many journalists have embraced the idea that Trump uniquely appealed to white working class voters". The spread of this narrative, the authors suggest in their article, is related to the content of Trump's speeches and texts, which have been often interpreted as targeting and mobilizing white working class voters (Lamont et al., 2017). However, Carnes and Lupu (2021) show that Trump's support had a much broader outreach than the working class. In such a context it is natural to conjecture that many non working-class voters considering supporting Trump must have also been subject to the same narrative (for example by following politics through the mainstream media). Therefore, we can expect that at least some of these voters believed Donald Trump to be almost exclusively targeting working class voters in his campaign speeches and rallies, and would evaluate the candidate's campaign accordingly. Specifically, believing the Trump-targeting-working-class narrative will lead voters to solve Poe's Law in the following, stereotypical way: If the candidate is only targeting naive voters (i.e., the working class) and he knows "what they want to hear", then an effective speech of complexity v can only inform that his valence is weakly above v. This line of reasoning seems very appealing, but might manifest a misperception of the surrounding environment: an observer that rationalizes any political speech as "this is what the voters want" infers valence as an econometrician of the game with fully-naive voters. In our setting with positive k, this inference is incorrect as it fails to realize that there is a positive mass of observers-voters who also care about the candidate's valence. However, it is well known that individuals misperceive and tend to overestimate their cognitive abilities relaive to those of their peers (for psychological evidence and the impact on various economic outcomes see resp., e.g., Moore and Healy (2008) and Malmendier and Taylor (2015)). Dismissal poses a significant threat to the validity of the results of Section 3: if sophisticated voters do not recognize that simplistic arguments help candidates masquerade poor valence (since no one but them cares about valence), they won't adjust their update, which will potentially foster the abuse of such arguments. In this section, we show that this possibility is indeed a real concern. First, we formalize our dismissal assumption and relate it to the literatures on overconfidence and level-k thinking. We then describe the outcomes of political communication in the presence of dismissive updating, contrasting with those of the rational model of Section 3. As it closely follows the one for the BNE, most of the formal development and discussion of the model with dismissal is relegated to Appendix B.

4.1 Dismissal: Definition and Discussion

Dismissive voters solve the Poe's Law in a stereotypical way: after hearing an effective speech of complexity v, they think that the candidate has expected valence

$$f^{dis}(v) = \mathbb{E}[\omega | \omega > v] = \frac{1+v}{2}. \tag{19}$$

Notice that the updating rule (19) is the correct updating rule in the fully-naive benchmark (Section 2.2). Indeed, following (19) manifests a bias that resembles the first stage of level-k thinking (Nagel (1995)): a dismissive voter indeed updates as if other citizens were instinctively responding to something – the rhetorical argument – that he (and only he) knows is payoff-irrelevant; thinking that he is surrounded by a naive crowd, he effectively underestimates the cognitive abilities of his peers. A feature that is wired in level-k models, and shared by cognitive hierarchy models (Camerer et al. (2004), which allows for heterogeneity in the beliefs on other players' levels), is that players never assign positive probability to other players' having the same cognitive level as themselves.[22] Dismissive agents in our model partially share this feature since they believe they are conceptualizing a more complex game than *any other citizen*; however, when it comes to the only player they have strategic uncertainty about, i.e., the candidate, dismissive agents place her in the same cognitive class as themselves, recognizing she also has private information (about her own valence and their peers' preferences). Because of this mismatch, our model is does not directly fit into the literature on level-k and cognitive hierarchy, though the behavioral evidence supporting level-1 thinking also explains the attitude of our dismissive agents towards their peers.[23] An alternative, and to some degree complementary, justification for (19) is that dismissive voters identify with their leader against a naive crowd of unknown bliss point, not even conceiving that she might be constrained by low valence.

Whatever the fundamental source of dismissal, voters that follow (19) use a misspecified model —the trivial game with $k = 0$— to guide their inference. In this sense, dismissal is a retrospective bias since voters reason about *realized* information in a systematically incorrect way.[24] As the real share k of sophisticated-but-dismissive voters grows away from 0, the misspecification becomes more and more severe, approaching a limit where

[22] A notable exception is the inclusive cognitive hierarchy model of Koriyama and Ozkes (2021) where players can conceptualize opponents with the same cognitive level. They argue that such inclusiveness substantially improves the explanatory power of the models of hierarchical thinking.

[23] See Crawford et al. (2013) for a review of applications of level-k and CH models. Breitmoser (2012) assesses their fit, relative to standard equilibrium models, for explaining actual behavior in certain games (concluding they perform better in auction settings as Crawford and Iriberri (2007) and coordination games, but worse in common interest and collective decision making models, e.g. Battaglini et al. (2010)).

[24] Benjamin (2019) surveys work that documents and models retrospective biases in updating, while Levy et al. (2022) provide a specific application of such biases to populism.

everyone is sophisticated but thinks that everyone around them is naive.[25] The analysis local to $k = 0$ shows that even a minimal misperception yields predictions that are opposite to those of BNE, both in terms of average simplism and in the interpretation of extreme arguments. As for the analysis with substantial misperceptions, we do not directly address the plausibility of the assumptions behind dismissive updating relative to those implicit in the BNE (and how such plausibility varies with the real k). Instead, we think of the results of this section as contributing to a vast literature that studies how overestimating one's own cognitive abilities can impact on various economic contexts. In particular, the contrast between Propositions 2 and 3 exemplifies how such biases might affect the type of rhetorical arguments employed in the political debate.

4.2 Analysis of the Game with Dismissal

Contrary to the BNE model, the expectation f is no longer endgenous, but rather pinned down by our behavioral assumption. As (19) is a strictly monotonic function, we can still partition the politician's type space according to the optimal policies presented in Lemma 2, though expressions (12) and (13) now deliver closed form characterization of the candidate's policy (see Proposition 4 in Appendix B). Given the candidate's strategy, we readily obtain the properties of the equilibrium under dismissal.

Proposition 3. *Under dismissive updating,*

i) $\bar{s}(k) \geq 0$ *for all* k*, strictly if and only if* $k \in \left(0, \frac{4}{5}\right)$.

ii) $\bar{s}(k)$ *is non-monotonic, decreasing for* $k > \frac{2}{3}$.

iii) For all k*,* $\bar{s}_l(k)$ *is increasing in* l.

iv) Let $v^\star(k) = \min\left\{\frac{k}{4(1-k)}, 1\right\}$*, increasing in* k*. Then,*

$$\bar{s}_l(k) > 0 \text{ if and only if } l > \max\left\{\frac{1}{2}, 1 - v^\star(k)\right\}.$$

We now discuss the properties of dismissive equilibria listed in Proposition 3, contrasting them with those of the Bayesian equilibria. Table VII summarizes this comparison, while Figure VIII focuses on our main measure, average simplism, showing how it changes with the share of sophisticated voters in the dismissive (blue line) relative to the rational (yellow line) equilibria.

[25] Recall that we call sophisticated the voters who use the political discourse to infer the candidate's valence rather than being swayed by her rhetorical argument. Clearly, under dismissal, sophisticated voters make wrong inference. We maintain this labeling for consistency with the model setup, and postpone to Section 5 a discussion of the electorate that sophisticated-but-dismissive voters might represent.

	BNE	Dismissal
Avg. Simplism: Level	Below naive benchmark	Above naive benchmark
Avg. Simplism: $k-$ Compstat	Monotone ↓	First ↑, then ↓ (to 0)
Extreme arguments ...	Repel simplism	Catalyze simplism

Figure VII: Comparison between BNE and Dismissal game

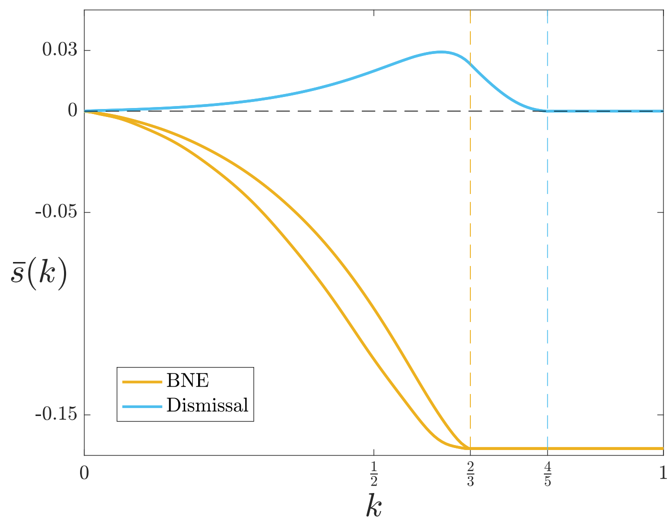


Figure VIII: Average simplism $\bar{s}(k)$ in the Bayes Nash vs dismissal game.

Average Simplism. In direct contrast with the Bayesian equilibrium, in which political communication is always a repellent of simplism ($\bar{s} \leq 0$), communication to dismissive agents catalyzes simplism ($\bar{s} \geq 0$, point i)). Hence, when compared with the fully naive benchmark, the presence of agents rewarding valence increases the complexity of the political debate only if there is common knowledge thereof. Dismissive voters are a vector of an increasingly simplistic political communication because simplistic arguments no longer exclusively target the instinctive response of naive voters, but also their (equally instinctive) updating. By point ii), increasing the share of sophisticated but dismissive voters has an ambiguous impact on the average level of simplism. Recall that in the BNE, the net simplism of political communication monotonically decreases with k as the updating premium scales down with the candidates' incentive to exploit it. Under dismissal the updating premium is instead fixed, which induces the abuse of valence matching strategies at simplistic arguments, leading to an initial increase in average simplism. To understand why $\bar{s}$ eventually decreases, it is useful to think about how large k equilibria become degenerate. Under dismissal, $\underline{v}$ is "absorbed" by 0, as politicians below the bound exploit the unresponsiveness of the updating premium by abusing valence matching strategies.[26]

[26]On the contrary, the Bayesian equilibrium becomes degenerate (for $k \geq \frac{2}{3}$) because $\underline{v}$ is "absorbed" by the bound $b(l)$: all politicians below the bound expose valence because no updating premium can be

At $k = \frac{2}{3}$, we have $\underline{v}(0) = 0$ but $v^\star < 1$, which means that all politicians below the bound, but not all politicians above it, choose valence matching. Therefore, not all the simplism contributed by politicians below the bound is offset by those above, which makes it immediate to conclude that $\bar{s}\left(\frac{2}{3}\right) > 0$. As k grows past $\frac{2}{3}$, the $\overline{v}(\cdot)$ function shifts up while $\underline{v}(\cdot)$ remains degenerate, which means that politicians above the bound offset an increasingly large share of simplism, causing $\bar{s}(k)$ to decrease. This process ends when $v^\star = 1$, i.e. at $k = \frac{4}{5}$, as all politicians (both above the below bound) choose an argument as complex as they can effectively defend. Beyond that threshold, the behavior of the candidate does not change and the equilibrium remains degenerate with all types choosing a matching-valence strategy (i.e. moving horizontally to the bound). As a result, in the game with dismissal the relationship between the share of sophisticated but dismissive agents and simplism is hump-shaped.

Argument-wise Simplism. Points *iii*) and *iv*) establish that, under dismissal, simplistic arguments become catalyst of further simplism and drive the aggregate results. By *iii*), the simpler an argument is, the more it catalyzes simplism: the average simplification discussed above is driven by the simplification catalyzed by extreme arguments (see Figure XIIb in Appendix B). This contrasts with the BNE model, where $\bar{s}_l$ had a downward jump, but is a natural result once we recognize that dismissive updating is particularly appealing for extreme arguments. Since the candidate gets a pass even there, those arguments become an excellent rhetoric shelter behind which to hide poor valence. Point *iv*) states that $\bar{s}_l(k)$ is positive if and only if no politician is using l as part of a valence-degrading strategy: under dismissal, an argument is defended effectively both by candidates that choose valence matching and by candidates that choose valence-degrading if and only if it repels simplism.[27] Arguments that catalyze simplism, instead, are only used by low valence politicians that do valence matching. Therefore a rational observer would face a trivial Poe's Law (the argument is genuine and the candidate is poor) upon hearing such arguments. Recall that in the BNE all arguments that induced a trivial Poe's law, i.e., those with $l > 1 - v^\star$, were necessarily *repellents* of simplism, while no definitive conclusion could be drawn about those that admitted a non-trivial one, see Proposition 2-*iii*). Clearly, these differences originate from the fact that the inference of a rational observer is inconsequential for the outcome of the dismissive game, while in BNE such observer coincides with a sophisticated voter, whose beliefs feed back in the equilibrium. Hence, the results on argument-wise simplism formalize the intuition that what drives the differences with the Bayesian model is that the updating premium is hard-wired in the minds of the dismissive voters, even for arguments where no rational premium can exist.

sustained in equilibrium.

[27]The *only if* direction requires that $k < \frac{2}{3}$ so the bound $\frac{1}{2}$ is not binding.

5 Alternative Interpretations of the Model

Simplism as Populism: An alternative interpretation of our model is in terms of extremism and/or populism as opposed to technocratic governance. Naive voters are those who are subject to the populist rhetoric, whereas sophisticated voters are *disenchanted*, i.e., they have a prejudice against all sorts of political debate and think that all that matters is the candidate's competence (technocratic view). This trait seems common especially in Western European democracies, where the emergence of populist movements was paralleled by a decline in the attachment of voters to ideological positions.[28]

Recall that the decreasing bound $b(l)$ implies that the lower its location l, the larger is the ability required to successfully defend a given argument: when interpreting the model in terms of extremism, this means that it takes much more valence to make a convincing moderate argument than it takes to repeat populist slogans. This seems a compelling characterization of extremism and populism: for example, Levy et al. (2022) define populism precisely as a political strategy offering a simplistic view of the world, ignoring some of the relevant dimensions of a given problem.

In addition to being simplistic, populist rhetoric crucially hinges on the narrative of leaders representing "the people" as opposed to the elite. Our model suggests that such narrative can be used to in fact mobilize the elite itself, especially when their members fail to realize that they are the targets of political propaganda, as in our dismissal setup. According to this view, the simplism of populist rhetoric need not be the manifestation of an intrinsic preference of voters for simplistic arguments, but rather a product of the interaction between the effort of mediocre politicians to masquerade their limited valence and the social mistrust and overconfidence driving the dismissive behavior of sophisticated voters.

The mechanism we just described would be even stronger under the realistic scenario in which the disenchanted proponents of technocratic governance are also the most moderate in the electorate, which would mechanically increase the bliss point of the (residual) naive voters. Once moderate naive agents become immune to political rhetoric, political communication loses a moderating anchor and might be pushed even further to the extreme, especially if the disenchantment (i.e., transition to sophistication) of moderate voters goes hand in hand with the belief that everyone else remained naive (with their preferences becoming more extreme).

[28] Although disenchantment aligns with our assumption that the argument employed is the payoff irrelevant, our model is flexible to account for situations where either the social planner or sophisticated voters are harmed by populist arguments. In the former case, the positive analysis goes through unchanged, and only $\bar{s}$ becomes a proper welfare variable. In the latter case we would need to augment the baseline model with a direct cost of populism based on the aversion of sophisticated voters for populist arguments, along the lines suggested for independent voters in the next paragraph (see fn. 29).

Partisan Base vs Independents: For whom the Politician Whistles? Our model is also open to the interpretation of mobilizing an electorate composed both of a partisan base and a portion of independent voters.[29] Partisans (the naive voters in the baseline model) only decide whether to support the candidate based on how close the policy proposed is to their bliss point. Independents (sophisticated voters in the baseline) do not intrinsically care about the policy but only about the valence of the politician. This assumption is not unreasonable because political campaigns often focus on a small set of partisan issues, say 'immigration' or 'civil rights'. Therefore, it is likely that a fraction of voters has very little interest in the campaign issue per se, but follows the political debate to understand how competent the politician would be in dealing with 'the real issues' (say 'the economy' or 'foreign policy').

Our results offer a novel interpretation of extremist rhetoric. We often think of extremist talk as "dog-whistling" to an extreme partisan base, exploiting their instinctive taste for divisive talk. In our model, instead, politicians use extremist arguments to whistle to the independents, exploiting the fact that they instinctively interpret political speeches as an attempt to mobilize the partisan crowd. Because they rationalize extreme speech as "this is what her base wants", dismissive independents perceive the partisans to be more extreme than they actually are. This logical trap might also explain why political platforms that are perceived to target more extreme (simplistic) partisan voters have an easier time getting a "pass" also from independent voters. In this view, our comparison of the Bayesian and dismissive equilibria indicates that the perception of extremism of the partisan base helps the politician masquerade poor valence, and fosters further extremism. In a dynamic setting, the success of simple-minded policies might set up a vicious cycle.

6 Conclusion

We have built a model of mobilization through rhetoric, rationalizing the emergence of simplistic arguments in political campaigns. A politician that talks to a cognitively heterogeneous audience must strike a balance between choosing an argument that resonates with the naive electorate and signaling her valence to sophisticated voters. The latter are faced with an interpretation dilemma that we call Poe's Law: when observing the effective defense of a rhetorical argument, they must ask whether this is the best argument the politician can defend or whether she is simplifying her speech in order to appeal to naive

[29] A slight modification of the baseline model is pheraps needed, as in this framework it is more natural to assume that independent voters have the most moderate bliss point $\lambda = 0$. In this parametrization k affects the mobilization part of the politician's objective directly, which is maximized at $l = \lambda(1 - k)$: Increasing k mechanically pushes towards more moderate policies, since it reduces the population's bliss points. Maintaining payoff-irrelevance even for independent voters, the baseline model isolates the communication-driven effects of voters' sophistication.

voters' gut feelings.

In equilibrium, some low-valence politicians exploit this ambiguity and appear more competent by choosing an argument that is more simplistic than the naive-preferred one. In spite of the fact that sophisticated voters are interested in valence, which should motivate politicians to employ complex arguments, their presence can fuel the use of simplistic arguments. Therefore, we uncover a novel explanation of demagoguery which, paraphrasing the words of Kraus (1990), can be described as "a politician appearing as dumb as you think their audience is, so that you will believe he is as smart as you are".[30] Our findings in the Bayesian setting go beyond pointing at this possibility: First, we show that the potential for abusing simplistic arguments to gain an updating premium is limited and a candidate that defends overly simplistic arguments outs herself as incompetent. Second, we show that political communication is on average less simplistic than the average naive-preferred argument, and that the greater the share of sophisticated voters, the more complex political communication becomes. Therefore, although valence signaling in equilibrium leads *some* candidates to simplify their arguments, the result is reversed *on average* and valence signaling leads to a complexification of the political discourse.

We have also addressed the robustness of our results to relaxing equilibrium assumptions that are challenged by the evidence that individuals tend to be overconfident in their cognitive abilities. Specifically, we have examined a variant of the model wherein voters interested in inferring valence overlook the potential for rhetorical manipulation aimed at themselves, because they think that others are naive.

Dismissive updating fundamentally alters how voters rewarding valence impact the simplicity of political discourse. Absent the discipline of equilibrium consistency, political communication tends indeed to be *more* simplistic than what would be preferred by naive voters. This occurs because low valence candidates target dismissive voters, who readily accept all arguments as long as they are rhetorically effective. Notably, it is the prevalence of overly-simplistic arguments that drives the disparity in aggregate outcomes between Bayesian and dismissive equilibria: in the former, those arguments are not abused as they out the candidate as incompetent; in the latter, they attract incompetent candidates that shelter behind dismissive updating.

In essence, our analysis shows that "politicians saying stupid things to mobilize their base" is a plausible equilibrium outcome. However, it also highlights that applying such justification acritically could inadvertently encourage the abuse of simplistic arguments, providing a potential explanation of the tendency towards rhetorical simplification observed in modern political discourse.

[30] The original quote, which we report in Section 3, goes: "The secret of the demagogue is to appear as dumb as his audience so that these people can believe themselves as smart as he".

References

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin, "A political theory of populism," *The Quarterly Journal of Economics*, 2013, *128* (2), 771–805. 6

Alesina, Alberto F and Richard T Holden, "Ambiguity and extremism in elections," Technical Report, National Bureau of Economic Research 2008. 7

Andonie, Costel and Daniel Diermeier, "Impressionable voters," *American Economic Journal: Microeconomics*, 2019, *11* (1), 79–104. 7

Andreottola, Giovanni, "Signaling valence in primary elections," *Games and Economic Behavior*, 2021, *126*, 1–32. 7

Battaglini, Marco, Rebecca B Morton, and Thomas R Palfrey, "The swing voter's curse in the laboratory," *The Review of Economic Studies*, 2010, *77* (1), 61–89. 26

Benjamin, Daniel J, "Errors in probabilistic reasoning and judgment biases," *Handbook of Behavioral Economics: Applications and Foundations 1*, 2019, *2*, 69–186. 26

Bonomi, Giampaolo, Nicola Gennaioli, and Guido Tabellini, "Identity, beliefs, and political conflict," *The Quarterly Journal of Economics*, 2021, *136* (4), 2371–2411. 7

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer, "Stereotypes," *The Quarterly Journal of Economics*, 2016, *131* (4), 1753–1794. 7

Breitmoser, Yves, "Strategic reasoning in p-beauty contests," *Games and Economic Behavior*, 2012, *75* (2), 555–569. 26

Brennan, Jason, "Why Do Politicians Say Stupid Things?," *Newsweek*, 2015. 4, 13

Camerer, Colin and Dan Lovallo, "Overconfidence and excess entry: An experimental approach," *American economic review*, 1999, *89* (1), 306–318. 7

Camerer, Colin F, Teck-Hua Ho, and Juin-Kuan Chong, "A cognitive hierarchy model of games," *The Quarterly Journal of Economics*, 2004, *119* (3), 861–898. 7, 26

Carnes, Nicholas and Noam Lupu, "The white working class and the 2016 election," *Perspectives on Politics*, 2021, *19* (1), 55–72. 4, 25

Crawford, Vincent P and Nagore Iriberri, "Level-k auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions?," *Econometrica*, 2007, *75* (6), 1721–1770. 26

Crawford, Vincent P, Miguel A Costa-Gomes, and Nagore Iriberri, "Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications," *Journal of Economic Literature*, 2013, *51* (1), 5–62. 26

Di Tella, Rafael and Julio J Rotemberg, "Populism and the return of the "paranoid style": Some evidence and a simple model of demand for incompetence as insurance against elite betrayal," *Journal of Comparative Economics*, 2018, *46* (4), 988–1005. 6

Djourelova, Milena, "Persuasion through Slanted Language: Evidence from the Media

Coverage of Immigration," *American Economic Review (forthcoming)*, 2022. 3

Eguia, Jon X and Francesco Giovannoni, "Tactical extremism," *American political science review*, 2019, *113* (1), 282–286. 7

Farrell, Joseph and Robert Gibbons, "Cheap talk with two audiences," *The American Economic Review*, 1989, *79* (5), 1214–1223. 6

Frankel, Alex and Navin Kartik, "Muddled information," *Journal of Political Economy*, 2019, *127* (4), 1739–1776. 6

Galasso, Alberto and Timothy S Simcoe, "CEO overconfidence and innovation," *Management Science*, 2011, *57* (8), 1469–1484. 7

Glaeser, Edward L, Giacomo AM Ponzetto, and Jesse M Shapiro, "Strategic extremism: Why Republicans and Democrats divide on religious values," *The Quarterly journal of economics*, 2005, *120* (4), 1283–1330. 7

Grossman, Gene M and Elhanan Helpman, "Electoral competition and special interest politics," *The review of economic studies*, 1996, *63* (2), 265–286. 7

Grossman, Sanford J, "The informational role of warranties and private disclosure about product quality," *The Journal of Law and Economics*, 1981, *24* (3), 461–483. 6

Hoorens, Vera, "Self-enhancement and superiority biases in social comparison," *European review of social psychology*, 1993, *4* (1), 113–139. 4

Jordan, Kayla N, Joanna Sterling, James W Pennebaker, and Ryan L Boyd, "Examining long-term trends in politics and culture through language of political leaders and cultural institutions," *Proceedings of the National Academy of Sciences*, 2019, *116* (9), 3476–3481. 2

Kartik, Navin and R Preston McAfee, "Signaling character in electoral competition," *American Economic Review*, 2007, *97* (3), 852–870. 7

Kartik, Navin, Francesco Squintani, Katrin Tinn et al., "Information revelation and pandering in elections," *Columbia University, New York*, 2015, *36.* 7

Koriyama, Yukio and Ali I Ozkes, "Inclusive cognitive hierarchy," *Journal of Economic Behavior & Organization*, 2021, *186*, 458–480. 26

Kraus, Karl, *Half-truths and one-and-a-half truths: Selected aphorisms*, University of Chicago Press, 1990. 23, 32

Kruger, Justin and David Dunning, "Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments.," *Journal of personality and social psychology*, 1999, *77* (6), 1121. 4

Lamont, Michèle, Bo Yun Park, and Elena Ayala-Hurtado, "Trump's electoral speeches and his appeal to the American white working class," *The British journal of sociology*, 2017, *68*, S153–S180. 25

Levy, Gilat and Ronny Razin, "When do simple policies win?," *Economic Theory*, 2012, *49* (3), 621–637. 7

Levy, Gilat, Ronny Razin, and Alwyn Young, "Misspecified politics and the recurrence of populism," *American Economic Review*, 2022, *112* (3), 928–62. 7, 26, 30

Malmendier, Ulrike and Timothy Taylor, "On the verges of overconfidence," *Journal of Economic Perspectives*, 2015, *29* (4), 3–8. 4, 7, 25

Moore, Don A and Paul J Healy, "The trouble with overconfidence.," *Psychological review*, 2008, *115* (2), 502. 25

Morelli, Massimo, Antonio Nicolò, and Paolo Roberti, "A commitment theory of populism," *Working Paper*, 2021. 7

Mudde, Cas, "The populist zeitgeist," *Government and opposition*, 2004, *39* (4), 541–563. 5

Nagel, Rosemarie, "Unraveling in guessing games: An experimental study," *The American economic review*, 1995, *85* (5), 1313–1326. 7, 26

Ortoleva, Pietro and Erik Snowberg, "Overconfidence in political behavior," *American Economic Review*, 2015, *105* (2), 504–35. 7

Rachman, Nathaniel, "The Simpleton Manifesto," *Persuasion*, 2020. 12

Spence, Michael, "Job market signaling," in "Uncertainty in economics," Elsevier, 1978, pp. 281–306. 6

Appendix

Table of Contents

A Proofs for the Bayesian Setting

Proof of Lemma 1

Take a politician (λ, ω) for which exposing valence is feasible, i.e. $\omega < 1 - \lambda$. We want to show that exposing valence (i.e., truthfully reporting $(l, v) = (\lambda, \omega)$) dominates all other ineffective speaking strategies that are feasible for her, namely pairs (l, v) with $v < \min\{\omega, 1 - l\}$.

Suppose per contra that exposing valence at some argument l' (not necessarily equal to λ) by displaying a valence level $v' < \omega$ was optimal. Then, since (l', ω) is feasible and $\tilde{f}((l', \omega)) \geq \omega$ by property $i)$ of the consistency requirement in Definition 3, it must be that $\tilde{f}((l', v')) \geq \omega > v'$. For this to be consistent –property $ii)$ of the consistency requirement in Definition 3 – there must be some politicians with valence $\omega' > \tilde{f}((l', v'))$ that choose (l', v'). But this is impossible, as such type could profitably deviate to (l', ω'), which is feasible for her and gives update $\tilde{f}((l', \omega')) \geq \omega' > \tilde{f}((l', v'))$. The first inequality uses again property i of the consistency requirement, which holds even off-equilibrium (i.e. does not require to assume that, absent the deviation, (l', ω') is played by some types). Hence, by a typical unraveling argument, all politicians who speak ineffectively use a speech of complexity that matches their valence. Therefore ineffective speaking induces a valence update that equals its complexity

$$(l, v) \in \mathcal{A}_{\neg ES} \implies \tilde{f}((l, v)) = v. \tag{20}$$

We are left to show that also the argument employed in a expose valence strategy coincides with the naive-preferred argument, that is $l = \lambda$. But this is immediate because we just showed that, among ineffective speaking strategies, the valence update (20) is independent of the argument l. Therefore, maximizing of (3) among ineffective speaking strategies requires to maximize the support π_N from naive voters, yielding $l = \lambda$.

Proof of Lemma 2

Following Lemma 1, expose valence strategies achieve complete support from naive voters but yield harsh update from sophisticated voters. Therefore, irrespectively of λ, candidates of valence v who expose valence get value

$$u_{EV}(v) = (1-k) + kv \tag{21}$$

The payoff from effective speaking at $(1-v, v)$ depends instead on the naive-preferred argument λ and is given by the following expression:

$$u_{ES}^{\lambda}(v) = (1-k)(1-(1-\lambda-v)^2) + kf(v) \tag{22}$$

now we obtain $\underline{v}(\lambda)$ by computing, the valence type who, when the preference type is λ, is indifferent between exposing and matching valence. That is, $\underline{v}(\lambda)$ solves $u_{EV}(v) = u_{ES}^{\lambda}(v)$, which can be written as

$$(1-k) + kv = (1-k)[1-(1-\lambda-v)^2] + kf(v) \tag{23}$$

rearranging,

$$\underline{v}(\lambda) = 1 - \lambda - \sqrt{\frac{k}{1-k}(f(\underline{v}(\lambda)) - \underline{v}(\lambda))} \tag{24}$$

so we have that:

$$u_{EV}(v) \geq u_{ES}^{\lambda}(v) \Leftrightarrow v \leq \underline{v}(\lambda) \tag{25}$$

Always following Lemma 1, from which we know that politicians choosing an expose valence strategy necessarily display $v = \omega$, we obtain that for each λ it is optimal to expose valence if and only if $\omega \leq \underline{v}(\lambda)$, whereas all types such that $\omega > \underline{v}(\lambda)$ will optimally choose effective speaking. Since $f(v) \geq v$, we have, as expected, that all types such that $\omega \geq 1-\lambda$ choose effective speaking. Moreover, as long as $f(v) > v$, there will be types such that $\omega < 1-\lambda$ also choosing effective speaking, despite the fact that they would be able to truthfully reveal their type by exposing valence.

Let us now focus on effective speaking. A politician doing effective speaking maximizes (22) choosing some v in the set $[0, \omega]$. This problem can either have a corner solution $v = \omega$, yielding the match valence strategy, or an interior solution, yielding the degrade valence strategy. This depends on whether the point satisfying the first order condition of (22) is feasible or not. This FOC reads:

$$2(1-k)(1-\lambda-v)+kf'(v)=0. \tag{26}$$

For each λ, we denote the solution to (26) as $\overline{v}(\lambda)$, yielding:

$$\overline{v}(\lambda)=\min\left\{1, 1-\lambda+\frac{k}{2(1-k)}f'(\overline{v}(\lambda))\right\} \tag{27}$$

which accounts for the possibility of $\overline{v}$ taking a corner value of 1. Notice that the problem of a politician choosing effective speaking (22) does not depend on ω other than through the valence constraint $v \leq \omega$. Therefore, for each λ, politicians with $\omega \leq \overline{v}(\lambda)$ are at a corner solution $v=\omega$, that is, they choose to match valence. For politicians with $\omega > \overline{v}(\lambda)$ effective speaking with $v=\overline{v}(\lambda)$ is feasible and therefore optimal: they choose to degrade valence not to lose too much support from naive voters.

Notice that our argument presumes that (26) characterizes a maximum of $u_{ES}^{\lambda}(v)$, i.e. we have taken for granted that the second-order condition of (22) is satisfied. In equilibrium, this is guaranteed by $v \geq v^* \geq \underline{V}_k$, where $\underline{V}_k$ is determined in the proof of Proposition 1 precisely from the SOC of $\max_v u_{ES}^{1}(v)$.

Proof of Lemma 3

Following Definition 3, given an effective speech $(1-v, v)$, sophisticated voters need to compute:

$$(\boldsymbol{l},\boldsymbol{v})^{-1}(1-v,v)=\{(\lambda,\omega):(\boldsymbol{l},\boldsymbol{v})(\lambda,\omega)=(1-v,v)\} \tag{28}$$

Using Lemma 2 and inverting the functions $\underline{v}(\lambda)$ and $\overline{v}(\lambda)$, for which we define:[31]

$$\overline{\lambda}(v):=\overline{v}^{-1}(v) \tag{29}$$

$$\underline{\lambda}(v):=\underline{v}^{-1}(v) \tag{30}$$

we obtain:

$$(\boldsymbol{l},\boldsymbol{v})^{-1}(1-v,v)=\left\{\lambda\in[\underline{\lambda}(v),\overline{\lambda}(v)]\cap\omega=v\right\}\cup\{\lambda=\overline{\lambda}(v)\cap\omega>v\} \tag{31}$$

[31]Notice that if there is no $\lambda\in[0,1]$ such that $\overline{v}(\lambda)=v$, then $\overline{\lambda}$ takes the corner value of 1. As far as $\underline{v}(\lambda)=v$ is concerned, there is always a solution $\lambda\in[0,1]$ for $v\geq v^*$.

Always following Lemma 2, we can write expressions for $\underline{\lambda}(v)$ and $\overline{\lambda}(v)$ as functions of v:

$$\underline{\lambda}(v) = 1 - v - \sqrt{\frac{k}{1-k}(f(v) - v)} \tag{32}$$

$$\overline{\lambda}(v) = \min\left\{1, 1 - v + \frac{k}{2(1-k)} f'(v)\right\} \tag{33}$$

From these expressions, we have that following effective speaking at a given v, the measure of types with valence equal to v is, as shown in Figure IV,

$$\phi_1 = \overline{\lambda}(v) - \underline{\lambda}(v), \tag{34}$$

which can be further written as:

$$\phi_1 = \min\left\{1, 1 - v + \frac{k}{2(1-k)} f'(v)\right\} - \left(1 - v - \sqrt{\frac{k}{1-k}(f(v) - v)}\right) \tag{35}$$

Similarly, when non-empty, the measure of types degrading valence ϕ_0, as shown again in Figure IV, is equal to $\phi_0 = 1 - v$. Finally, the expected valence of the politician conditional on degrading valence is $\mathbb{E}[\omega|\omega > v] = \frac{1+v}{2}$. Concerning the set of types degrading valence, in order for $\{\lambda = \overline{\lambda}(v) \cap \omega > v\}$ to be non-empty it must hold that $v \geq \bar{v}(1)$. If $v < \bar{v}(1)$, then $\phi_0 = 0$. We will discuss this issue in detail in the proof of Proposition 1, since this is crucially tied with the characterization of the value of v^* at which Poe's Law becomes non-trivial. Having derived the values of ϕ_0 and ϕ_1, we can construct the update conditional on effective speaking at v,

$$f(v) = v\phi_1(v) + \mathbb{E}[\omega|\omega > v]\phi_0(v) \tag{36}$$

which corresponds with condition (15).

Proof of Proposition 1

The proof proceeds as follows. We first derive (18), combining regions where ϕ_0 is zero and strictly positive. To prove (18) constitutes an equilibrium we need to check optimality and consistency. Consistency of expectations is guaranteed by Lemma 3. Optimality is guaranteed by Lemma 2, but we need to check that *i*) candidates for which Poe's Law is degenerate ($\phi_0 = 0$) also respect the equilibrium prescriptions, and *ii*) the hypothesis of Lemma 2 are verified, so that also candidates that choose speeches $\phi_0 > 0$ are optimizing.

Step 1: Deriving equation (18) for generic v^*. Building on Lemma 3, let us now first focus on the case in which $\phi_0 = 1 - v$. Using (35) to substitute the expressions for $\phi_1(v)$ in condition (36) — (15) in the main text — and solving for $f'(v)$, yields the differential equation (17):

$$f'(v) = 2\frac{1-k}{k}\left[\frac{(1-v)\left(\frac{1+v}{2}-f(v)\right)}{f(v)-v} - \sqrt{\frac{k}{1-k}(f(v)-v)}\right] \tag{37}$$

Therefore, all we are left to do to find expectations $f(v)$ is solve the differential equation (17). Before we can do that, however, we need an initial value. To this end, we define

$$v^* := \overline{v}(1) \tag{38}$$

the valence degrading strategy chosen by politicians with the most extreme naive electorate $\lambda = 1$. Clearly, $\overline{v}(\cdot)$ depends on $f'(v)$ and hence on $f(v)$. Using the expression for $\overline{v}$, i.e., (27) yields:

$$v^* = \frac{k}{2(1-k)}f'(v^*) \tag{39}$$

which gives the initial value $f'(v^*) = \frac{2(1-k)}{k}v^*$ from which the differential equation can be solved forward. By standard arguments, the IVP admits a unique solution for all k, v^*, which we denote with $\hat{f}_{v^*,k}(v)$, a self-map on $[v^*, 1]$ that characterizes the equilibrium above v^*, i.e. where $\phi_0 > 0$ in (36). If $v < v^*$ we have, following Lemma 3, that $\phi_0 = 0$, and (36) gives $f(v) = v$.

Step 2: Determining the set $V_k = [\underline{V}_k, \overline{V}_k]$ of feasible v^*. We now turn to characterizing the set V_k, ruling out v^* that induce policies are inconsistent with optimality of some candidate's types. Both the upper and the lower bound on feasible v^* rule out deviations from candidates with valence above $b(\lambda)$, i.e. for which exposing valence is not feasible.[32] Those candidates must choose valence matching if their valence is below v^* and must choose to degrade valence at the $\overline{v}$ obtained from the first-order condition (26). The former requirement gives the upper bound, the latter the lower bound.

Upper bound $\overline{V}_k$ To have $f(v) = v$ for all $v \leq v^*$, it must be that no politician uses speeches $\{ES(v)\}_{v \leq v^*}$ as part of a valence-degrading strategy. Indeed, if (a positive measure of) candidates with valence above v chose $ES(v)$, then in equilibrium it must be $f(v) > v$. Even if $f(v) = v$, i.e. absent any updating premium, the presence of sophisticated voters gives a reason to relatively valent politicians that face a high-λ naive crowd to sacrifice some naive support and pursue valence matching. Hence, we obtain $\overline{V}_k$

[32] Candidates (λ, ω) with $\omega < \min\{b(\lambda), v^*\}$ obviously optimize choosing to expose valence: there is no feasible updating premium for them, so no motive to distort the argument from λ. For candidates with valence above v^* but below $b(\lambda)$, it is sufficient to check that $\underline{v}$ is decreasing.

by finding the effective speech that a candidate with $\lambda = 1$ would select if i) her valence constraint was inactive and ii) sophisticated voters updated $f(v) = v$. If v^* was above $\overline{V}_k$, then candidates with $\lambda = 1$ and $\omega \in [\overline{V}_k, v^*]$ would prefer $ES(\overline{V}_k)$ to $ES(\omega)$, breaking the equilibrium. Substituting $\lambda = 1$ and $f(v) = v$ in (22) we obtain

$$\overline{V}_k = \arg\max_v (1-k) \cdot (1-v^2) + kv = \frac{k}{2(1-k)},$$

from which we get that $\overline{V}_k$ is 0 for $k = 0$ and increasing in k, reaching 1 for $k = \frac{2}{3}$.

In sum, if v^* is below $\overline{V}_k$, then the effective speeches that should not be chosen for valence degrading ($\phi_0 = 0$) are not attractive and confirm harsh updating.

Lower bound $\underline{V}_k$ The lower bound $\underline{V}_k$ comes from the requirement that the first order condition (26) characterizes a maximum of $u^{\lambda}_{ES}(v)$. Differentiating (26) we obtain the second order condition for a maximum:

$$f''(v) \leq \frac{2(1-k)}{k} \tag{40}$$

Condition (40) must hold for all v that are used as part of a valence-degrade strategy that is for $v \geq v^*$. Intuitively, if $f''(v)$ evaluated at $\overline{v}(\lambda)$ is too large, then the objective $u^{\lambda}_{ES}(v)$ is (locally) convex and (26) characterizes a minimum rather than a maximum.

Differentiating (17) we can express $f''(v)$ in $[v^*, 1]$ as a second-order ODE (that we do not report because it is analytically cumbersome) which allows to check whether (40) holds.[33] We can evaluate f'' at v^*, taking $f'(v^*)$ from (16) and $f(v^*)$ from Poe's Law (15). This gives that $f''(v^*)$ satisfies (40) if and only if $v^* \geq \underline{V}_k$, where $\underline{V}_k$ has no immediate analytical expression, but is characterized in a Mathematica file, available upon request, and computed numerically as the red line plotted in the left panel of Figure IX. Also numerically, we check that when $v^* \geq \underline{V}_k$ then the second-order condition holds also for all $v \geq v^*$,[34] and that $\underline{V}_k$ is increasing in k and it reaches to 1 at $k = \frac{2}{3}$.

In sum, the intuition for why there is a lower bound on v^* is as follows. By (16), $f'(v^*)$ is directly proportional to v^*: Reducing v^* requires a decrease in the updating premium resulting from a local upward deviation, which depresses $f'(v^*)$. In turns, through (17), this pushes $f''(v^*)$ above the upper bound $\frac{2(1-k)}{k}$ implied by the second order condition of $\max u^1_{ES}(v)$.

Figure IX above gives a graphical representation of the set $V_k = [\underline{V}_k, \overline{V}_k]$ of feasible

[33] In the linear branch, $f'' = 0$ so the objective is strictly increasing over the feasible actions, establishing that valence matching is optimal in the triangle between the bound and v^*.

[34] In particular, in Section C we analytically show that the SOC is satisfied with equality, along all equilibria, at $v = 1$.

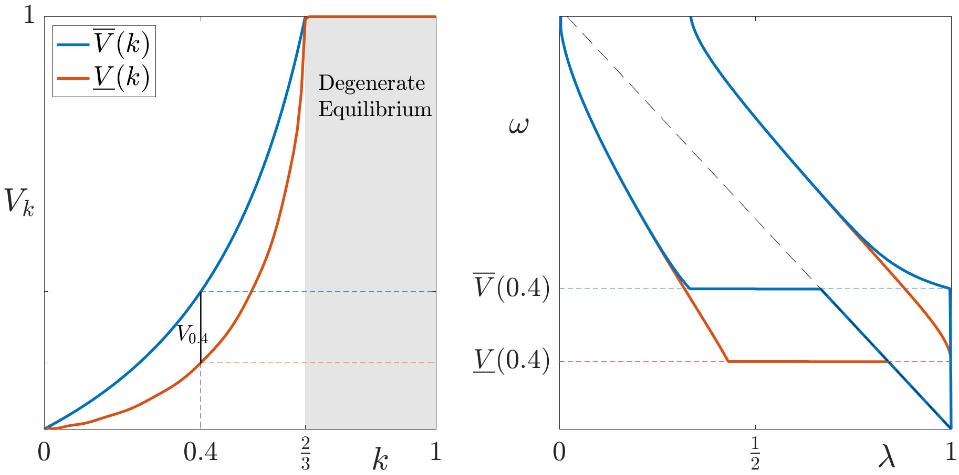


Figure IX: *Left panel:* The higher and lower bounds of V_k as function of k. *Right panel:* equilibria for fixed $k = 0.4$ with highest and lowest feasible $v^\star$

v^*, and show the equilibrium policies in the two extrema equilibria for $k = 0.4$.[35]

By Step 1, (18) gives an expectation which is consistent with the functions $\overline{v}(\lambda)$ and $\underline{v}(\lambda)$ it induces; by Step 2, if $v^* \in V_k$ then all politicians above the bound (both those with valence above v^* and those with valence below v^*) optimally follow the equilibrium prescriptions. To close the proof we need to show that the properties of $\overline{v}(\lambda)$ and $\underline{v}(\lambda)$ assumed in Lemma 2 are verified.

Step 3: Check that $\overline{v}(\lambda) \geq \underline{v}(\lambda)$, and both are decreasing. To see that $\overline{v}(\lambda) \geq \underline{v}(\lambda)$, take the differenct between expressions (24) and (27). This yields

$$\frac{k}{2(1-k)} f'(v) \geq -\sqrt{\frac{k}{1-k}(f(v) - v)} \tag{41}$$

Using (17) we can see that (41) is always satisfied, and in fact it holds strictly for all $v \in (v^*, 1)$.

$\overline{v}(\lambda)$ *strictly decreasing:* Recall that:

$$\overline{v}(\lambda) = 1 - \lambda + \frac{k}{2(1-k)} f'(\overline{v}(\lambda)) \tag{42}$$

Totally differentiating the above yields:

$$\frac{d\overline{v}}{d\lambda} = \left(f''(\overline{v}) \frac{k}{2(1-k)} - 1 \right)^{-1} \tag{43}$$

[35] Recall from the text that those equilibria also correspond to outcomes of the fixed point procedure initialized, respectively, at maximally harsh and maximally benevolent updating $f(v) = v, f(v) = \frac{v+1}{2}$.

and from this we obtain:

$$\frac{d\overline{v}}{d\lambda} < 0 \iff f''(\overline{v}) < \frac{2(1-k)}{k} \forall \lambda \tag{44}$$

which corresponds to the second-order condition (40), implied at all v by $v^* > \underline{V}_k$.[36]

$\underline{v}(\lambda)$ *strictly decreasing (for* $v \neq v^*$*):* For $v < v^*$, the result is obvious since $\underline{v}(\lambda) = 1 - \lambda$. For $v > v^*$, it is convenient to first consider the inverse of $\underline{v}(\lambda)$, denoted by $\underline{\lambda}(v)$, for which we can establish that for each v, the value of $\underline{\lambda}(v) := \underline{v}^{-1}(v)$ is unique. This can be immediately concluded from the condition:

$$\underline{\lambda}(v) = 1 - v - \sqrt{\frac{k}{1-k}(f(v) - v)} \tag{45}$$

Going back to $\underline{v}(\lambda)$, we can show that increasing λ, the maximum (if not unique) value that solves (24) cannot increase.[37] The reason is that as λ increases and the displayed valence v is kept fixed, the utility from exposing valence is constant, whereas the utility from effective speaking at $(1 - v, v)$ increases, since the cost in terms of naive voters' support is smaller. Therefore, a politician that was indifferent between exposing and matching valence at some λ will strictly prefer to match her valence following a small increase in λ. Hence, whenever it is defined, $\underline{v}$ is strictly decreasing in λ, concluding the proof.

Proof of Proposition 2

Proof of point i): First of all, it is easy to see that $\bar{s}(0) = 0$. This happens since politicians choose $l = \lambda$ and follows immediately from 4. Moreover, given that $V_{k=0} = \{0\}$, the correspondence $\bar{s}(k)$ is single-valued at $k = 0$.

At the opposite extreme of k, we have that for all $k \geq \frac{2}{3}$, $\bar{s}(k)$ is equal to $-\frac{1}{6}$. This follows from the discussion done in Section 2.2 and Section 2.3 about the case of all sophisticated voters and the fact that for $k \geq \frac{2}{3}$, $V_k = \{1\}$, $f(v) = v$ and politicians optimally choose to display $v = \omega$ and $l = \min\{\lambda, 1 - \omega\}$, leading to $\mathbb{E}[l|\lambda] = \frac{1}{3}$ and $\bar{s}(k) = \frac{1}{3} - \mathbb{E}[\lambda] = \frac{1}{3} - \frac{1}{2} = -\frac{1}{6}$.

We now prove the result for values of $k \in \left(0, \frac{2}{3}\right)$. First of all, notice that when choosing to report valence v, a politician doing effective speaking chooses location $1 - v$. Hence,

[36] Indeed, the two are the same condition, and we could derive the lower bound $\underline{V}_k$ from the requirement that $\overline{v}$ is decreasing.

[37] Notice that $\underline{v}(\lambda)$ being well defined as a function is not necessary for our characterization of the equilibrium in terms of a partition, neither for Poe's Law.

what we want to show is the following:

$$\mathbb{E}[1-v|MV \cup DV] < \mathbb{E}[\lambda|MV \cup DV] \tag{46}$$

where MV and DV respectively denote *matching valence* and *degrading valence*. This can be rearranged to:

$$\mathbb{E}[v|MV \cup DV] > \mathbb{E}[1-\lambda|MV \cup DV] \tag{47}$$

Let's start computing the right-hand side of (47). Notice that in the following we will work with the bliss-point λ and the reported valence v rather than then valence type ω, since we are interested in finding the expected argument chosen by politicians as a function of their reported valence. Denoting by $\mu(MV)$ and $\mu(DV)$ the measure of the type space in which politicians do effective speaking and degrading valence respectively, we have:

$$\mathbb{E}[1-\lambda|MV \cup DV] = \frac{\int_0^1 \int_{\underline{\lambda}}^1 (1-\lambda) d\lambda dv}{\mu(MV)+\mu(DV)} \tag{48}$$

Solving the integrals on the numerator yields:

$$\int_0^1 \int_{\underline{\lambda}}^1 (1-\lambda) d\lambda dv = \int_0^1 (1-\underline{\lambda}-\frac{1}{2}\lambda^2|_{\underline{\lambda}}^1) dv = \int_0^1 (\frac{1}{2}-\underline{\lambda}+\frac{1}{2}\underline{\lambda}^2) dv = \frac{1}{2}\int_0^1 (1-\underline{\lambda})^2 dv \tag{49}$$

Finally, notice that this can be split in the following way, exploiting the fact that $\underline{\lambda}(v) = 1-v$ for $v < v^*$.

$$\mathbb{E}[1-\lambda|MV \cup DV](\mu(MV)+\mu(DV)) = \frac{1}{2}\int_0^{v^*} v^2 dv + \frac{1}{2}\int_{v^*}^1 (1-\underline{\lambda}(v))^2 dv \tag{50}$$

Coming to the expected reported valence (i.e., the left-hand side of (47)), we need to use the density of each reported valence v, which is equal to:

$$\mu(v) = \bar{\lambda}(v) - \underline{\lambda}(v) + \mathbb{I}_{v \geq v^*}(1-v) \tag{51}$$

noticing further that in order to derive conditional expectations, we use the same measure that we used for the right-hand side, that is $\mu(MV)+\mu(DV)$. As a matter of fact, notice that the following holds:

$$\int_0^1 \mu(v) dv = \int_0^1 \int_{\underline{\lambda}}^1 d\lambda dv \tag{52}$$

Now, the expected reported valence can be written in the following way:

$$\mathbb{E}[v|MV \cup DV](\mu(MV)+\mu(DV)) = \int_0^1 v\mu(v) dv \tag{53}$$

and going further:

$$\mathbb{E}[v|MV\cup DV](\mu(MV)+\mu(DV)) = \int_0^{v^*} v(1-(1-v))dv + \int_{v^*}^1 v(\bar{\lambda}(v)-\underline{\lambda}(v)+(1-v))dv \quad (54)$$

The condition we want to prove therefore boils down to the following (cancelling out the measures $\mu(MV)+\mu(DV)$ from the denominators):

$$\int_0^{v^*} v(1-(1-v))dv+\int_{v^*}^1 v(\bar{\lambda}(v)-\underline{\lambda}(v)+(1-v))dv > \frac{1}{2}\int_0^{v^*} v^2dv+\frac{1}{2}\int_{v^*}^1(1-\underline{\lambda}(v))^2dv \quad (55)$$

which can be rearranged to:

$$\int_0^{v^*} v^2dv + 2\int_{v^*}^1 v(1-v)dv > \int_{v^*}^1[(1-\underline{\lambda})^2 - 2v[\bar{\lambda}-\underline{\lambda}]]dv \quad (56)$$

Now, notice that

$$((1-\underline{\lambda})-v)^2 = \frac{k}{1-k}(f(v)-v) \quad (57)$$

We rewrite the first condition as:

$$\int_0^{v^*} v^2dv > \int_{v^*}^1\{(1-\underline{\lambda})^2 - 2v[\bar{\lambda}-\underline{\lambda}] - 2v(1-v)\}dv \quad (58)$$

which can be rewritten as:

$$\int_0^{v^*} v^2dv > \int_{v^*}^1\{(1-\underline{\lambda})^2 + 2v\underline{\lambda} - 2v + v^2 + v^2 - 2v\bar{\lambda}\}dv \quad (59)$$

and further as:

$$\int_0^{v^*} v^2dv > \int_{v^*}^1\{\frac{k}{1-k}(f(v)-v) + v^2 - 2v\bar{\lambda}\}dv \quad (60)$$

Now, using the definition of $\bar{\lambda} = 1 - v + \frac{k}{2(1-k)}f'(v)$, we can rewrite $2v\bar{\lambda}$ which yields:

$$\int_0^{v^*} v^2dv > \int_{v^*}^1\{\frac{k}{1-k}(f(v)-v) + v^2 - 2v + 2v^2 - \frac{k}{1-k}vf'(v)\}dv \quad (61)$$

Finally, we can integrate by parts $vf'(v)$, yielding:

$$\int_0^{v^*} v^2dv > \int_{v^*}^1\{\frac{k}{1-k}(f(v)-v) + 3v^2 - 2v + \frac{k}{1-k}f(v)\}dv - \frac{k}{1-k}vf(v)|_{v^*}^1 \quad (62)$$

We can now solve all the integrals not involving $f(v)$, writing:

$$\frac{1}{3}v^* > v^3|_{v^*}^1 - v^2|_{v^*}^1 - \frac{k}{2(1-k)}v^2|_{v^*}^1 - \frac{k}{1-k} + \frac{k}{1-k}v^* f(v^*) + 2\frac{k}{1-k}\int_{v^*}^1 f(v)dv \tag{63}$$

yielding:

$$\frac{4}{3}(v^*)^3 - (v^*)^2 > -\frac{3}{2}\frac{k}{1-k} + \frac{k}{2(1-k)}(v^*)^2 + \frac{k}{1-k}v^* f(v^*) + 2\frac{k}{1-k}\int_{v^*}^1 f(v)dv \tag{64}$$

and finally:

$$\frac{3}{2}\frac{k}{1-k} + \frac{4}{3}(v^*)^3 - (v^*)^2\left(1 + \frac{k}{2(1-k)}\right) > \frac{k}{1-k}v^* f(v^*) + 2\frac{k}{1-k}\int_{v^*}^1 f(v)dv \tag{65}$$

Now, we can observe that the right-hand side is always smaller than if $f(v) = \frac{1+v}{2}$ for all v. Hence, we can try and see if the condition holds when evaluating the right-hand side at its upper bound. Notice that this can be done since at $k = 0$, where we know the condition holds as an equality, using the upper bound still makes the condition hold.

Making the substitutions yields:

$$\frac{3}{2}\frac{k}{1-k} + \frac{4}{3}(v^*)^3 - (v^*)^2\left(1 + \frac{k}{2(1-k)}\right) > \frac{k}{1-k}\frac{1+v^*}{2}v^* + 2\frac{k}{1-k}\left(\frac{1-v^*}{2} + \frac{1}{4}v^2|_{v^*}^1\right) \tag{66}$$

which can be further simplified to:

$$\frac{3}{2}\frac{k}{1-k} + \frac{4}{3}(v^*)^3 - (v^*)^2(1 + \frac{k}{2(1-k)}) > \frac{k}{1-k}\frac{1+v^*}{2}v^* + 2\frac{k}{1-k}(\frac{1-v^*}{2} + \frac{1}{4}(1-(v^*)^2)) \tag{67}$$

With some rearrangement, this gives:

$$\frac{k}{2(1-k)}v^* + \frac{4}{3}(v^*)^3 - (v^*)^2\frac{2-k}{2(1-k)} > 0 \tag{68}$$

Taking common factors, this finally yields:

$$\frac{v^*}{6(1-k)}(3k + 8(1-k)(v^*)^2 - 3(2-k)v^*) > 0 \tag{69}$$

Notice that the only negative term is the last one, and in order to be problematic that requires a large v^* and a low k, which is somehow a contradiction. Notice further that if $k \geq 0.4$, this condition is satisfied for all v^*. Moreover, solving this quadratic equation yields that the condition is satisfied as long as v^* is low enough (we can see this since the condition is a parabola with the vertex facing down, and that the second solution is not relevant since it is always larger than our upper bound v^*, for the relevant values of k. In particular, it can be checked that one solution is always larger than $\frac{1}{2}$, which is larger than the upper bound for all $k \in [0, 0.4]$.). Hence, if we can show that the condition is satisfied at the upper bound of v^*, it will be satisfied for any v^*. Substituting the value of the upper bound on v^*, which is $\frac{k}{2(1-k)}$, into the condition, yields the following to be satisfied:

$$3k + 8(1-k)\left(\frac{k}{2(1-k)}\right)^2 - 3(2-k)\frac{k}{2(1-k)} > 0 \tag{70}$$

With some algebra this can be shown to yield:

$$\frac{k^2}{2(1-k)} > 0 \tag{71}$$

Proof of Point ii): Consider effective speaking at a location l such that $l < 1 - v^\star$ and therefore $f(v)$ satisfies the differential equation. For convenience, consider $v = 1 - l$ so our desideratum is:

$$\overbrace{1-v}^{l} > \overbrace{\bar{\lambda}(v)}^{\lambda \text{ that DV at } v} \frac{1-v}{1-v+\left(\bar{\lambda}(v)-\underline{\lambda}(v)\right)} + \underbrace{\frac{\bar{\lambda}(v)+\underline{\lambda}(v)}{2}}_{\text{Midpoint of MV}} \frac{\left(\bar{\lambda}(v)-\underline{\lambda}(v)\right)}{1-v+\left(\bar{\lambda}(v)-\underline{\lambda}(v)\right)} \tag{72}$$

Rearranging this we obtain:

$$(1-v)\left[(1-v-\underline{\lambda}(v))\right] < \frac{\bar{\lambda}(v)+\underline{\lambda}(v)}{2}\left(\bar{\lambda}(v)-\underline{\lambda}(v)\right) \tag{73}$$

and further:

$$\frac{\left(\bar{\lambda}(v)+\underline{\lambda}(v)\right)}{2} < \frac{(1-v)-\underline{\lambda}(v)}{\bar{\lambda}(v)-\underline{\lambda}(v)}(1-v) \tag{74}$$

Now, using the expressions for $\bar{\lambda}$ and $\underline{\lambda}$, which, recall, read:

$$\underline{\lambda}(v) = 1 - v - \sqrt{\frac{k}{1-k}\left(f(v)-v\right)} \tag{75}$$

$$\bar{\lambda}(v) = 1 - v + \frac{kf'(v)}{2(1-k)} \tag{76}$$

we obtain, for the left-hand side and right-hand side of (74) respectively:

$$\frac{\left(\overline{\lambda}\left(v\right)+\underline{\lambda}\left(v\right)\right)}{2}=\frac{1-v+\frac{kf'(v)}{2(1-k)}+1-v-\sqrt{\frac{k}{1-k}\left(f\left(v\right)-v\right)}}{2}$$
$$=\left(1-v\right)+\frac{\frac{kf'(v)}{2(1-k)}-\sqrt{\frac{k}{1-k}\left(f\left(v\right)-v\right)}}{2} \tag{77}$$

and

$$\frac{\left(1-v\right)-\underline{\lambda}\left(v\right)}{\overline{\lambda}\left(v\right)-\underline{\lambda}\left(v\right)}\left(1-v\right)=\frac{\sqrt{\frac{k}{1-k}\left(f\left(v\right)-v\right)}}{\frac{kf'(v)}{2(1-k)}+\sqrt{\frac{k}{1-k}\left(f\left(v\right)-v\right)}}\left(1-v\right) \tag{78}$$

Substituting these into (74) yields:

$$\left(1-v\right)+\frac{\frac{kf'(v)}{2(1-k)}-\sqrt{\frac{k}{1-k}\left(f\left(v\right)-v\right)}}{2}<\frac{\sqrt{\frac{k}{1-k}\left(f\left(v\right)-v\right)}}{\frac{kf'(v)}{2(1-k)}+\sqrt{\frac{k}{1-k}\left(f\left(v\right)-v\right)}}\left(1-v\right) \tag{79}$$

which can be rearranged to:

$$4\left(1-k\right)\left[f'\left(v\right)\left(1-v\right)-\left(f\left(v\right)-v\right)\right]+kf'\left(v\right)^{2}<0 \tag{80}$$

We now want to show that:

$$4\left(1-k\right)\left[f'\left(v^{\star}\right)\left(1-v^{\star}\right)-\left(f\left(v^{\star}\right)-v^{\star}\right)\right]+kf'\left(v^{\star}\right)^{2}<0 \tag{81}$$

Recall that

$$f'\left(v^{\star}\right)=\frac{2\left(1-k\right)}{k}$$

so:

$$4\left(1-k\right)\left[\frac{2\left(1-k\right)}{k}v^{\star}\left(1-v^{\star}\right)-\left(f\left(v^{\star}\right)-v^{\star}\right)\right]+k\left[\frac{2\left(1-k\right)}{k}v^{\star}\right]^{2}<0 \tag{82}$$

With some algebra, the former condition can be rearranged to:

$$\frac{\left(f\left(v^{\star}\right)-v^{\star}\right)}{v^{\star}\left(1-\frac{1}{2}v^{\star}\right)}>\frac{2\left(1-k\right)}{k} \tag{83}$$

Using (17) and (16) to find a closed-form expression for $f(v^{*})$ – which we do not report here – we obtain that condition (83) is satisfied for v^{*} sufficiently small. Finally – using the upper bound $\frac{2(1-k)}{k}$ on the second order condition $f''(v^{*})$ that pins down $\underline{V}_{k}$ – we prove that for k small, the threshold implied by condition (83) is greater than $\underline{V}_{k}$. This means that for small k we can always find a feasible v^* such that

$$\mathbb{E}\left[\lambda\left|\left(l\left(\lambda,\omega\right),v\left(\lambda,\omega\right)\right)=\left(l,1-l\right)\right]<l\right.$$

hence closing the proof.

B Proofs and Additional Results: Dismissal Game

We first characterize and discuss the candidate's best response to dismissive updating, then prove the properties of dismissive equilibria that are discussed in the main text, Proposition 3. To make the comparison with the BNE more immediate and avoid heavy notation, we keep the same notation $\underline{v},\overline{v},f,\dots$ in the game with dismissal. Only when explicitly comparing expressions related to the dismissal game with their Bayesian counterparts, we adopt the subscript *dis*. Recall that dismissive agents always update according to $f(v)=\frac{1+v}{2}$, which readily gives the politician's best response.

Proposition 4. *In the game with dismissal, the behavior of politicians follows the structure of Lemma 2, with $\overline{v}$ and $\underline{v}$ decreasing functions of λ given by:*

$$\underline{v}(\lambda)=\max\left\{0,1-\lambda-\frac{k}{4(1-k)}-\sqrt{\frac{k}{2(1-k)}\left(\frac{k}{8(1-k)}+\lambda\right)}\right\} \tag{84}$$

$$\overline{v}(\lambda)=\min\left\{1,1-\lambda+\frac{k}{4(1-k)}\right\} \tag{85}$$

Proof. An argument analogous to that in Lemma 1 establishes that even under dismissal ineffective speeches are valence exposing. Along the lines of Lemma 2 (which only requires that f is a strictly increasing function generating strictly decreasing loci), a candidate communicating to a dismissive audience compares the value of exposing valence $u_{EV}(v)=(1-k)+kv$ with the optimal effective speaking strategy. Substituting $f(v)=\frac{1+v}{2}$ in the expression for the returns on effective speaking (10) gives that, under dismissal

$$u_{ES}^{\lambda}(v)=(1-k)(1-((1-v)-\lambda)^2)+k\frac{1+v}{2}. \tag{86}$$

Imposing $u_{ES}^{\lambda}(v)=u_{EV}(v)$, and censoring the solution at zero, yields that $\underline{v}(\lambda)$ satisfies (84). Maximizing (86) gives that valence-degrading strategies satisfy the first-order condition (26) with $f'(v)=\frac{1}{2}$. Rearranging, we obtain that $\overline{v}(\lambda)$ satisfies (85).

By direct inspection of (84)-(85), $\overline{v}(\lambda)$ and $\underline{v}(\lambda)$ are – unless degenerate – strictly decreasing in λ, validating the construction of Lemma 2.

□

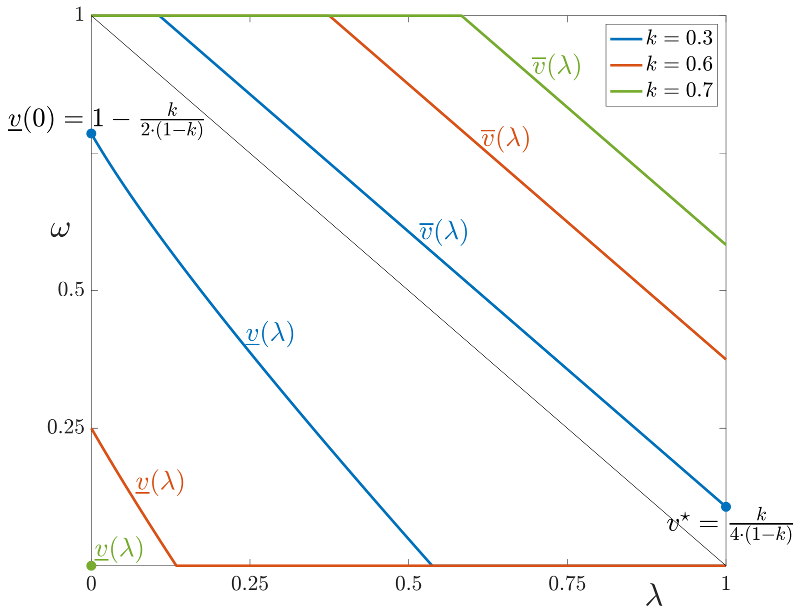
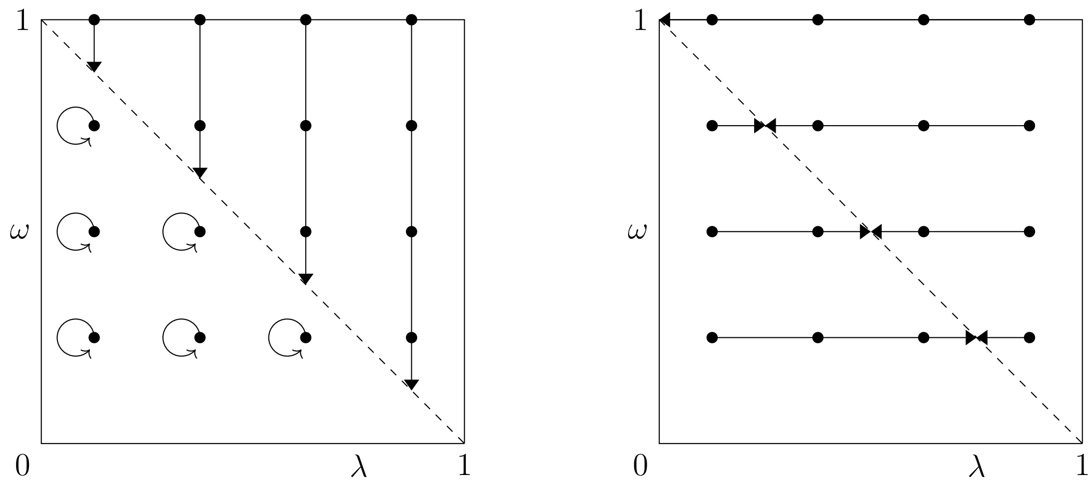


Figure X: Partition of the (λ, ω) space into *expose valence*, *match valence* and *degrade valence* regions, for three different values of k.

Figure X displays a typical resulting partition. For a given λ, low valence politicians, i.e. those with $\omega < \underline{v}(\lambda)$, still expose their valence: despite the high premium for effective speaking, the argument that matches their (poor) valence would alienate too many naive voters. For intermediate valence $\underline{v}(\lambda) \leq \omega \leq \overline{v}(\lambda)$, politicians pick the argument that matches their valence. High valence politicians with $\omega > \overline{v}(\lambda)$ degrade their valence, speaking effectively with complexity $\overline{v}(\lambda)$. Notice that $\overline{v}(1)$ is the most simplistic argument that is used in a valence degrading strategy. For this reason, we denote $v^* := \overline{v}(1)$ in order to create a direct parallel with the Bayes Nash game. This also helps highlighting an important difference between the two games: in the Bayes Nash game, v^* marked the point below which effective speaking is interpreted harshly. In the dismissal game, instead, expectations are always benevolent – even following the most simplistic speeches – despite the existence of an equivalent for v^*.

We now contrast the outcomes of the dismissal game with the Bayesian benchmark. If $k = 0$, the two games coincide: with no sophisticated agents who aim to infer the valence of the politician, expectations do not play a role and the outcome pictured in Figure XIa is the same as that of Figure IIa. Moving on to interior values of k, consider first politicians choosing between exposing and matching valence. Since $f^{dis}(v) > f(v)$ for any equilibrium f and $v < 1$, then $\underline{v}_{dis}(\lambda) < \underline{v}(\lambda)$: any politician that exposes valence in the game with dismissal also expose valence in the BNE. As k increases, the difference between the two games is more and more apparent. The benefits of effective speaking grow larger,



(a) Location-complexity choices for $k = 0$.

(b) Location-complexity choices for $k \geq \frac{4}{5}$.

Figure XI: Policies in the polar cases of the game with dismissal. Points represent types, arrows represent policies.

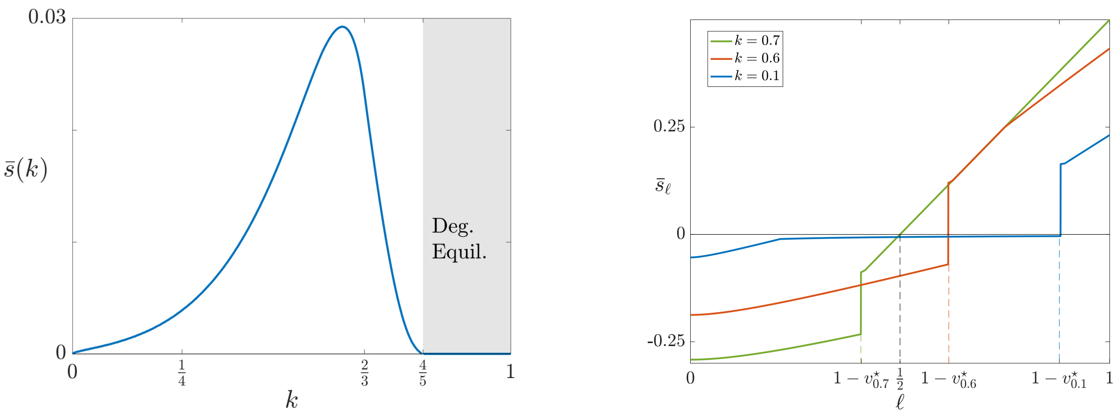
since the increase in the weight on the updating premium is not paralleled by equilibrium adjustments in the expectations depressing its value. As a result, for large k no candidate exposes valence: from (84) it is immediate to see that $\underline{v}(0) = 0$ if $k \geq \frac{2}{3}$. Notice the contrast with the Bayes Nash Equilibrium: in that case, when $k \geq \frac{2}{3}$ all politicians who are not constrained by their naive-preferred argument choose to expose their valence and the effective speaking premium disappears precisely because politicians want to exploit it.

Consider now the problem of (relatively) high valence politicians who have to choose an effective speech. Plugging the dismissive updating rule in the returns from effective speaking $u^{\lambda}_{ES}(v)$ we obtain expression (85) for the valence-degrading locus $\overline{v}(\lambda)$. Although, contrary to $\underline{v}$, there is no immediate comparison with the BNE counterpart,[38] $\overline{v}$ is still increasing in k (strictly if non-trivial): a higher share of sophisticated voters makes it cheaper to effectively defend complex arguments and push valent candidates further away from the naive-preferred argument. It is easy to check that $v^{\star} = \overline{v}(1) = \min\left\{\frac{k}{4(1-k)}, 1\right\}$ reaches 1 as $k = \frac{4}{5}$: At very high levels of sophistication, even candidates with $\lambda = 1$ want to display as much valence as possible (using arguments close to 0). For $k \geq \frac{4}{5}$, all politicians choose valence matching (see Fig. XIb) and $\bar{s} = 0$ because the simplification caused by politicians below the diagonal is perfectly offset by the complexification of politicians above the diagonal. If the share of dismissive voters is large, politicians choose an argument equal on average (though almost never point-wise) to the naive-preferred

[38]In particular, it is not true that the types choosing to match their valence under dismissal would always follow the same strategy in a Bayesian equilibrium. This is because, contrary to f, it is not possible to rank f' in the dismissive and rational model.

one.[39] This is again in sharp contrast to what happens in the Bayes Nash equilibrium, where $\bar{s}(k) < 0$ for all $k > 0$ and it was minimal at large values of k.

With the candidate's policies (84)-(85) in closed form, it is straightforward to compute the aggregate moments and establish the comparative statics presented in the main text, Proposition 3. Before proving them, we show in Figure XIIa a summary of these results. The left panel (replicated in Figure VIII in the main text) represent the average simplism as we vary the share of sophisticated (but dismissive) voters: *i*) always positive and *ii*) hump-shaped. The right panel instead focuses on the simplism contributed by each argument in three dismissive equilibria: *iii*) the simpler the argument, the more simplism it contributes, and *iv*) only arguments above $1 - v^\star(k)$ contribute positive simplism. The positive average simplism is therefore driven by the simplism catalyzed by extreme arguments. We now formally establish the results.



(a) Average net simplism of the political speech $\bar{s}$ as a function of k.

(b) Argument-specific simplism $\bar{s}_l$ for low, medium and high k

Figure XII: Global and local measures of simplism in the game with dismissal.

Proof of Proposition 3

Proof of parts *i*) and *ii*) We now start with the proof of the first two statements. Recall that

$$\bar{s}_l = l - \mathbb{E}[\lambda | (l, v)(\lambda, \omega) = (l, 1 - l)] \tag{87}$$

Since the expectation conditional on effective speaking requires to know the type λ of those that use it as a part of an effective speaking strategy, the algebraic manipulations require to invert the $\underline{v}, \overline{v}$ functions presented in the text. For this reason, it is convenient

[39] The fact that the average simplism of political communication when all politicians pursue effective speaking equals zero is a consequence of the symmetric distribution of types above and below the bound, i.e., $\mathbb{E}[\lambda] = \mathbb{E}[b(\omega)]$. By focusing on this balanced benchmark we are implicitly setting supply-driven simplism to zero. We do so in order to isolate the amount of simplism generated by the strategic rhetorical choices of politicians.

to work in a reported valence space v. Hence we denote $\tilde{s}_v = \bar{s}_{1-v}$ be the average simplification of those that do effective speaking at valence level v, i.e., at location $1 - v$. The proposition would equivalently read that $\tilde{s}_v$ is decreasing in v, and that it is negative if and only if $v < \min\{\frac{1}{2}, \overline{v}(1)\}$. The proof is conceptually straightforward but a bit involved since, depending on whether the (inverse) $\underline{v}, \overline{v}$ functions hit the bounds we have to distinguish 4 cases. We show that $\tilde{s}$ is locally decreasing in v in each of these cases (Lemma 4), and that whenever there there is a jump, the function $\tilde{s}_v$ jumps downwards (Lemma 6).

First, we derive expressions for $\bar{\lambda}(v)$ and $\underline{\lambda}(v)$, inverting the conditions in (84):

$$\bar{\lambda}(v) = \min\left\{1, 1 - v + \frac{k}{4(1-k)}\right\},$$

and

$$\underline{\lambda}(v) = \max\left\{0, 1 - v - \sqrt{2}\sqrt{\frac{k}{4(1-k)}}\sqrt{1-v}\right\},$$

from which we obtain

$$\tilde{s}_v = \begin{cases} (1-v) - \left[\bar{\lambda}(v)\frac{1-v}{1-v+(\bar{\lambda}(v)-\underline{\lambda}(v))} + \frac{\underline{\lambda}(v)+\bar{\lambda}(v)}{2}\frac{(\bar{\lambda}(v)-\underline{\lambda}(v))}{1-v+(\bar{\lambda}(v)-\underline{\lambda}(v))}\right] & v \geq \overline{v}(1) \\ (1-v) - \frac{\underline{\lambda}(v)+1}{2} & v < \overline{v}(1) \end{cases} \tag{88}$$

Lemma 4. *Given the expression for $\tilde{s}_v$ given by (88) and the possibility of corner solutions for $\bar{\lambda}(v)$ and $\underline{\lambda}(v)$, there are 4 cases to consider. Within each of these cases, $\tilde{s}_v$ is decreasing in v.*

Proof. **Case 1: $\underline{\lambda}(v) = 0$ and $\bar{\lambda}(v) = 1$.** The expression for $\tilde{s}_v$ becomes:

$$\tilde{s}_v = (1-v) - \frac{1}{2} = \frac{1}{2} - v, \tag{89}$$

decreasing in v.

Case 2: Both $\underline{\lambda}(v)$ and $\bar{\lambda}(v)$ interior. The expression for $\tilde{s}_v$ becomes:

$$\tilde{s}_v = \frac{k^2}{8(1-k)\left(k(3-4v) + 2\sqrt{2}\sqrt{k(1-k)(1-v)} + 4v - 4\right)} \tag{90}$$

Differentiating with respect to v yields:

$$-\frac{k^2\left(\frac{\sqrt{2}\sqrt{1-k}k}{\sqrt{k(1-v)}} + 4(1-k)\right)}{8(1-k)\left(k(3-4v) + 2\sqrt{2}\sqrt{k-1}\sqrt{k(v-1)} + 4v - 4\right)^2} < 0 \tag{91}$$

Case 3: $\underline{\lambda}(v) = 0$, $\bar{\lambda}(v)$ interior. The expression for $\tilde{s}_v$ becomes:

$$\tilde{s}_v = (1-v) - \left(\bar{\lambda}(v)\frac{1-v}{1-v+\bar{\lambda}(v)} + \frac{\bar{\lambda}^2}{2(1-v+\bar{\lambda}(v))})\right) \tag{92}$$

Differentiating with respect to v yields:

$$\frac{1}{4}\left(\frac{k^2}{(8-7k-8(1-k)v)^2} - 1\right) \tag{93}$$

It is immediate to see that (93) is (strictly) increasing in v and evaluates to 0 at $v = 1$. Hence, the expression is negative for $v < 1$.

Case 4: $\bar{\lambda}(v) = 1$, $\underline{\lambda}(v)$ interior. The expression for $\tilde{s}_v$ becomes:

$$\tilde{s}_v = (1-v) - \frac{1+\underline{\lambda}(v)}{2} \tag{94}$$

which can be written as:

$$\frac{1}{4}\left(\frac{\sqrt{2}\sqrt{k(1-v)}}{\sqrt{1-k}} - 2v\right) \tag{95}$$

Differentiating with respect to v yields:

$$\frac{1}{8}\left(\frac{\sqrt{2}k}{\sqrt{k(1-k)(1-v)}} - 4\right) \tag{96}$$

This can be rearranged as:

$$-\frac{1}{2} + \frac{1}{4}\sqrt{\frac{k}{2(1-k)}\frac{1}{1-v}} = -\frac{1}{2} + \frac{1}{4}\sqrt{\frac{1-\underline{v}(0)}{1-v}} \tag{97}$$

which is always negative for $v < \underline{v}(0)$.

□

Having shown that $\tilde{s}_v$ is continuous and decreasing in v within of the four cases, in order to complete the proof we need to establish that at the transition between cases is either continuous or features a downward jump in $\tilde{s}_v$. To do so, we analyze the sequence of scenarios as v changes, as a function of k.

Lemma 5. *Depending on k the following sequence of cases determine $\tilde{s}_v$ as v increases*

1	*for* $k \geq \frac{4}{5}$	
$1 \to 3$	*for* $k \in \left[\frac{2}{3}, \frac{4}{5}\right)$	*switching at* $v = \bar{v}(1)$
$4 \to 1 \to 3$	*for* $k \in \left(\frac{4}{7}, \frac{2}{3}\right)$	*switching at* $v = \underline{v}(0)$ *and* $v = \bar{v}(1)$
$4 \to 2 \to 3$	*for* $k \leq \frac{4}{7}$	*switching at* $v = \bar{v}(1)$ *and* $v = \underline{v}(0)$

Proof. In the following, we present the four possible scenarios which depend on the value of k.

a. $k \geq \frac{4}{5}$ In this case, $\underline{\lambda} = 0$ for all v and $\bar{\lambda} = 1$ for all v. Hence, we are in the trivial case 1 for all values of v.

b. $k \in \left[\frac{2}{3}, \frac{4}{5}\right)$ In this interval, $\underline{\lambda}$ is never interior (since $k \geq \frac{2}{3}$), whereas $\bar{\lambda}$ is interior for $v > \bar{v}(1)$. Hence, as v increases the scenario switches from scenario 1 to scenario 3, with the switch taking place at $\bar{v}(1)$.

c. $k \in \left(\frac{4}{7}, \frac{2}{3}\right)$ Notice that as long as $k > \frac{4}{7}$, we have that $\bar{v}(1) > \underline{v}(0)$. To see this, recall that $\bar{v}(1) = \frac{k}{4(1-k)}$, whereas $\underline{v}(0) = 1 - \frac{k}{2(1-k)}$. Comparing the two, we get that:

$$\bar{v}(1) > \underline{v}(0) \iff \frac{k}{4(1-k)} > 1 - \frac{k}{2(1-k)} \iff \frac{3}{4}\frac{k}{1-k} > 1 \iff k > \frac{4}{7} \quad (98)$$

Since $\bar{v}(1) > \underline{v}(0)$, increasing v the transition is now between case 4, case 1 and case 3, with the two switches taking place at $\underline{v}(0)$ and $\bar{v}(1)$ respectively.

d. $k \leq \frac{4}{7}$ For these values of k, we have that $\bar{v}(1) \leq \underline{v}(0)$, as we have shown in the discussion of point c. Therefore, the sequence of cases is now case 4, case 2 and finally case 3.

□

Lemma 6 (Proof of Point iv)). *When switching from one case to another, $\tilde{s}_v$ is either continuous, or it features a downward jump. Moreover,*

$$\tilde{s}_v > 0 \iff v < \min\left\{\frac{1}{2}, \bar{v}(1)\right\}$$

Proof. We provide the proof separately for each sequence of cases proved possible in Lemma 5.

a. In this case, there are no switching points, and $\tilde{s}_v = \frac{1}{2} - v$, is positive if and only if $v < \frac{1}{2}$.

b. In this case, the switch from case 1 to case 3 takes place at $v = \bar{v}(1) = \frac{k}{4(1-k)}$. Because $k > \frac{2}{3}$, we have that $\bar{v}(1) > \frac{1}{2}$. Hence, $\tilde{s}_v$ turns from positive to negative at $v = \frac{1}{2}$, that is before the switch. Finally, notice that when switching from case 1 to case 3, $\tilde{s}_v$ has an downward jump, since it switches from

$$(1 - v) - \bar{\lambda}|_{=1}\frac{1}{2}$$

to

$$(1-v) - \bar{\lambda}|_{=1}\left[\frac{1}{2}+\frac{\Theta}{2}\right]$$

where $\Theta \equiv \frac{1-v}{1-v+\bar{\lambda}|_{=1}}$ is the weight assigned to $\bar{\lambda}$ in the expression for $\tilde{s}_v$. These conditions are derived from (92) evaluated at $v = \bar{v}(1)$. Hence, $\tilde{s}_v$ is decreasing and positive if and only if $v < \frac{1}{2}$.

c. In this scenario, the first switch is from case 4 to case 1, and there is then a further switch from case 1 to case 3. Consider the latter switch first: following the same reasoning as in the previous point, the switch entails an downward jump in $\tilde{s}_v$. Moreover, since the switch happens for $v < \frac{1}{2}$, we have that $\tilde{s}_v > 0$ before the switch. We therefore need to show that $\tilde{s}_v$ is negative after the switch. Evaluating condition (92) at $v = \frac{k}{4(1-k)}$ yields:

$$\frac{17k^2 - 24k + 8}{36k^2 - 68k + 32} \tag{99}$$

which is always negative for $k \in ((\frac{4}{7}, \frac{2}{3})$.

Concerning the first switch, which takes place at $v = \underline{v}(0)$, notice that by construction the value of $\tilde{s}_v$ is continuous at the switch, since the switch takes place at the point where, in case 4, $\underline{\lambda} = 0$, just like in case 1. This means that $\tilde{s}_v$ is positive at the switch (since $\underline{v}(0) < \frac{1}{2}$ for $k > \frac{4}{7}$), and, as we have already proved, decreasing both before and after the switch.

d. Finally, in this case we start the analysis from the first switch. This takes place at $\bar{v}(1)$ and it entails a discontinuity, in particular an downward jump in $\bar{s}_v$. As a matter of fact, defining by $\Theta_d \equiv \frac{1-\bar{v}(1)}{1-\bar{v}(1)+1-\underline{\lambda}(\bar{v}(1))}$, we have that across the threshold $1-\bar{v}(1)$ the value of $\tilde{s}_v$ goes from:

$$(1-v) - \frac{1-\underline{\lambda}}{2}$$

to

$$(1-v) - \left[\frac{1-\underline{\lambda}}{2} + \Theta_d\left(\frac{1}{2} - \frac{\underline{\lambda}}{2}\right)\right],$$

which is smaller since $\underline{\lambda}(\bar{v}(1)) < 1$. We need to prove that the value of $\tilde{s}_v$ is positive before the switch and negative afterwards. To do this, we first evaluate $\tilde{s}_v$ in case 4, that is expression (95) evaluating it at $\bar{v}(1) = \frac{k}{4(1-k)}$. It can be easily shown that its value is positive for all $k \in [0, \frac{4}{7}]$:

$$\frac{1}{8}\left[1 - \frac{1}{1-k}\left(1 - \sqrt{2k(4-5k)}\right)\right] > 0 \iff k < \frac{8}{11} \tag{100}$$

and $\frac{8}{11} > \frac{4}{7}$. To show that the value of $\tilde{s}_v$ is negative after the switch, that is at $\overline{v}(1)$ we evaluate the expression for case 2, that is (90), at $\overline{v}(1)$. This yields:

$$\frac{k^2}{8(1-k)\left(-4(1-k)+\sqrt{2k(4-5k)}\right)} \tag{101}$$

which can be shown to always be negative, as desired.

Finally, the second switch is from case 2 to case 3, and it takes place at $\underline{v}(0)$. This switch entails no discontinuity in $\tilde{s}_v$, since $\underline{\lambda}(v) = 0$ at $\underline{v}(0)$ in both cases. Moreover, since $\tilde{s}_v$ is case-wise decreasing and we showed that $\tilde{s}_v$ is negative at $\overline{v}(1) < \underline{v}(0)$, then the switch must occur at a negative value of $\tilde{s}_v$.

To sum up, we showed that also in this case $\tilde{s}_v$ is decreasing in v, and that it changes sign at $v = \overline{v}(1)$.

□

Proof of Point ***iii*****).** If $k = 0$, $l = \lambda$ for all politician types (λ, ω), hence $\bar{s}_l(0) = l - \mathbb{E}[\lambda|l] = l - l = 0$ for all l and $\bar{s}(0) = 0$.

Similarly, if $k \geq \frac{4}{5}$, $l = 1 - \omega$ for all (λ, ω). It follows that $\mathbb{E}[\lambda|l] = \frac{1}{2}$, and hence $\bar{s}_l = l - \frac{1}{2} = \frac{1}{2} - \omega$, and integrating across ω we obtain $\bar{s} = 0$.

Furthermore, we can see that $\bar{s}(\frac{2}{3}) > 0$. This happens since at $k = \frac{2}{3}$, all types such that $\omega < 1 - \lambda$ choose to match their valence. This happens since $\underline{v}(\lambda)$ is now constrained at 0 for all λ. At the same time, however, when $k = \frac{2}{3}$ there are values of λ such that $\overline{v}(\lambda) < 1$. This means that some types degrade their valence. As a result, not all the simplism caused by types $\omega < 1 - \lambda$ is offset by the complexification of types $\omega > 1 - \lambda$, leading to $\bar{s} > 0$.

For $k \in [\frac{2}{3}, \frac{4}{5})$, the value of $\bar{s}(k)$ is decreasing. This is due to the fact that whereas all types $\omega < 1 - \lambda$ choose to match their valence (defending a more simplistic argument than λ), as k increases $\overline{v}(\lambda)$ increases, meaning that some politicians switch from valence degrading to valence matching (and that those who still degrade valence choose a more complex argument than with lower values of k). Therefore, $\bar{s}(k)$ decreases when $k \in [\frac{2}{3}, \frac{4}{5})$, reaching 0 at $k = \frac{4}{5}$.

To see this argument more formally we can write out the integral for the case $k \geq \frac{2}{3}$:

$$\int_0^1 \left[\int_0^{\overline{v}(\lambda)} (1-\omega) d\omega + \int_{\overline{v}(\lambda)}^1 (1 - \overline{v}(\lambda)) d\omega \right] d\lambda \tag{102}$$

which can be rearranged to:

$$\frac{1}{2}+\int_0^1\int_{\overline{v}(\lambda)}^1(\omega-\overline{v}(\lambda))d\omega d\lambda>\frac{1}{2} \tag{103}$$

To see that the derivative is decreasing in k, we need to apply Leibniz Rule:

$$\int_0^1\left[(1-\overline{v}(\lambda))\frac{\partial 1}{\partial k}-(\overline{v}(\lambda)-\overline{v})\lambda)\frac{\partial\overline{v}(\lambda,k)}{\partial k}+\int_{\overline{v}(\lambda)}^1(-\frac{\partial\overline{v}(\lambda,k)}{\partial k})\right]d\lambda<0 \tag{104}$$

Finally, for $0<k<\frac{2}{3}$, we can use the closed-form for $\bar{s}(k)$ to show that it is first increasing in k and then decreasing, reaching the maximum for k strictly below $\frac{2}{3}$.

C Proofs of Additional Results (Bayes Nash Game)

Limit of $f'(v)$ at 1

Proof. Take the definition of the expectations:

$$f(v)=v\frac{\bar{\lambda}(v)-\underline{\lambda}(v)}{\bar{\lambda}(v)-\underline{\lambda}(v)+(1-v)\mathbb{I}_{v\geq v^*}}+\frac{1+v}{2}\frac{(1-v)\mathbb{I}_{v\geq v^*}}{\bar{\lambda}(v)-\underline{\lambda}(v)+(1-v)\mathbb{I}_{v\geq v^*}} \tag{105}$$

Denote now $\frac{\bar{\lambda}(v)-\underline{\lambda}(v)}{\bar{\lambda}(v)-\underline{\lambda}(v)+(1-v)\mathbb{I}_{v\geq v^*}}:=\beta(v)$ With this notation, the expression for $f(v)$ can be rewritten as:

$$f(v)=v\beta(v)+\frac{1+v}{2}(1-\beta(v)) \tag{106}$$

Differentiating:

$$f'(v)=\beta'(v)v+\beta(v)+\frac{1}{2}(1-\beta(v))+\frac{1+v}{2}(-\beta'(v)) \tag{107}$$

Taking the limit as v goes to 1, and noting that:

$$\lim_{v\to 1}f'(v)=\lim_{\to 1}\beta(v)\frac{1}{2}+\frac{1}{2} \tag{108}$$

Using the definition of β, we have that this limit is:

$$\lim_{v\to 1}\beta(v)=\lim_{v\to 1}\frac{\bar{\lambda}(v)-\underline{\lambda}(v)}{\bar{\lambda}(v)-\underline{\lambda}(v)+(1-v)\mathbb{I}_{v\geq v^*}}=\lim_{v\to 1}\frac{\bar{\lambda}(v)-\underline{\lambda}(v)}{\bar{\lambda}(v)-\underline{\lambda}(v)}=1 \tag{109}$$

Hence we have that:

$$\lim_{v\to 1}f'(v)=1 \tag{110}$$

Also, notice that $\beta(v)$ has to converge to 1 as v goes to 1. The reason is that when

$v = 1$, the politician having valence $v = 1$ is indeed the only possible explanation. In other words, $\beta(1) = 1$ by construction. So unless $\beta(v)$ is discontinuous at 1, we have that the limit of $\beta(v)$ at one is one. $\square$

Derivative of $\underline{\lambda}(v)$ at 1

Proof. Start from the expression for $\underline{\lambda}$:

$$\underline{\lambda}(v) = 1 - v - \sqrt{\frac{k}{1-k}(f(v) - v)} \tag{111}$$

We want to show that, in the (λ, v) plot, $\underline{\lambda}$ gets infinitely steep as λ approaches 0. In other words, we want to show that:

$$\frac{\partial \underline{\lambda}(v)}{\partial v}|_{v=1} \approx 0 \tag{112}$$

Differentiating $\underline{\lambda}(v)$ we obtain:

$$\frac{\partial \underline{\lambda}(v)}{\partial v} = -1 - \sqrt{\frac{k}{1-k}} \frac{f'(v) - 1}{2\sqrt{f(v) - v}} \tag{113}$$

Taking the limit as $v \to 1$, we can see that

$$\lim_{v \to 1} \frac{f'(v) - 1}{2\sqrt{f(v) - v}} = \frac{0}{0} \tag{114}$$

Reverse engineering the result we want to have, we want to show that:

$$\lim_{v \to 1} \frac{f'(v) - 1}{\sqrt{f(v) - v}} = -2\sqrt{\frac{1-k}{k}} \tag{115}$$

Using the De l'Hopital rule, we obtain:

$$\lim_{v \to 1} \frac{f'(v) - 1}{\sqrt{f(v) - v}} = \lim_{v \to 1} \frac{f''(v)}{\frac{1}{2}\frac{f'(v)-1}{\sqrt{f(v)-v}}} = \frac{\lim_{v \to 1} 2f''(v)}{\lim_{v \to 1} \frac{f'(v)-1}{\sqrt{f(v)-v}}} \tag{116}$$

which gives us, noticing that $f'(v)$ must be approaching 1 from below, that is $f'(v) - v < 0$ in a neighborhood of 1:

$$\lim_{v \to 1} \frac{f'(v) - 1}{\sqrt{f(v) - v}} = -\sqrt{\lim_{v \to 1} 2f''(v)} = -2\sqrt{\frac{(1-k)}{k}} \tag{117}$$

which is precisely what we need.

□

SOC of (10) $= 0$ **at** $v = 1$

Proof. We can show analytically that the second-order condition is exactly met at $v = 1$. Take condition (17) and use De L'Hopital rule (twice) at $v = 1$. We get:

$$\begin{aligned}\lim_{v\to 1} f'(v) &= \lim_{v\to 1} 2\frac{1-k}{k}\left[\frac{-1\left(\frac{1+v}{2} - f(v)\right) + (1-v)\left(\frac{1}{2} - f'(v)\right)}{f'(v) - 1}\right] \\ &= 2\frac{1-k}{k}\lim_{v\to 1}\frac{(-1)\left(\frac{1}{2} - f'(v)\right) - (1-v)f''(v) - 1\left(\frac{1}{2} - f'(v)\right)}{f''(v)}\end{aligned}$$

Using $f'(v) = 1$ (see Appendix C), we finally get:

$$1 = 2\frac{1-k}{k}\frac{1}{f''(1)} \Leftrightarrow f''(1) = 2\frac{1-k}{k} \tag{118}$$

□